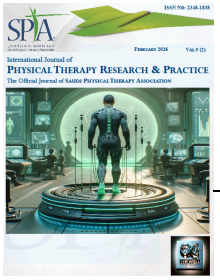

INTERNATIONAL JOURNAL OF
PHYSICAL THERAPY RESEARCH & PRACTICE
AN OFFICIAL JOURNAL OF SAUDI PHYSICAL THERAPY ASSOCIATION
SPTA
الجمعية السعودية للعلاج الطبيعي
Saudi Physical Therapy Association

Invited Review

High-Intensity Interval Training Versus Moderate-Intensity Continuous Training in Older Adults With Heart Failure: A Systematic Review of Comparative Randomized Trials

Mohammad Abu Shaphe[1*]; Shadab Uddin[1]; Rashid Ali Beg[1]

1. Department of Physical Therapy, College of Nursing and Health Sciences, Jazan University, Jazan, Saudi Arabia.

*Corresponding Author: mshaphe@jazanu.edu.sa **DOI:** https://doi.org/10.62464/p770w219

Abstract

Background: High-intensity interval training (HIIT) has been proposed as a superior alternative to moderate-intensity continuous training (MICT) for improving exercise capacity in patients with heart failure (HF). Early small, randomized trials suggested substantial gains in peak oxygen uptake (peak VO_2) and left ventricular ejection fraction (LVEF) with HIIT, but larger multicenter trials have reported neutral findings. **Objective:** To systematically evaluate the comparative effects of HIIT versus MICT on exercise capacity and LVEF in older adults with heart failure, and to assess the certainty of evidence across heart failure phenotypes. **Methods:** A systematic review of randomized controlled trials comparing HIIT and MICT in adults aged ≥60 years with heart failure was conducted in accordance with PRISMA 2020 guidelines. Databases were searched from inception to the final search date without language restriction. Primary outcomes were change in peak VO_2 (mL/kg/min) and LVEF (%). Secondary outcomes included six-minute walk distance (6MWT) and serious adverse events. Random-effects meta-analyses were performed where appropriate. Risk of bias was assessed using the Cochrane RoB 2 tool, and certainty of evidence was evaluated using GRADE. **Results:** Sixteen comparative randomized studies met inclusion criteria. In heart failure with reduced ejection fraction (HFrEF), pooled meta-analyses demonstrated a modest improvement in peak VO_2 favoring HIIT (mean difference approximately +1.7 to +2.1 mL/kg/min). However, the largest multicenter trial (n=261) showed no significant between-group difference. Pooled improvements in LVEF were small (~3%) and inconsistent across studies. In heart failure with preserved ejection fraction (HFpEF), randomized evidence did not demonstrate superiority of HIIT over MICT, and adjusted analyses suggested exercise volume rather than intensity drove adaptation. Improvements in 6MWT (~28 meters) modestly favored HIIT but approached minimal clinically important thresholds. Serious adverse events were infrequent and comparable between groups under supervised conditions. **Conclusions:** In older adults with heart failure, HIIT confers modest improvements in peak VO_2 in HFrEF populations according to pooled analyses, but these advantages are not consistently replicated in large pragmatic multicenter trials and are not observed in HFpEF. Structural improvements in LVEF remain uncertain.

Keywords: Heart failure; high-intensity interval training; moderate-intensity continuous training; peak oxygen uptake; left ventricular ejection fraction; cardiac rehabilitation; HFrEF; HFpEF; systematic review.

Article info : Received: Published: **Feb. 28, 2026**

To Cite: Mohammad Abu Shaphe, Shadab Uddin, Rashid Ali Beg. High-Intensity Interval Training Versus Moderate-Intensity Continuous Training in Older Adults With Heart Failure: A Systematic Review of Comparative Randomized Trials. (2026). International Journal of Physical Therapy Research & Practice, 5(2), 150-178. https://doi.org/10.62464/p770w219

1. Introduction

1.1 Burden of Heart Failure in Older Adults

Heart failure (HF) is among the most consequential cardiovascular syndromes affecting older adults, with prevalence rising sharply with advancing age and contributing substantially to recurrent hospitalization, health-system costs, and long-term mortality risk. Although contemporary trials in both HFrEF and HFpEF emphasize physiological and functional outcomes rather than epidemiologic estimates, they consistently underscore that older HF populations experience clinically meaningful limitations in daily activity and overall well-being, reinforcing the public health and clinical significance of this syndrome (Ellingsen et al., 2017; Mueller et al., 2021; Stensvold et al., 2024). Reduced exercise tolerance is a defining hallmark of HF and is strongly linked to adverse prognosis. In this context, peak oxygen consumption (peak VO_2), measured via cardiopulmonary exercise testing (CPET), is widely used as a robust integrative indicator of functional capacity and a central endpoint in HF rehabilitation trials, reflecting its clinical relevance for risk stratification and therapeutic evaluation (Wisløff et al., 2007; Haykowsky et al., 2013; Xie et al., 2017).

