

Multimodal Digital Health Interventions versus Standard Perioperative Management in Lung Cancer Surgery: A Systematic Review, Meta-Analysis, and Evidence Gap Map

Huanzhi Peng

Army Medical University

Shengliang Zhao

Army Medical University

Lili Tang

Army Medical University

Kai Wang

Army Medical University

Mengyang Chang

Army Medical University

Quanxing Liu

quanxing9999@tmmu.edu.cn

Army Medical University

Research Article

Keywords: Lung cancer, Digital health, Perioperative care, Supportive care, Systematic review, Meta-analysis

Posted Date: August 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10143256/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Purpose: Multimodal digital health interventions (DHIs) may extend perioperative monitoring, rehabilitation, symptom support, and care coordination for patients undergoing lung cancer surgery, but their comparative effectiveness has not been systematically mapped. This review evaluated multimodal DHIs versus conventional perioperative management across clinical, patient-reported, and implementation outcomes.

Methods: Ten English and Chinese databases, trial registries, grey literature sources, conference proceedings, and reference lists were searched from inception to June 2026. Eligible studies were randomized or nonrandomized controlled studies of adults undergoing lung cancer surgery that used at least two digital modality categories. Risk of bias was assessed using RoB 2.0 and ROBINS-I logic. Compatible randomized trials were synthesized using random-effects meta-analysis; nonrandomized evidence was summarized narratively using SWiM. Certainty was rated with GRADE.

Results: Eleven controlled studies including 2,088 participants were included. Randomized-trial syntheses favored multimodal DHIs for pulmonary function (3 RCTs; $n = 345$; SMD 0.705, 95% CI 0.137 to 1.272) and health-related quality of life (4 RCTs; $n = 438$; SMD 0.647, 95% CI 0.276 to 1.018), but certainty was very low and heterogeneity was substantial. Length of stay was directionally shorter but inconclusive. Anxiety, depression, and prolonged hospitalization favored DHIs in single-study analyses; postoperative pulmonary complications and pain could not be pooled.

Conclusion: Multimodal DHIs show promising but very low-certainty evidence as adjuncts to perioperative lung cancer supportive care. Future trials should standardize outcome reporting, directly compare DHIs with multidisciplinary management, and evaluate implementation, equity, and cost outcomes.

PROSPERO registration: CRD420261417858.

1. INTRODUCTION

1.1 Background

Lung cancer remains the leading cause of cancer-related mortality worldwide, accounting for approximately 1.8 million deaths annually [1]. Surgical resection offers the best prospect of cure for patients with early-stage non-small cell lung cancer (NSCLC) [2]. The perioperative period is characterized by substantial clinical challenges: postoperative pulmonary complications (PPCs) occur in 19%–59% of patients [3], pulmonary function recovery is protracted, and patients experience significant symptom burden—pain, anxiety, depression, and fatigue—that impairs health-related quality of life (HRQoL) [4, 5]. Enhanced recovery after surgery (ERAS) protocols demand multidimensional, individualized, and resource-efficient perioperative care strategies [6].

Traditional multidisciplinary team management (MM)—integrating thoracic surgeons, anesthesiologists, respiratory therapists, specialist nurses, dietitians, and psychologists—is the standard of care [6]. Despite well-established benefits, MM is constrained by substantial human resource demands (16.67 ± 6.41 min per patient per encounter) [7], inequitable geographic accessibility, fragmented information transfer, and inadequate post-discharge continuity [4].

1.2 Multimodal Digital Health as a Theoretical Framework

Multimodal digital health interventions (DHIs)—defined as interventions combining two or more distinct digital modality categories from the following: (a) wearable devices for continuous physiological monitoring, (b) mobile applications for rehabilitation guidance and symptom tracking, (c) remote monitoring or telerehabilitation platforms, (d) virtual/augmented reality, (e) artificial intelligence (AI)-based decision support, and (f) electronic patient-reported outcome (ePRO) systems with automated clinical alerts—have emerged as promising alternatives to traditional care models [8]. The COVID-19 pandemic accelerated digital health adoption globally [9], and China’s national “Internet + Healthcare” policy has created an enabling environment for early multimodal DHI adoption.

From an implementation science perspective, multimodal DHIs rely on synergy effects across digital components: a single digital tool addresses only one clinical function (e.g., SMS reminders, telephone coaching, or standalone wearables), whereas integrated modalities—combining monitoring, education, clinical alerts, and rehabilitation—form a closed-loop perioperative care pathway. This theoretical framework underpins the operational definition of multimodal interventions (≥ 2 distinct digital modality categories) adopted in this review, distinguishing synergistic multi-component platforms from interventions that merely digitize a single care function.

1.3 Current Evidence and Research Gaps

Individual trials have demonstrated promising efficacy. A non-inferiority RCT demonstrated that digital therapeutics (DTx)-assisted management achieved comparable forced expiratory volume in 1 second (FEV₁%) recovery to MM while reducing management time by 91.1% [7]. The POPPER trial demonstrated that a smartphone application integrating symptom reporting, exercise guidance, and education significantly improved pulmonary function recovery ratio following video-assisted thoracoscopic surgery (VATS) lobectomy [10]. A DTx-based cardio-oncology rehabilitation trial showed significant improvements in peak oxygen uptake (VO_{2peak}) and HRQoL in NSCLC survivors [11]. The TD-WELLBEING multicenter RCT reported that ePRO-based collaborative symptom management significantly reduced symptom threshold events and improved emotional function through 12 months [12, 13].

Lu et al. [14] published the first systematic review and meta-analysis of digital pulmonary rehabilitation in lung cancer surgery (11 studies, 5 RCTs + 6 quasi-experimental), reporting a 45% reduction in PPC risk. However, their review did not restrict to multimodal interventions, did not compare DHIs with formal MM, and did not examine healthcare workforce efficiency or the evidence landscape systematically. He et al. [15] conducted a meta-analysis of smartphone app-based interventions across cancer types and found

improved HRQoL and self-efficacy, but did not focus on the perioperative period or the multimodal design question.

Notwithstanding these advances, four critical evidence gaps persist. First, no review has quantitatively compared multimodal DHIs against MM, the clinical question most relevant to perioperative supportive-care systems. Second, prior reviews often mixed single- and multi-component digital interventions, limiting inference about synergistic platforms. Third, perioperative oncology outcomes have not been mapped across hard endpoints, patient-reported outcomes, and implementation outcomes. Fourth, Chinese and international evidence has not been integrated into a unified evidence gap map.