The pathophysiology of exercise intolerance in HF is multifactorial and extends beyond impaired cardiac pumping capacity to include chronotropic incompetence, endothelial dysfunction, abnormal ventilatory responses, skeletal muscle deconditioning, and altered peripheral oxygen extraction. Contemporary randomized trials and mechanistic analyses evaluating exercise training in HF emphasize that both central and peripheral mechanisms contribute to functional limitation and that training-induced benefits may reflect improvements across these interacting pathways (Wisløff et al., 2007; Mueller et al., 2022; Nasser et al., 2025). In HFrEF, reduced systolic function limits stroke volume augmentation and cardiac output reserve during exertion; in HFpEF, impaired ventricular compliance, microvascular dysfunction, and peripheral abnormalities appear to play dominant roles in limiting exercise capacity, which may partly explain differences in responsiveness to intensity-based training prescriptions across phenotypes (Mueller et al., 2021; Mueller et al., 2022; Angadi et al., 2015). Accordingly, structured exercise training has emerged as a cornerstone non-pharmacological therapy capable of targeting central and peripheral determinants of functional impairment and remains a consistent component of evidence-informed HF management strategies (Hannan et al., 2018; Yang et al., 2024).

International practice recommendations support aerobic exercise training for clinically stable HF patients, reflecting a broad evidence base demonstrating improvements in exercise capacity and patient-centered outcomes. Nevertheless, the optimal intensity prescription remains debated. Moderate-intensity continuous training (MICT) has traditionally formed the foundation of cardiac rehabilitation programming, but high-intensity interval training (HIIT) has gained increasing attention as a potentially more potent stimulus for cardiovascular and skeletal muscle adaptation. Early randomized data suggested marked improvements with HIIT in HFrEF, including large gains in peak VO_2 and favorable remodeling indices (Wisløff et al., 2007), whereas later pragmatic multicenter trials reported neutral between-group differences, highlighting the importance of adherence fidelity and achieved training intensity (Ellingsen et al., 2017). Meta-analyses continue to report modest pooled advantages of HIIT for peak VO_2, although findings are influenced by trial size and heterogeneity (Haykowsky et al., 2013; Xie et al., 2017; Yang et al., 2024; Fauzi, 2025). Collectively, these data frame the clinical relevance of refining exercise prescription in older adults with HF while recognizing that phenotype, adherence, and real-world implementation constraints may shape the comparative effectiveness of HIIT and MICT (Mueller et al., 2021; Mueller et al., 2022).

1.2 Rationale for High-Intensity Interval Training

High-intensity interval training (HIIT) consists of repeated bouts of exercise performed at high relative intensities typically 85–95% of peak heart rate or peak oxygen uptake (VO_2peak) interspersed with structured periods of active recovery. Compared with moderate-intensity continuous training (MICT), HIIT is theorized to elicit greater shear stress–mediated endothelial adaptations, enhanced mitochondrial biogenesis, improved stroke volume augmentation, and superior skeletal muscle oxidative capacity. These physiological mechanisms provide the conceptual basis for proposing HIIT as a potentially more potent stimulus for cardiovascular and peripheral adaptation in patients with heart failure (Haykowsky et al., 2013; Xie et al., 2017).

Enthusiasm for HIIT in heart failure was catalyzed by the landmark randomized trial conducted by Wisløff et al. (2007). In this single-center study of patients with heart failure with reduced ejection fraction (HFrEF), a 12-week 4×4-minute HIIT protocol produced a 46% increase in peak VO_2 compared with a 14% increase in the MICT group. Notably, HIIT was also associated with an approximately 10-percentage-point absolute increase in left ventricular ejection fraction (LVEF), accompanied by evidence of reverse ventricular remodeling. These findings suggested that exercise intensity rather than duration alone might play a critical role in promoting central cardiac recovery and improving systolic function (Wisløff et al., 2007).

Subsequent smaller randomized trials reinforced this hypothesis. In a pilot study of patients with heart failure with preserved ejection fraction (HFpEF), Angadi et al. (2015) reported significant improvements in peak VO_2 following HIIT, whereas MICT resulted in minimal change. Similarly, Benda et al. (2015) observed improvements in physical fitness and cardiovascular function following interval-based training protocols. These early mechanistic and small-scale clinical investigations contributed to a growing body of evidence suggesting that higher-intensity training might induce superior functional and vascular adaptations compared with continuous moderate exercise.

Early meta-analytic syntheses further consolidated this perception. Haykowsky et al. (2013) reported a weighted mean difference of 2.14 mL/kg/min in peak VO_2 favoring HIIT in HFrEF populations. Xie et al. (2017) demonstrated a pooled mean difference of 1.76 mL/kg/min in favor of HIIT across cardiac populations, including heart failure cohorts. Trejos-Montoya et al. (2019) similarly observed greater improvements in LVEF and aerobic capacity with HIIT compared with MICT in pooled analyses.

Collectively, these early trials and meta-analyses fostered the prevailing view that HIIT may be superior to MICT for improving aerobic capacity and potentially systolic function in patients with heart failure. However, as later multicenter trials demonstrated, the magnitude and consistency of this superiority appear sensitive to study design, sample size, adherence fidelity, and phenotype-specific factors, underscoring the need for careful interpretation of early high-intensity findings.

1.3 Emerging Conflicting Evidence From Multicenter Trials

Despite strong early signals suggesting superiority of high-intensity interval training (HIIT), subsequent large, multicenter randomized controlled trials challenged this hypothesis. The SMARTEX-HF trial, which enrolled 261 patients with heart failure with reduced ejection fraction (HFrEF) across multiple centers, compared a 4×4-minute HIIT protocol with moderate-intensity continuous training (MICT) and a home-based exercise recommendation group (Ellingsen et al., 2017; Støylen et al., 2012). In contrast to earlier single-center findings, SMARTEX-HF reported no statistically significant between-group difference in peak oxygen uptake (peak VO_2) at 12 weeks (mean difference −0.4 mL/kg/min; $p = .70$), nor did it demonstrate superiority of HIIT in improving left ventricular ejection fraction (LVEF) (Ellingsen et al., 2017).

A critical methodological observation from SMARTEX-HF was substantial intensity contamination. Approximately 51% of participants randomized to HIIT failed to achieve the prescribed high-intensity threshold, whereas nearly 80% of participants assigned to MICT exceeded their target moderate intensity. This bidirectional deviation from protocol effectively narrowed the physiological contrast between groups and likely attenuated measurable differences in outcomes (Ellingsen et al., 2017). These findings underscored the practical challenges of maintaining strict intensity separation in multicenter, pragmatic rehabilitation settings.

Similarly, the OptimEx-Clin trial, which enrolled 180 patients with heart failure with preserved ejection fraction (HFpEF), found no significant superiority of HIIT over MICT in improving peak VO_2 (Mueller et al., 2021). Neither intervention achieved the prespecified minimal clinically important difference compared with guideline-based advice alone. A mechanistic subanalysis further demonstrated that after adjustment for total energy expenditure, the difference between HIIT and MICT was essentially null, suggesting that cumulative exercise volume rather than interval intensity may be the principal driver of physiological adaptation in HFpEF populations (Mueller et al., 2022).

The emergence of these rigorously conducted multicenter trials introduced uncertainty regarding whether earlier positive findings reflected true physiological superiority or were influenced by small-study effects, limited sample sizes, or highly controlled single-center environments. Nevertheless, subsequent meta-analyses have continued to report pooled advantages of HIIT. For example, Yang et al. (2024) identified a mean difference of 1.78 mL/kg/min in peak VO_2 and 3.13% in LVEF favoring HIIT, while Fauzi (2025) reported similar pooled estimates. However, these analyses were largely driven by smaller single-center trials, raising concerns regarding heterogeneity, publication bias, and the relative weighting of early mechanistic studies compared with large pragmatic trials (Haykowsky et al., 2013; Xie et al., 2017).

Collectively, the juxtaposition of early single-center superiority findings with later multicenter neutrality highlights a central tension in the evidence base and underscores the importance of considering trial scale, methodological rigor, and intervention fidelity when interpreting comparative effectiveness in heart failure rehabilitation.

1.4 Phenotype-Specific Considerations: HFrEF Versus HFpEF

The heterogeneity of heart failure phenotypes complicates interpretation of exercise training effects and may partially explain inconsistencies observed across trials. In heart failure with reduced ejection fraction (HFrEF), systolic dysfunction limits stroke volume augmentation and cardiac output during exertion. Because HIIT imposes repeated periods of near-maximal cardiovascular stress, it has been hypothesized that this modality may preferentially stimulate central cardiac adaptations in HFrEF. Early mechanistic evidence supporting this hypothesis was provided by Wisløff et al. (2007), who reported substantial improvements in peak VO_2 and left ventricular ejection fraction (LVEF), along with indices of reverse remodeling, following a 12-week HIIT protocol. Subsequent meta-analyses have generally demonstrated modest pooled advantages of HIIT over moderate-intensity continuous training (MICT) in HFrEF populations for both peak VO_2 and LVEF (Haykowsky et al., 2013; Xie et al., 2017; Yang et al., 2024; Fauzi, 2025). Although the magnitude of these effects is smaller in pooled analyses than in early single-center trials, the direction of effect in HFrEF cohorts has typically favored higher-intensity training.