1.4 Study Aims

This systematic review, meta-analysis, and evidence gap map addresses four questions: (1) What is the pooled effectiveness of multimodal DHIs across clinical endpoints, patient-reported outcomes, and healthcare resource utilization? (2) Do comparator type, digital modality count, and surgical approach moderate intervention effectiveness? (3) Do multimodal DHIs show comparable clinical effectiveness to MM while reducing management burden? (4) Where are the most important evidence gaps for future trials?

2. METHODS

2.1 Protocol and Registration

This review was conducted in accordance with PRISMA 2020 [16] (Online Resource 1) and registered with PROSPERO (CRD420261417858; amendment v2.0). The amendment expanded the comparator framework to three tiers, added Scopus and IEEE Xplore, and incorporated evidence gap mapping. A targeted supplementary search was integrated into the same PRISMA flow, eligibility assessment, extraction table, and statistical analysis rather than reported as a separate review.

2.2 Eligibility Criteria

Population

Adults (≥ 18 years) undergoing lung cancer surgery (lobectomy, sublobar resection, or pneumonectomy) via any approach (VATS, robot-assisted thoracic surgery [RATS], or thoracotomy), regardless of histological type or tumor stage.

Intervention: Multimodal DHIs incorporating ≥ 2 distinct digital modality categories from: (a) wearable devices (activity trackers, smartwatches, pulse oximeters, ECG patches); (b) mobile applications for rehabilitation guidance, symptom tracking, or health education; (c) remote monitoring or telerehabilitation platforms with bidirectional data flow; (d) virtual or augmented reality; (e) AI-based clinical decision support or personalized exercise prescriptions; (f) ePRO systems with active symptom

reporting and automated clinical alerts. The ≥ 2 threshold distinguishes synergistic multi-component interventions from single-function digital tools.

Comparator — three-tier stratified framework: C1 (primary): traditional MM, defined as formal coordinated care involving ≥ 3 disciplines; C2 (secondary): conventional perioperative management without formal MM (usual care, standard ERAS, or center-based pulmonary rehabilitation); C3 (exploratory): single-modality digital interventions (for qualitative benchmarking where data available).

Outcomes: Tier 1 (hard endpoints): PPCs (Clavien-Dindo grade $\geq$ II within 30 days), pulmonary function recovery ($FEV_1\%$ predicted or FVC at 3–4 weeks post-surgery), hospital length of stay (LOS, days). Tier 2 (PROs): pain intensity (VAS/NRS, standardized to 0–10), anxiety (HADS-A/GAD-7), depression (HADS-D/PHQ-9), HRQoL (EORTC QLQ-C30/WHOQOL-BREF). Tier 3 (implementation): healthcare personnel management time (min/patient), treatment adherence (%).

Study design: Parallel-group RCTs (any blinding status) and quasi-experimental studies with a clearly defined concurrent or historical comparator group (non-randomized controlled trials, controlled before-after studies, prospective comparative cohort studies). Minimum sample size: ≥ 10 participants per arm. English and Chinese language publications.

Exclusion criteria

(1) Non-lung cancer or mixed cohorts without extractable lung cancer subgroup data; (2) non-perioperative interventions; (3) single-modality digital interventions; (4) comparator receiving an active digital intervention (head-to-head digital comparison); (5) case reports, reviews, protocols, single-arm studies without comparators; (6) conference abstracts with insufficient data; (7) pediatric populations; (8) non-English/Chinese publications.

2.3 Search Strategy

Ten bibliographic databases were searched from inception to June 2026: PubMed, EMBASE, Cochrane CENTRAL, Web of Science, CINAHL, Scopus, IEEE Xplore, CNKI, Wanfang Data, and CBM. Strategies combined MeSH/Emtree terms and free-text keywords across lung cancer surgery, perioperative care, digital health, multimodal/multi-component interventions, and controlled-study design. The full strategies and supplementary search audit trail are provided in Online Resource 2.

Additionally, ClinicalTrials.gov, ChiCTR, EU Clinical Trials Register, ANZCTR, and WHO ICTRP were searched; grey literature was screened through manual browsing of conference proceedings (ATS, ESTS, IASLC, ASCO, ESMO, ICCN; 2023–2026); and reference lists of included studies and key reviews [14, 15] were hand-searched. No language restrictions were applied during the search phase.

2.4 Study Selection and Data Extraction

Search results were imported into reference management software for automatic deduplication (DOI → PMID → title + year) supplemented by manual verification. Two reviewers independently screened titles/abstracts, then full texts against eligibility criteria. Disagreements were resolved by discussion or third-reviewer consultation. Cohen's κ was calculated at both screening stages.

A standardized, piloted extraction form was used for independent dual extraction of study identification, population, intervention, comparator, outcome data, and risk-of-bias information. For continuous outcomes reported as median and interquartile range, Wan et al. methods [17] were used only for sensitivity analysis when original mean $\pm$ SD data were unavailable. Missing data were pursued by contacting corresponding authors up to two times.

2.5 Risk of Bias Assessment

Risk of bias was independently assessed by two reviewers. RCTs were assessed using Cochrane RoB 2.0 [18] across five domains; quasi-experimental studies were assessed using ROBINS-I [19] across seven domains. Risk-of-bias data structures were prepared using robvis-compatible coding and finalized as custom ggplot2/grid visualizations for readability. Subjective PROs (pain, anxiety, HRQoL) were interpreted with particular caution under open-label designs—an inherent limitation of digital health trials where blinding participants to receiving a digital device or application is infeasible. Objective outcomes (spirometry, LOS, mortality) were considered less susceptible to performance bias. This differential interpretation was reflected in GRADE ratings and sensitivity analyses.

2.6 Statistical Analysis

Effect measures and pooling: For dichotomous outcomes, risk ratios (RR) with 95% CI were calculated when extractable events were available. For continuous outcomes, standardized mean differences (SMD) with Hedges' g correction were applied when different instruments assessed the same construct; mean differences (MD) were used when instruments were identical. Cohen's conventions were used for SMD interpretation (0.2 = small, 0.5 = medium, 0.8 = large). A random-effects model with restricted maximum likelihood (REML) estimator for τ^2 was used for pooled analyses, given the clinical heterogeneity inherent in diverse digital modalities, surgical contexts, and healthcare settings. Meta-analysis was conducted only when at least two source-verifiable studies reported compatible outcome data; $k = 2$ syntheses were explicitly interpreted as exploratory. When fewer than two compatible studies were available, narrative synthesis following the Synthesis Without Meta-analysis (SWiM) guideline [20] was performed.

Heterogeneity

Cochran's Q test ($\alpha = 0.10$), the I^2 statistic with 95% confidence intervals, and τ^2 were reported. No rigid I^2 thresholds were applied; clinical interpretation of τ^2 magnitude alongside I^2 was emphasized. 95% prediction intervals were reported for pooled estimates with $k \geq 3$ to indicate the range of true effects expected in future studies.

Subgroup and sensitivity analyses: Pre-specified subgroup analyses examined comparator type (MM vs. non-MM), number of digital modalities (≥ 3 vs. 2), surgical approach, and study design. Because only RCTs were quantitatively pooled and nonrandomized studies were narratively synthesized, study design was handled by stratification rather than statistical interaction testing. Given the small number of compatible studies per outcome, formal interaction testing was not performed; subgroup analyses are descriptive and hypothesis-generating only. Feasible sensitivity analyses included leave-one-out diagnostics when k permitted, τ^2 estimator comparison when pooling was statistically meaningful, fixed-effect comparison, and a pulmonary-function sensitivity model adding Lv 2025 through median/IQR conversion. PPC sensitivity restricted to Clavien-Dindo grade $\geq$ II was not performed because no included study had extractable Clavien-Dindo PPC counts.

Publication bias: For outcomes with $k \geq 10$, Egger's regression test; for $5 \leq k < 10$, Doi plot with Luis Furuya-Kanamori (LFK) index [21] supplemented by funnel plot visual inspection; for $k < 5$, visual inspection only with explicit acknowledgment of insufficient statistical power. Trim-and-fill was applied when funnel plot asymmetry was detected.

Evidence Gap Map

An evidence gap map was constructed using ggplot2, mapping outcome domains (Y-axis) against comparator types (X-axis), with bubble size proportional to sample size and fill color representing GRADE certainty. Blank cells represent evidence gaps.

Analyses were performed using R 4.4.3 and metafor 5.0–1. Evidence quality was rated with GRADE [22], starting at high for RCTs and low for nonrandomized studies, with downgrading for risk of bias, inconsistency, indirectness, imprecision, and publication bias, and upgrading where justified. Data and code are provided in Online Resource 8.

3. RESULTS

3.1 Study Selection

The original search and supplementary gap-fill search were handled as one unified identification process. Across database, registry, grey-literature, citation, and supplementary searches, approximately 3,300 records were identified before final deduplication. After title/abstract screening and full-text assessment, 11 controlled studies met eligibility criteria. Common exclusions included single-modality interventions, single-arm studies, protocol-only records, and studies without extractable lung cancer surgery data; candidate and ongoing studies are summarized in Online Resource 3. The PRISMA flow diagram is presented in Fig. 1.

3.2 Study Characteristics

The 11 included studies enrolled or assigned 2,088 participants. Eight RCTs randomized or enrolled 1,294 participants, and three nonrandomized controlled studies contributed 794 participants. Nine

studies were conducted in China [7, 10–13, 23–26], one in Canada [27], and one in South Korea [28]. Comparator types included formal MM in one RCT [7], usual or routine perioperative care in most RCTs [10–13, 23–25, 29], and historical or matched standard-care comparators in the nonrandomized studies [26–28]. Follow-up ranged from discharge or early postoperative hospitalization to 12 months. Detailed study characteristics are presented in Online Resource 3.

The interventions clustered into three configurations: monitoring plus education/symptom management [12, 23, 24]; monitoring plus rehabilitation feedback through wearables, apps, or connected devices [10, 11, 26, 28, 29]; and integrated digital therapeutics or remote-care platforms combining symptom surveillance, education, exercise guidance, and clinician feedback [7, 13, 27].

3.3 Risk of Bias

Among the eight RCTs, one [10] was low risk overall and seven had some concerns, primarily because digital intervention blinding was infeasible and subjective PROs were vulnerable to expectation effects. The three nonrandomized studies had serious risk of bias due to selection, confounding, and missing or variable follow-up denominators. Domain-level assessments are provided in Online Resource 4 and visualized in Fig. 2.

3.4 Primary Outcomes

3.4.1 Pulmonary Function Recovery

Three RCTs ($n = 345$; Fig. 3A) contributed source-verifiable mean/SD-based pulmonary-function data to the unified quantitative synthesis: Xu [7] (FEV1% predicted), Jiang [23] (PaO₂), and Chen 2026 [25] (FEV1). The direction favored multimodal DHIs (SMD = 0.705, 95% CI 0.137 to 1.272; $P = 0.015$; $I^2 = 83.9\%$; $\tau^2 = 0.207$; GRADE: very low). Heterogeneity remained substantial, and the 95% prediction interval was wide (-0.353 to 1.762), because the model combined different pulmonary constructs and time points. Adding Lv 2025 [10] through median/IQR conversion produced a similar estimate (4 RCTs; $n = 481$; SMD = 0.607, 95% CI 0.198 to 1.016; $P = 0.0036$; $I^2 = 78.8\%$), supporting a directionally favorable but still heterogeneous signal; the 95% prediction interval remained wide (-0.220 to 1.434). The unified meta-analysis and SWiM summary is presented in Online Resource 5.

3.4.2 Postoperative Pulmonary Complications

Postoperative pulmonary complications could not be meta-analyzed. Xu et al. [7] excluded patients who experienced severe postoperative complications, precluding extraction of per-arm Clavien-Dindo PPC counts. Dai/Jing 2025 [12] reported readmission rather than Clavien-Dindo-graded PPCs. Patel 2023 [29] reported intraoperative complications and adverse events, but not the pre-specified PPC endpoint. Therefore, no included study had extractable PPC counts suitable for quantitative synthesis.

3.4.3 Hospital Length of Stay

Two RCTs ($n = 242$; Fig. 3B) reported compatible mean/SD LOS data [7, 29]. The pooled direction favored shorter hospitalization with multimodal DHIs but was imprecise and highly heterogeneous ($MD = -0.804$ days, 95% CI -2.542 to 0.934 ; $P = 0.365$; $I^2 = 89.6\%$; GRADE: very low). Xu 2025 showed almost identical LOS between DTx and MM (3.85 ± 0.99 vs. 3.84 ± 1.07 days), whereas Patel 2023 showed shorter LOS with Move For Surgery (2.67 ± 1.61 vs. 4.44 ± 3.48 days). Lv 2025 and Dai/Jing 2025 reported LOS as median(IQR) and were retained descriptively.