In contrast, heart failure with preserved ejection fraction (HFpEF) is characterized predominantly by diastolic dysfunction, impaired ventricular compliance, endothelial and microvascular dysfunction, skeletal muscle abnormalities, and

reduced peripheral oxygen extraction. These pathophysiological mechanisms may not respond preferentially to short bursts of high intensity, as limitations are often peripheral and multifactorial rather than centrally systolic in origin. Supporting this distinction, Wang et al. (2022) demonstrated that although HIIT conferred significant benefits in coronary artery disease populations, the advantage was not statistically significant when analyses were restricted to heart failure cohorts. Furthermore, the OptimEx-Clin randomized trial found no superiority of HIIT over MICT in HFpEF patients (Mueller et al., 2021), and subsequent mechanistic analyses indicated that once total energy expenditure was accounted for, differences between modalities were negligible (Mueller et al., 2022).

These phenotype-specific findings underscore the importance of avoiding generalized intensity-based recommendations across all heart failure populations. While HIIT may confer modest central benefits in selected HFrEF patients, current evidence suggests equivalence between modalities in HFpEF, where cumulative workload, adherence, and peripheral conditioning may be more influential determinants of functional adaptation than peak interval intensity alone.

1.5 Safety Considerations in Older Adults

Given the advanced age, multimorbidity, and polypharmacy commonly observed in heart failure (HF) populations, safety considerations are central to exercise prescription. Across randomized trials comparing high-intensity interval training (HIIT) with moderate-intensity continuous training (MICT), serious adverse events have generally been infrequent and comparable between groups. In the multicenter SMARTEX-HF trial, no statistically significant difference in serious cardiovascular events was observed between HIIT and MICT during the intervention period (Ellingsen et al., 2017).

Similarly, long-term data from the Generation 100 study in older adults with cardiovascular disease reported no acute cardiovascular events attributable to HIIT over extended follow-up, further supporting the safety profile of structured high-intensity exercise when appropriately supervised (Stensvold et al., 2024). Meta-analytic syntheses have reached comparable conclusions, indicating that HIIT does not appear to increase the risk of serious adverse events relative to MICT in clinically stable HF populations (Fauzi, 2025; Hannan et al., 2018).

However, interpretation of safety findings requires caution. Most trials were not powered specifically to detect rare but clinically significant adverse outcomes, and event counts were generally low. Moreover, HIIT in these studies was delivered within structured cardiac rehabilitation environments, characterized by close monitoring, heart rate surveillance, and gradual progression. Extrapolation to unsupervised or community-based settings may therefore be limited. In older adults with HF particularly those with frailty, multiple comorbidities, or chronotropic incompetence careful screening, individualized intensity prescription, and ongoing supervision remain essential to ensure safe implementation. Adherence challenges and the need for monitoring infrastructure may also constrain the scalability of HIIT beyond structured clinical programs.

1.6 Evidence Gap and Rationale for the Present Review

Despite substantial investigation over the past two decades, the comparative efficacy of high-intensity interval training (HIIT) versus moderate-intensity continuous training (MICT) in older adults with heart failure remains incompletely resolved. Early small randomized trials reported striking improvements with HIIT, including marked gains in peak oxygen uptake (peak VO_2) and left ventricular ejection fraction (LVEF) in patients with heart failure with reduced ejection fraction (HFrEF) (Wisløff et al., 2007) and significant improvements in peak VO_2 in heart failure with preserved ejection fraction (HFpEF) cohorts (Angadi et al., 2015). These findings generated considerable enthusiasm for intensity-driven cardiac rehabilitation strategies.

However, subsequent large, multicenter randomized trials have yielded neutral or attenuated results. The SMARTEX-HF trial found no significant superiority of HIIT over MICT in improving peak VO_2 or LVEF in HFrEF populations (Ellingsen et al., 2017), while the OptimEx-Clin trial similarly demonstrated no meaningful advantage of HIIT in HFpEF patients (Mueller et al., 2021). Mechanistic analyses further suggested that when total energy expenditure is accounted for, differences between HIIT and MICT may disappear, particularly in HFpEF (Mueller et al., 2022).

Although multiple meta-analyses continue to report modest pooled advantages of HIIT for peak VO_2 and, in some analyses, LVEF (Haykowsky et al., 2013; Xie et al., 2017; Yang et al., 2024; Fauzi, 2025), these estimates are heavily influenced by small, single-center trials. The attenuation of effect sizes in larger pragmatic studies raises concerns regarding small-study effects, publication bias, and the reproducibility of early mechanistic findings in real-world rehabilitation settings. Moreover, heterogeneity in protocol design, supervision intensity, adherence fidelity, and participant phenotype complicates interpretation of pooled results.

Uncertainty also persists regarding phenotype-specific responsiveness, particularly between HFrEF and HFpEF populations, as well as the extent to which observed benefits reflect exercise intensity per se versus cumulative workload or energy expenditure. These unresolved questions underscore the need for a comprehensive synthesis that explicitly integrates trial size, methodological rigor, heart failure phenotype, adherence patterns, and safety considerations. Such an approach is essential to provide a balanced and clinically meaningful appraisal of the comparative effectiveness of HIIT and MICT in older adults with heart failure and to inform evidence-based rehabilitation practice.