3.5 Secondary Outcomes

3.5.1 Pain Intensity

Pain was not meta-analyzed after full-text audit. Lv 2025 provided extractable MDASI-LC pain mean $\pm$ SD data (1.71 ± 1.81 vs. 2.05 ± 1.45 ; Hedges $g = -0.21$, $SE = 0.17$), favoring the app intervention but with a small effect. Dai/Jing 2025 reported pain threshold values as median(IQR), and Jiang 2024 reported VAS as median(Q1,Q3), precluding direct pooling with mean/SD SMDs. Pain evidence was therefore synthesized descriptively and retained as low/insufficient certainty.

3.5.2 Health-Related Quality of Life

Four RCTs ($n = 438$; Fig. 4) contributed source-verifiable mean/SD HRQoL data [11, 23–25]. Multimodal DHIs improved HRQoL ($SMD = 0.647$, 95% CI 0.276 to 1.018 ; $P < 0.001$; $I^2 = 67.9\%$; GRADE: very low), with a 95% prediction interval from -0.056 to 1.351 . Instruments included EORTC QLQ-C30 global health or total scores and WHOQOL-100 physical domain. Dai/Jing 2025 additionally reported adjusted HRQoL benefits but was not pooled because raw mean/SD data were unavailable.

3.5.3 Anxiety and Depression

Sui 2020 [24] provided extractable dichotomous HADS anxiety and depression outcomes at 12 months. The WeChat education and rehabilitation program reduced anxiety rate ($19/100$ vs. $41/100$; $RR = 0.463$, 95% CI 0.290 to 0.740) and depression rate ($20/100$ vs. $36/100$; $RR = 0.556$, 95% CI 0.347 to 0.890). Li 2025 [11] and Dai/Jing 2025 [12] also reported favorable psychological or emotional-function signals, but reporting formats were not sufficiently compatible for pooled quantitative synthesis.

3.5.4 Treatment Adherence and Management Time

Adherence definitions were heterogeneous (proportion completing scheduled ePRO assessments, proportion using app at least once, proportion completing prescribed rehabilitation); descriptive synthesis is presented in Online Resource 6. Only Xu et al. [7] directly compared management time: DTx 1.48 ± 3.22 min versus MM 16.67 ± 6.41 min per patient per encounter ($MD = -15.19$ min, 95% CI -17.12 to -13.26 , $P < 0.001$), a 91.1% reduction.

3.6 Nonrandomized Controlled Evidence

Three nonrandomized controlled studies ($n = 794$) were included in the qualitative evidence base. Lee et al. [28] reported higher daily steps and moderate-to-vigorous physical activity with wearable-supported

exercise; Tang/Wang [27] reported remote pulmonary-function monitoring and emergent complication detection; and Chen 2024 [26] reported better FEV1/FVC and shorter postoperative hospitalization. These studies supported the direction of benefit but were not pooled with RCTs because of serious confounding, selection, or measurement concerns (Online Resource 7).

3.7 Subgroup Analyses and Exploratory Trends

3.7.1 Comparator Type

The single study with an MM comparator [7] could not be formally compared with the non-MM subgroup due to the absence of replication. Descriptive comparison showed numerically consistent FEV1% recovery in Xu 2025 and favorable non-MM pulmonary-function estimates in the RCT and NRSI evidence. This observation supports the hypothesis of comparable clinical effectiveness but is insufficient to draw definitive conclusions about non-inferiority.

3.7.2 Number of Digital Modalities

A descriptive modality-count comparison was not stable because intervention complexity was confounded by comparator type, outcome instrument, and study design. The available pattern supports clinical plausibility for synergistic multi-component interventions but does not establish a dose-response relationship between modality count and effect size.

3.7.3 Study Design

RCTs and nonrandomized controlled studies were synthesized separately by design. RCTs formed the quantitative evidence base, whereas nonrandomized studies contributed supportive SWiM evidence because of serious confounding, selection, and missing-data concerns.

3.8 Evidence Gap Map

The evidence gap map (Fig. 5) revealed a developing but still sparse evidence base. Key evidence gaps were: (1) DHI versus MM comparison: evidence exists only for pulmonary function and management time from a single RCT, a critical gap for practice change; (2) geographic diversity: despite inclusion of studies from China, Canada, and South Korea, evidence remains dominated by China; (3) implementation outcomes: no study reported formal cost-effectiveness or cost-utility outcomes; (4) long-term outcomes: survival and durable functional recovery remain rarely reported; and (5) underserved populations: older adults, patients with cognitive impairment, and those with low digital literacy were systematically excluded or underrepresented.

3.9 Sensitivity Analyses and Publication Bias

Sensitivity analyses were limited by the small number of source-verifiable studies. Including Lv 2025 by median/IQR conversion in the pulmonary-function model produced SMD = 0.607 (95% CI 0.198 to 1.016; $I^2 = 78.8\%$), consistent with the main direction of effect. Leave-one-out diagnostics (Online Resource 5) indicated that heterogeneity was materially influenced by the mixture of FEV1%, PaO2, and FEV1

outcomes. PPC and pain meta-analytic sensitivity analyses were not performed because quantitative synthesis was not justified for those outcomes.

Publication bias could not be formally assessed for any outcome due to insufficient study numbers (all $k < 10$). The pulmonary-function funnel plot (Online Resource 5) is provided for transparency only and should not be interpreted as a reliable asymmetry test.

3.10 GRADE Evidence Quality

GRADE ratings ranged from insufficient to very low (Online Resource 5). Primary downgrading factors were open-label designs, inconsistency, imprecision, indirectness from heterogeneous instruments and time points, and serious confounding risk for nonrandomized studies. No outcome met upgrading criteria.

4. DISCUSSION

4.1 Principal Findings

This unified systematic review integrates the original search and supplementary gap-fill search into a single evidence base of 11 controlled studies involving 2,088 participants. Across eight RCTs, the most defensible quantitative signals favored multimodal DHIs for pulmonary function (3 RCTs; SMD = 0.705, 95% CI 0.137 to 1.272), HRQoL (4 RCTs; SMD = 0.647, 95% CI 0.276 to 1.018), anxiety and depression rates in one RCT, and prolonged hospital stay in one RCT. LOS direction favored DHIs but remained statistically uncertain and heterogeneous. PPCs and pain could not be pooled because endpoint definitions and reporting formats were incompatible. Three nonrandomized controlled studies provided supportive but very low-certainty signals for activity, dyspnea, pulmonary function, and hospitalization outcomes, and were synthesized narratively because of serious confounding and missing-data risks.

4.2 Comparison with Existing Evidence

The present findings extend prior reviews in three respects. First, Lu et al. [14] focused on digital pulmonary rehabilitation, whereas this review required multimodal DHI architecture and audited PPC definitions before pooling. Second, prior app-based recovery reviews [15] pooled mixed surgical oncology populations, whereas this review isolated lung cancer surgery. Third, evidence gap mapping shows that the most clinically important comparison, multimodal DHI versus formal MM, remains supported by only one RCT.

4.3 Interpretation of Heterogeneity

Heterogeneity was substantial for pulmonary function and LOS. For pulmonary function, FEV1%, PaO₂, PFRR, and FEV1 are related but not interchangeable physiological constructs. This heterogeneity supports interpreting pooled SMDs as broad recovery signals rather than instrument-specific effects. For LOS, Xu 2025 compared DTx with formal MM and showed almost no difference, while Patel 2023 compared preconditioning with usual care and favored the digital preconditioning program. Comparator

type, intervention phase, surgical pathway, and health-system context likely explain part of the inconsistency.

4.4 Evidence Gap Map: Charting the Research Frontier

The evidence gap map reveals a nascent but rapidly developing field. The most clinically important gap remains DHI-versus-MM comparison: only one RCT directly compared multimodal DHI with formal MM, and no study replicated management-time or cost-effectiveness outcomes. Evidence is also sparse for PPCs, pain, adherence mechanisms, and digitally underserved populations.

4.5 Clinical Translation Value

For clinical implementation, multimodal DHIs are best considered adjunctive supportive-care tools rather than replacements for perioperative multidisciplinary care. Tertiary centers may use them to extend monitoring and routine rehabilitation feedback, while resource-limited settings may use them to improve post-discharge continuity, provided escalation pathways and digital access barriers are addressed.

4.6 Strengths

This review has several strengths. First, it integrates an initial search and a supplementary gap-fill search into one coherent evidence base, reducing the risk of presenting missed studies as a separate post hoc analysis. Second, the strict multimodal definition distinguishes synergistic multi-component digital systems from single-function digital tools. Third, all included full texts used for extraction were locally archived and source-verified. Fourth, the design-stratified analysis preserves causal interpretability by pooling RCTs while narratively synthesizing nonrandomized studies. Fifth, the evidence gap map provides a structural contribution beyond traditional meta-analysis.

4.7 Limitations

Several limitations warrant consideration. First, even after gap-fill searching, compatible quantitative data remain sparse and all pooled estimates are exploratory. Second, most RCTs were open-label, which is difficult to avoid in digital health trials but increases concern for subjective PROs. Third, pulmonary function and HRQoL measures were heterogeneous, limiting clinical interpretability of SMD-based estimates. Fourth, nonrandomized studies were at serious risk of bias and cannot establish causal effects. Fifth, only one study used a formal MM comparator, limiting conclusions about whether DHIs can substitute for or complement multidisciplinary teams. Sixth, no included study reported formal cost-effectiveness, and older adults with frailty, cognitive impairment, or low digital literacy were underrepresented.

4.8 Future Research Agenda

Future trials should prioritize direct comparisons between multimodal DHIs and formal MM, prespecified core outcomes for thoracic surgery digital health, blinded adjudication of complications, full reporting of means/SDs or event counts, and economic evaluation. Hybrid effectiveness-implementation trials using CFIR or RE-AIM frameworks are needed to determine how digital platforms can be integrated into routine

thoracic oncology pathways. Dedicated studies in older, frail, rural, and low-digital-literacy populations are particularly important.

5. CONCLUSIONS

After integrating the original and supplementary gap-fill searches, 11 controlled studies suggest that multimodal DHIs may improve pulmonary-function recovery, HRQoL, psychological outcomes, and selected hospitalization outcomes after lung cancer surgery. Certainty remains very low, so current evidence supports cautious adjunctive integration into perioperative supportive care and underscores the need for rigorous, standardized, implementation-focused trials.

Statements and Declarations

Funding

This work was supported by the Chongqing Science and Health Joint Medical Science and Technology Innovation and Key Project (2025GGXM001).

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval

This systematic review and meta-analysis used published aggregate data and did not involve new human participant recruitment, human biological material, or animal research; therefore, ethics approval was not required.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Data availability

The extraction datasets, analysis code, search materials, and supporting files are provided with the submission as Online Resource 8. Public repository deposition can be added before acceptance if required by the journal.

Author contributions

PHZ: Conceptualization, methodology, data curation, formal analysis. ZSL: writing – original draft, writing – review & editing, visualization. Conceptualization, supervision, writing – review & editing, funding acquisition, project administration. TLL: Investigation, data curation, validation. WK: Investigation, data curation. CMY: Investigation, data curation. LQX: Conceptualization, methodology, supervision, writing – review & editing, funding acquisition, project administration. All authors reviewed and approved the final manuscript.

Use of AI tools

AI tools were used for language editing, formatting support, and document organization. All scientific content, literature decisions, extracted data, analyses, interpretations, and the final submitted version must be verified and approved by the human authors.