1.7 Objective

The objective of this systematic review is to systematically evaluate and compare the effects of high-intensity interval training (HIIT) and moderate-intensity continuous training (MICT) on exercise capacity primarily assessed by peak oxygen uptake (peak VO_2) and left ventricular ejection fraction (LVEF) in older adults with heart failure. In addition to quantifying pooled effects, this review aims to critically appraise the methodological quality of included trials, examine the influence of study size and intervention fidelity, and explore potential differences in response between heart failure phenotypes, specifically heart failure with reduced ejection fraction (HFrEF) and heart failure with preserved ejection fraction (HFpEF). Through an integrated analysis of efficacy, safety, and certainty of evidence, this review seeks to provide a clinically relevant synthesis to inform evidence-based exercise prescription in this population.

2. Methods

2.1 Review Design and Reporting Framework

This systematic review was conducted in accordance with the methodological standards outlined in the Cochrane Handbook for Systematic Reviews of Interventions and reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines. The review was designed to evaluate the comparative efficacy and safety of high-intensity interval training (HIIT) versus moderate-intensity continuous training (MICT) in older adults with heart failure. The research question was structured using the PICO framework, with predefined eligibility criteria, outcomes, and analytic strategies established prior to data synthesis.

The population of interest comprised older adults diagnosed with stable heart failure. Studies were eligible if they enrolled participants aged 60 years or older, or if the reported mean or median age of the cohort was at least 60 years. Both heart failure with reduced ejection fraction (HFrEF) and heart failure with preserved ejection fraction (HFpEF)

phenotypes were eligible. HFrEF was generally defined as left ventricular ejection fraction (LVEF) ≤40%, whereas HFpEF was typically defined as LVEF ≥50%, although definitions were accepted as reported by individual trials. Studies enrolling patients with acute decompensated heart failure, recent myocardial infarction within four weeks, or unstable cardiovascular conditions were excluded to ensure clinical homogeneity and safety comparability. Trials exclusively enrolling younger populations were excluded unless subgroup data for older adults were separately extractable.

2.2 Eligibility Criteria

Eligible interventions consisted of structured HIIT programs characterized by repeated bouts of high-intensity exercise interspersed with defined recovery intervals. High intensity was operationalized as exercise performed at or above 80% of peak heart rate, 85% of maximum heart rate, 80% of heart rate reserve, or comparable relative intensity thresholds based on peak oxygen uptake. Protocols varied across trials but frequently included the 4×4-minute interval model performed at 85–95% of peak heart rate. Alternative interval formats, including short-interval or sprint-based protocols, were also considered eligible provided they met intensity criteria. The comparator condition required structured moderate-intensity continuous training prescribed at approximately 50–75% of peak heart rate, heart rate reserve, or VO_2peak, delivered continuously without high-intensity intervals. Studies comparing HIIT to usual care, no exercise, or non-standardized physical activity advice without a direct MICT arm were excluded from quantitative synthesis to preserve internal validity of direct comparisons.

The primary outcomes were change in peak oxygen consumption (peak VO_2), measured in mL/kg/min using cardiopulmonary exercise testing, and change in LVEF, expressed as percentage points and assessed via echocardiography or other validated imaging modalities. Secondary outcomes included change in six-minute walk test distance, ventilatory efficiency (VE/VCO_2 slope), peak workload, and incidence of serious adverse cardiovascular events. To ensure sufficient exposure to the intervention, studies were required to have an intervention duration of at least four weeks.

2.3 Information Sources and Search Strategy

A comprehensive and systematic search strategy was developed to identify all relevant studies evaluating the comparative effects of high-intensity interval training (HIIT) and moderate-intensity continuous training (MICT) in older adults with heart failure. Electronic searches were conducted in the following databases: MEDLINE (via PubMed), Embase, Cochrane Central Register of Controlled Trials (CENTRAL), Scopus, and Web of Science. These databases were selected to ensure broad coverage of biomedical, cardiovascular, rehabilitation, and exercise science literature.

The search strategy combined controlled vocabulary terms (e.g., MeSH and Emtree terms) and free-text keywords related to: (1) heart failure (“heart failure,” “HFrEF,” “HFpEF,” “cardiac failure”); (2) exercise interventions (“high-intensity interval training,” “HIIT,” “interval training,” “moderate-intensity continuous training,” “MICT,” “aerobic training,” “exercise therapy”); and (3) outcomes (“peak oxygen consumption,” “VO2,” “VO2max,” “cardiopulmonary exercise testing,” “left ventricular ejection fraction,” “LVEF,” “exercise capacity”). Boolean operators (AND, OR) were used to appropriately combine search terms, and filters were applied to identify randomized controlled trials where applicable.