References

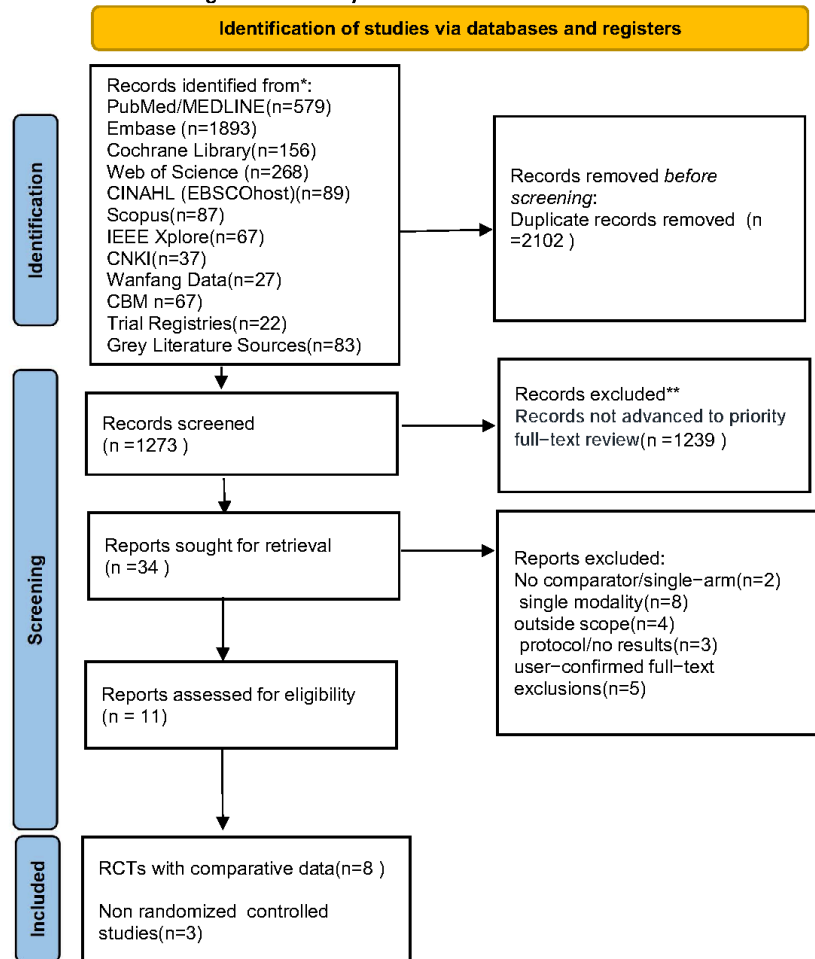
1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249.
2. Howington JA, Blum MG, Chang AC, et al. Treatment of stage I and II non-small cell lung cancer. *Chest*. 2013;143(5 Suppl):e278S–e313S.
3. Agostini P, Cieslik H, Rathinam S, et al. Postoperative pulmonary complications following thoracic surgery: are there any modifiable risk factors? *Thorax*. 2010;65(9):815–818.
4. Yang M, Zhong JD, Zhang JE, et al. Systematic review on the role of technology in supporting lung cancer patients. *Support Care Cancer*. 2022;30(1):25–38.
5. Li J, Liang R, Song Y, et al. Systematic review and meta-analysis of digital interventions for mental health in cancer patients. *Psychooncology*. 2021;30(12):1995–2010.
6. Batchelor TJP, Rasburn NJ, Abdelnour-Berchtold E, et al. Guidelines for enhanced recovery after lung surgery. *Eur J Cardiothorac Surg*. 2019;55(1):91–115.
7. Xu J, Ni H, Zhan H, et al. Efficacy of digital therapeutics for perioperative management in patients with lung cancer: a randomized controlled trial. *BMC Med*. 2025;23:186.
8. Keesara S, Jonas A, Schulman K. Covid-19 and health care's digital revolution. *N Engl J Med*. 2020;382(23):e82.
9. Marwaha JS, Raza MM, Kvedar JC. The digital transformation of surgery. *NPJ Digit Med*. 2023;6:103.
10. Lv C, Lu F, Zhou X, et al. Efficacy of a smartphone application assisting home-based rehabilitation after VATS lobectomy (POPPER study). *Int J Surg*. 2025;111(1):597–608.
11. Li G, Zhou X, Deng J, et al. Digital therapeutics-based cardio-oncology rehabilitation for lung cancer survivors: randomized controlled trial. *JMIR Mhealth Uhealth*. 2025;13:e60115.
12. Dai W, Feng W, Zhang Y, et al. Physiological and psychological symptom management based on ePRO: the TD-WELLBEING randomized clinical trial. *Br J Cancer*. 2025;133(7):937–944.

13. Dai W, Wang Y, Liao J, et al. Electronic patient-reported outcome-based symptom management versus usual care after lung cancer surgery: long-term results. *J Clin Oncol*. 2024;42(15):1781–1792.
14. Lu T, Deng T, Long Y, et al. Effectiveness and feasibility of digital pulmonary rehabilitation in patients undergoing lung cancer surgery: systematic review and meta-analysis. *J Med Internet Res*. 2024;26(1):e56795.
15. He M, Chen M, Ji Y, Lu G. Effectiveness of smartphone app-based interventions after surgery on quality of recovery among cancer patients: a systematic review and meta-analysis. *Ann Med*. 2024;56(1):2390167.
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement. *BMJ*. 2021;372:n71.
17. Wan X, Wang W, Liu J, et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol*. 2014;14:135.
18. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
19. Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
20. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ*. 2020;368:l6890.
21. Furuya-Kanamori L, Barendregt JJ, Doi SAR. A new improved graphical and quantitative method for detecting bias in meta-analysis. *Int J Evid Based Healthc*. 2018;16(4):195–203.
22. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence. *BMJ*. 2008;336(7650):924–926.
23. Jiang Y, Jin Y, Cai J, et al. A mobile health-based management for patients with limited lung function after minimally invasive thoracoscopic lobectomy. *Altern Ther Health Med*. 2024;30(3):133–139.
24. 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;44:101707. doi:10.1016/j.ejon.2019.101707.
25. Chen L, Liu J, Xia H, Jiang L, Shen S. Intelligent breathing training using a digital device for postoperative lung cancer patients: A randomized controlled trial. *Biol Res Nurs*. 2026;0(0):1-13. doi:10.1177/10998004251411972.
26. Chen Z, Wang Y, Qiu P, Tan L, Hu Y. Pulmonary rehabilitation exercise based on wearable device pedometer improved lung cancer patients with impaired pulmonary function. *Altern Ther Health Med*. 2024;30(1):78-82.
27. Tang D, Wang J, Fu W, et al. Applying digital technologies for remote care in the real life context: A 3-year experimentation with postoperative lung cancer patients. *Medicine*. 2026;105(21):e48750. doi:10.1097/MD.00000000000048750.

28. Lee J, Kong S, Shin S, et al. Wearable device-based intervention for promoting patient physical activity after lung cancer surgery: A nonrandomized clinical trial. *JAMA Netw Open*. 2024;7(9):e2434180. doi:10.1001/jamanetworkopen.2024.34180.
29. Patel YS, Sullivan KA, Churchill IF, et al. Preconditioning program reduces the incidence of prolonged hospital stay after lung cancer surgery: Results from the Move For Surgery randomized clinical trial. *BJS*. 2023;110:1-6. doi:10.1093/bjs/znad252.