The search covered all records from database inception to the predefined final search date, which was documented prior to data synthesis to ensure transparency and reproducibility. No language restrictions were imposed. In addition to electronic database searches, reference lists of included trials and relevant systematic reviews were manually screened to identify potentially

eligible studies not captured in the initial search. Grey literature sources and trial registries (e.g., ClinicalTrials.gov) were also reviewed to reduce the risk of publication bias.

Conference abstracts were considered eligible if sufficient methodological detail and outcome data were available for extraction and risk-of-bias assessment. The full search strategy for each database, including search strings and applied filters, is provided in the Supplementary Appendix to facilitate reproducibility.

2.4 Study Selection Process

All retrieved records were exported into a reference management system, and duplicate entries were removed. Study selection proceeded in two phases. First, two independent reviewers screened titles and abstracts against predefined eligibility criteria. Second, full-text articles were obtained for all potentially eligible studies and independently assessed for inclusion. Disagreements were resolved through discussion and consensus; when necessary, a third reviewer adjudicated unresolved discrepancies. A PRISMA 2020 flow diagram was constructed to document the number of records identified, screened, excluded (with explicit reasons), and included in qualitative and quantitative synthesis.

2.5 Data Extraction

Data extraction was performed independently by two reviewers using a standardized, pilot-tested extraction form. Extracted data included study design, setting, sample size, participant characteristics (including age, sex distribution, heart failure phenotype, and New York Heart Association functional class), baseline exercise capacity and LVEF, detailed HIIT and MICT protocol characteristics (intensity prescription, interval structure, session duration, weekly frequency, total intervention duration, supervision level, and progression strategy), adherence rates, completion rates, and reported adverse events. Outcome data were extracted as change-from-baseline values whenever available; when only post-intervention values were reported, these were recorded with appropriate notation. Where key statistics were missing or unclear, attempts were made to contact corresponding authors for clarification. Extracted data were cross-checked for accuracy, and discrepancies were resolved by consensus review of the source text.

2.6 Risk of Bias Assessment

Risk of bias for randomized controlled trials was assessed using the Cochrane Risk of Bias 2 (RoB 2) tool. Five domains were evaluated for each trial: bias arising from the randomization process; bias due to deviations from intended interventions; bias due to missing outcome data; bias in measurement of the outcome; and bias in selection of the reported result. Each domain was rated as low risk, some concerns, or high risk based on responses to signaling questions specified in the RoB 2 framework. An overall risk-of-bias judgment was assigned for each study according to Cochrane guidance. Particular attention was given to deviations from intended interventions, given the inherent challenges of intensity adherence in exercise trials, and to measurement bias, although peak VO_2 and LVEF were considered objective outcomes with low susceptibility to observer bias when assessed using standardized procedures. Risk-of-bias assessment was performed independently by two reviewers, with disagreements resolved through discussion.

2.7 Data Synthesis and Statistical Analysis

Quantitative synthesis was undertaken when at least three clinically and methodologically homogeneous trials reported comparable outcomes. Continuous outcomes were pooled using mean differences (MD) with 95% confidence intervals when outcomes were measured on the same scale. Standardized mean differences were considered if outcome scales differed across studies. Given anticipated clinical heterogeneity in training protocols, supervision levels, and patient characteristics, a random-effects model was employed as the primary analytic approach.

Statistical analyses were conducted using validated meta-analytic software.

2.8 Assessment of Heterogeneity

Statistical heterogeneity was assessed using the I^2 statistic, τ^2, and the chi-square test for heterogeneity, with a significance threshold of $p<0.10$ for the latter. I^2 values were interpreted as low (<30%), moderate (30–60%), substantial (60–90%), or considerable (>90%) heterogeneity. Where substantial heterogeneity was observed, potential sources were explored through subgroup and sensitivity analyses. Clinical heterogeneity was also evaluated qualitatively, particularly regarding heart failure phenotype, baseline functional capacity, intervention duration, and protocol structure.

2.9 Subgroup Analyses

Pre-specified subgroup analyses were conducted according to heart failure phenotype (HFrEF versus HFpEF), intervention duration (<12 weeks versus ≥12 weeks), interval structure (long-interval versus short-interval HIIT), and supervision level (fully supervised versus partially home-based programs). These subgroup analyses were interpreted cautiously and regarded as exploratory rather than confirmatory.

2.10 Sensitivity Analyses

Sensitivity analyses were performed to assess robustness of pooled estimates. This included exclusion of trials judged to be at high risk of bias, exclusion of small trials with fewer than 30 participants per arm, and comparison of fixed-effect and random-effects models to evaluate the influence of analytic assumptions on effect estimates.

2.11 Publication Bias

Publication bias was assessed when at least ten studies were included in a meta-analysis. Funnel plots were visually inspected for asymmetry, and Egger's regression test was conducted where appropriate. Asymmetry was interpreted cautiously in the context of clinical heterogeneity and sample size variation.

2.12 Certainty of Evidence (GRADE)

The certainty of evidence for each primary outcome was evaluated using the GRADE framework. Evidence from randomized trials began as high certainty and was downgraded based on risk of bias, inconsistency, indirectness, imprecision, and publication bias. Downgrading decisions were made transparently and justified in accompanying Summary of Findings tables. Certainty was categorized as high, moderate, low, or very low, reflecting confidence in the estimated effect and the likelihood that future research would alter the estimate.

All methodological decisions were made prior to outcome interpretation to minimize bias. Throughout the review process, transparency, reproducibility, and adherence to Cochrane standards were prioritized to ensure methodological rigor appropriate for high-impact cardiology publication.

3. Results

A total of 500 records were identified through the systematic search strategy. Following screening and eligibility assessment, 25 sources were included in the qualitative synthesis, comprising 15 primary randomized controlled trials (RCTs), 8 systematic reviews or meta-analyses, and 2 secondary sub analyses derived from a large multicenter trial. The included studies spanned a broad temporal range and demonstrated substantial heterogeneity in sample size, heart failure phenotype, intervention duration, supervision intensity, and outcome assessment methods.

The majority of primary RCTs enrolled patients with heart failure with reduced ejection fraction (HFrEF), whereas a smaller subset specifically targeted heart failure with preserved ejection fraction (HFpEF) (Ellingsen et al., 2017; Mueller et al., 2021). Across studies, mean participant age ranged from approximately 59 to 75 years, with

most cohorts clustered in the mid-to-late 60s. Baseline peak oxygen consumption (peak VO_2) values were generally consistent with moderate functional impairment, although inter-study variability reflected differences in heart failure severity, New York Heart Association functional class, and inclusion criteria (Wisløff et al., 2007; Haykowsky et al., 2013).

Intervention duration most commonly ranged from 8 to 12 weeks. However, the OptimEx-Clin trial incorporated a 12-month follow-up period with a partially home-based training component (Mueller et al., 2021), and long-term observational follow-up data from the Generation 100 study extended to five years (Stensvold et al., 2024). HIIT protocols most frequently employed the 4×4-minute interval model performed at 85–95% of peak heart rate, interspersed with active recovery at approximately 60–70% of peak heart rate (Ellingsen et al., 2017; Wisløff et al., 2007). In contrast, MICT protocols typically consisted of continuous aerobic exercise prescribed at 60–75% of peak heart rate for 30–47 minutes per session, reflecting conventional cardiac rehabilitation practice.

3.1 Exercise Capacity Outcomes

Peak oxygen uptake (peak VO_2) was the most consistently reported primary outcome across randomized controlled trials. Early single-center studies demonstrated substantial improvements favoring high-intensity interval training (HIIT). The landmark trial by Wisløff et al. (2007) reported a 46% increase in peak VO_2 following 12 weeks of HIIT compared with a 14% increase in the moderate-intensity continuous training (MICT) group ($p < .001$). The magnitude of improvement observed in this study was considerably larger than that typically reported in subsequent trials and was accompanied by significant improvements in left ventricular ejection fraction (LVEF) and indices of reverse ventricular remodeling.

Donelli da Silveira et al. (2020) similarly reported significant improvements favoring HIIT in patients with heart failure with preserved ejection fraction (HFpEF), with approximately a 22% increase in peak VO_2 compared with 11% in the MICT group ($p < .001$). In a small pilot study of HFpEF patients, Angadi et al. (2015) observed that HIIT increased peak VO_2 from 19.2 to 21.0 mL/kg/min, whereas MICT produced no statistically significant change. Although these early trials suggested superiority of HIIT, their small sample sizes generally fewer than 30 participants per group limit statistical precision and raise the possibility of small-study effects.

In contrast, the largest multicenter randomized trial in HFrEF, SMARTEX-HF (n = 261), found no statistically significant difference between HIIT and MICT for change in peak VO_2 at 12 weeks. The reported between-group difference was −0.4 mL/kg/min (95% CI −1.7 to 0.8; $p = .70$) (Ellingsen et al., 2017). Both exercise modalities improved peak VO_2 relative to baseline, but superiority of HIIT was not demonstrated. Notably, adherence analyses revealed substantial intensity contamination: approximately 51% of participants randomized to HIIT trained below the prescribed high-intensity threshold, while nearly 80% of those assigned to MICT exceeded their prescribed moderate-intensity targets. This bidirectional deviation likely narrowed the effective exposure differential and attenuated measurable between-group differences.