Figures

PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only



*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: <http://www.prisma-statement.org/>

Figure 1

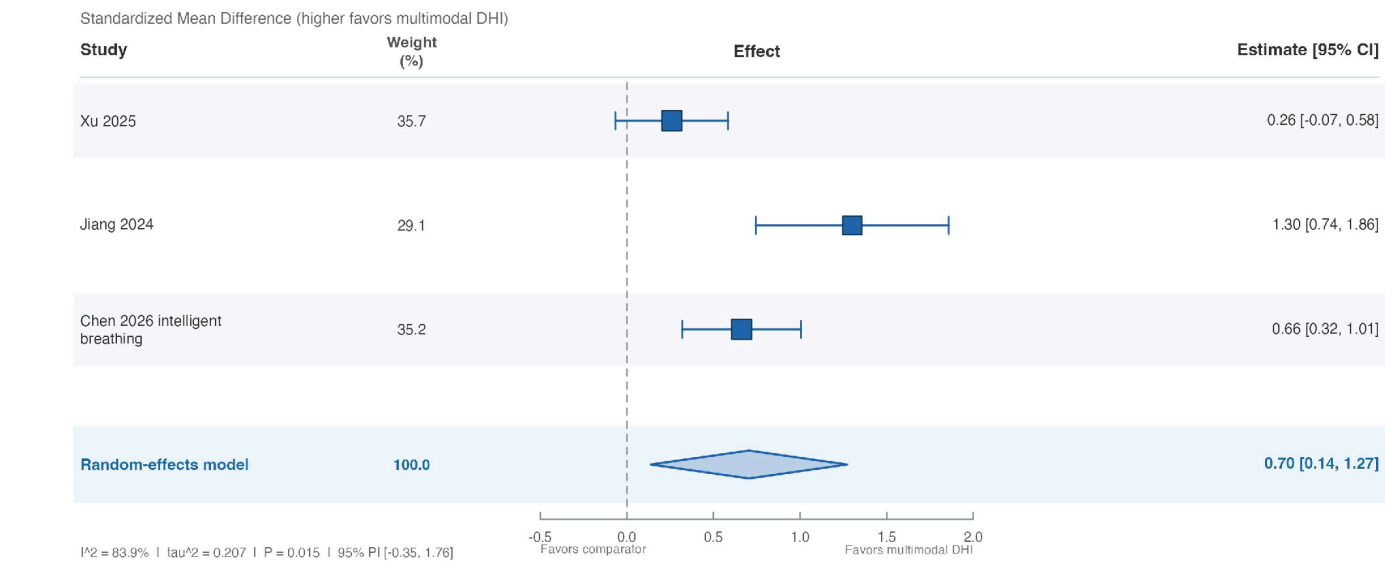
See image above for figure legend.



Figure 2

Figure legend not provided with this version

A **Pulmonary Function: Multimodal DHI vs. Comparator**



B **Hospital Length of Stay: Multimodal DHI vs. Comparator**

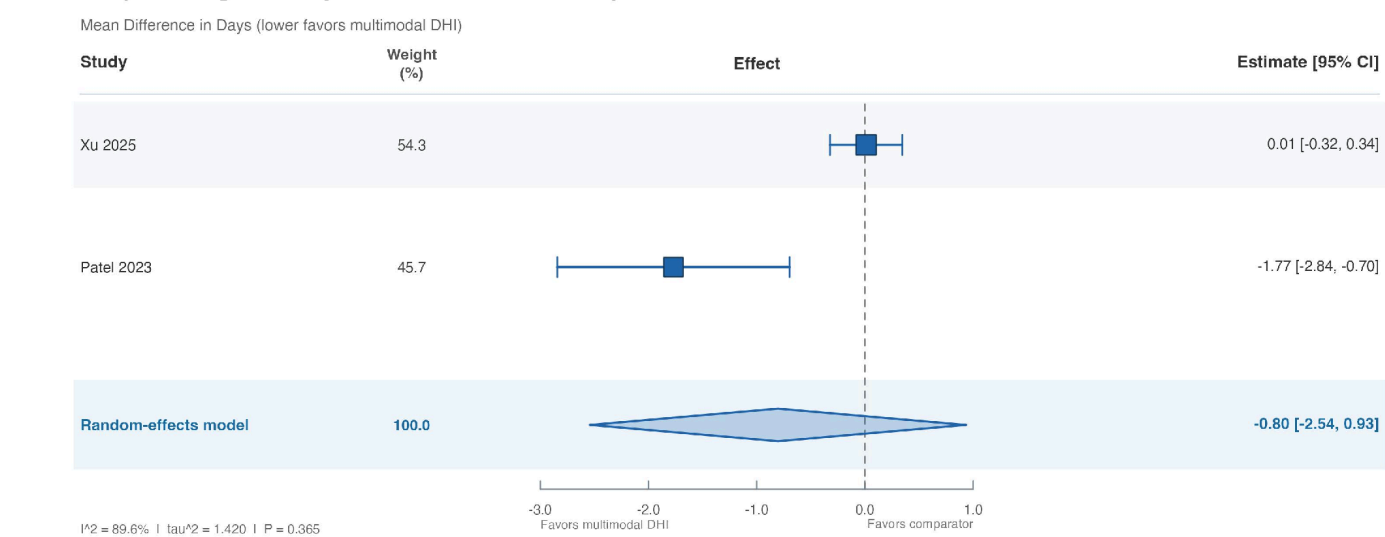


Figure 3

Figure legend not provided with this version

Health-Related Quality of Life: Multimodal DHI vs. Comparator

Standardized Mean Difference (higher favors multimodal DHI)

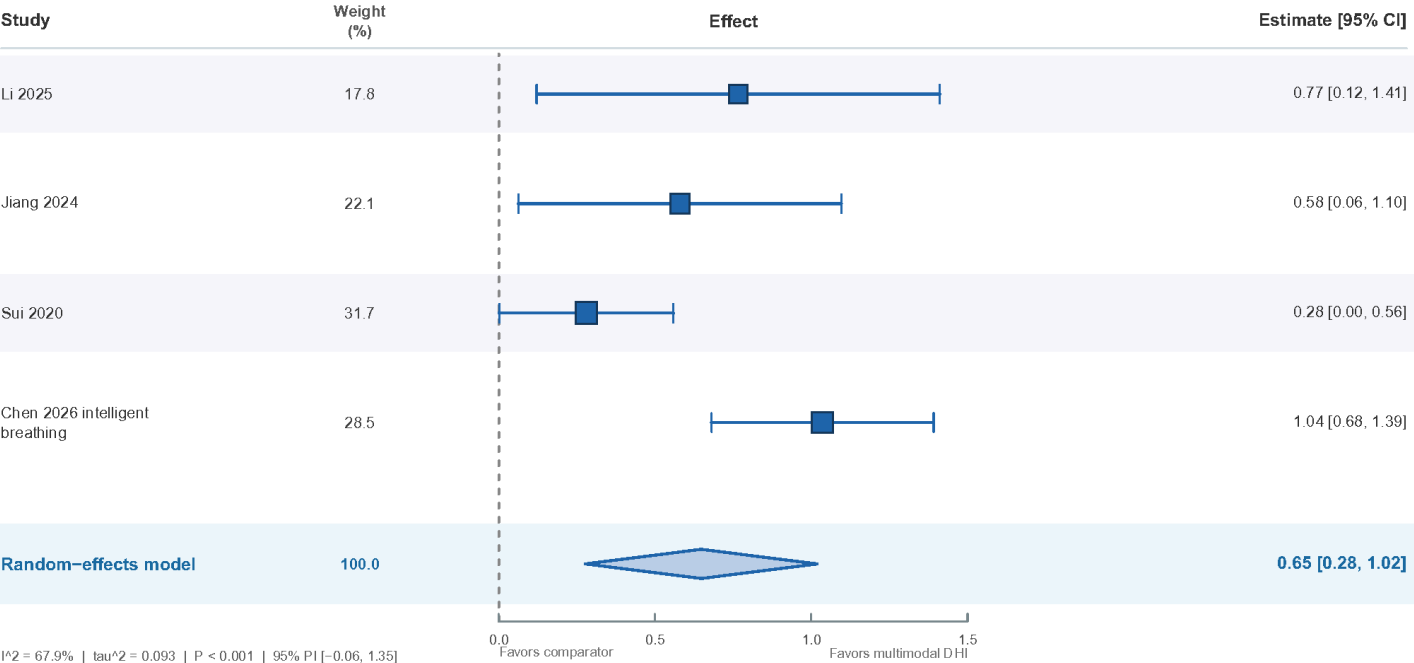


Figure 4

Figure legend not provided with this version

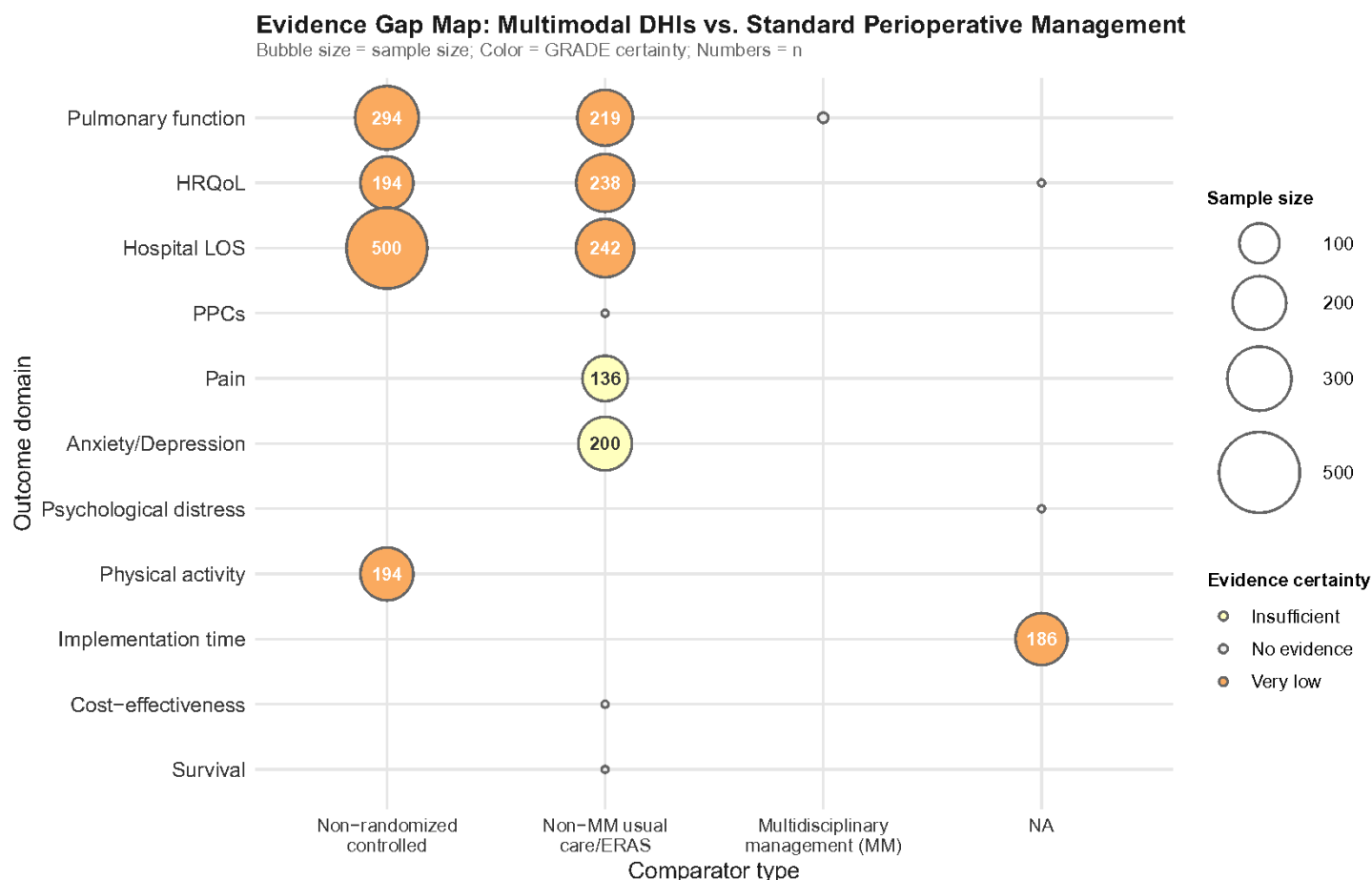


Figure 5

Figure legend not provided with this version

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM1PRISMA2020checklist.pdf](#)
- [ESM5bLeaveoneoutpulmonaryfunction.pdf](#)
- [ESM3Studycharacteristicsandexcludedcandidates.pdf](#)
- [ESM4Riskofbiasdetail.pdf](#)
- [ESM5aFunnelplotpulmonaryfunction.pdf](#)
- [ESM5MetaanalysisGRADEsensitivity.pdf](#)
- [ESM7NonrandomizedSWiMevidence.pdf](#)
- [ESM6Adherenceandmanagementtime.pdf](#)
- [ESM2Searchstrategiesandaudit.pdf](#)