Similarly, the OptimEx-Clin trial in HFpEF (n = 180) failed to demonstrate superiority of HIIT over MICT for peak VO_2 at three months (Mueller et al., 2021). The between-group difference was −0.4 mL/kg/min (95% CI −1.4 to 0.6), and neither intervention achieved the prespecified minimal clinically important difference of 2.5 mL/kg/min relative to guideline-based advice. A secondary mechanistic analysis adjusting for total energy expenditure showed that the adjusted difference between HIIT and MICT was only 0.09 mL/kg/min ($p = .98$), suggesting that cumulative exercise volume rather than intensity per se may drive physiological adaptation in HFpEF populations (Mueller et al., 2022).

Meta-analytic syntheses have generally favored HIIT for improving peak VO_2. Xie et al. (2017) reported a pooled mean difference of 1.76 mL/kg/min (95% CI 1.06–2.46; $p < .001$) favoring HIIT across cardiac populations, including heart failure cohorts. Haykowsky et al. (2013) reported a weighted mean difference of 2.14 mL/kg/min (95% CI 0.66–3.63) in HFrEF populations. Trejos-Montoya et al. (2019) also demonstrated pooled superiority of HIIT. More recent meta-analyses by Yang et al. (2024) and Fauzi (2025) reported pooled mean differences of approximately 1.78 mL/kg/min favoring HIIT, with statistically significant results. However, these pooled estimates were largely driven by smaller single-center trials, and statistical heterogeneity was moderate to substantial in several models, reflecting variability in study design, intervention fidelity, and patient characteristics.

Secondary functional outcomes demonstrated broadly similar patterns. Six-minute walk distance (6MWT) showed significant pooled improvements favoring HIIT in some meta-analyses, with Yang et al. (2024) reporting a mean difference of 28.13 meters (95% CI 14.56–41.70). Ventilatory efficiency, as assessed by the VE/VCO_2 slope, did not consistently differ between modalities across studies (Donelli da Silveira et al., 2020; Xie et al., 2017). Improvements in anaerobic threshold were reported in selected trials but were not uniformly assessed or consistently replicated.

3.2 Left Ventricular Ejection Fraction

Left ventricular ejection fraction (LVEF) was reported primarily in trials enrolling patients with heart failure with reduced ejection fraction (HFrEF). The seminal study by Wisløff et al. (2007) documented a 10-percentage-point absolute increase in LVEF following 12 weeks of high-intensity interval training (HIIT), compared with minimal change in the moderate-intensity continuous training (MICT) group ($p < .01$). This improvement was accompanied by evidence of reverse ventricular remodeling and enhanced stroke volume, contributing to early enthusiasm regarding the potential central cardiac benefits of intensity-driven training.

However, subsequent trials have not consistently replicated these pronounced effects. In the large multicenter SMARTEX-HF trial, no statistically significant between-group difference in LVEF was observed at 12 weeks, with a reported mean difference of 1.5% (95% CI −2.1 to 5.1) between HIIT and MICT (Ellingsen et al., 2017). Similarly, Benda et al. (2015) reported no significant difference between training modalities in LVEF change.

Meta-analytic findings have been mixed. Haykowsky et al. (2013) reported a non-significant weighted mean difference of 3.29% (95% CI −0.7 to 7.28) favoring HIIT. Xie et al. (2017) likewise observed a non-significant trend toward greater LVEF improvement with HIIT. In contrast, more recent meta-analyses by Yang et al. (2024) and Fauzi (2025) reported pooled mean differences of approximately 3.13% favoring HIIT, reaching statistical significance. Notably, these pooled estimates were influenced by inclusion of smaller single-center trials that reported comparatively large effect sizes, potentially contributing to heterogeneity and small-study effects.

In heart failure with preserved ejection fraction (HFpEF) populations, LVEF generally remained unchanged across trials, consistent with preserved baseline systolic function. Neither the OptimEx-Clin trial nor smaller HFpEF studies demonstrated superiority of HIIT over MICT in altering LVEF (Mueller et al., 2021; Angadi et al., 2015).

3.3 Safety and Adverse Events

Safety outcomes across the included randomized controlled trials were generally reassuring. No consistent excess of serious adverse cardiovascular events was observed in participants assigned to high-intensity interval training (HIIT) compared with those undergoing moderate-intensity continuous training (MICT). In the multicenter SMARTEX-HF trial, no statistically significant differences in serious adverse events

were detected between intervention groups during the supervised training phase, although a numerical trend toward more cardiovascular events in the HIIT arm during extended follow-up was noted (Ellingsen et al., 2017).

Similarly, the OptimEx-Clin trial in patients with heart failure with preserved ejection fraction reported comparable rates of acute coronary syndromes and other serious adverse events across HIIT, MICT, and guideline-based physical activity advice groups (Mueller et al., 2021). Long-term observational data from the Generation 100 study, which included older adults with cardiovascular disease, demonstrated no acute cardiovascular events directly attributable to HIIT over five years of follow-up (Stensvold et al., 2024).

Meta-analytic syntheses have likewise concluded that HIIT does not appear to increase the risk of serious adverse events compared with MICT in clinically stable heart failure populations (Fauzi, 2025; Hannan et al., 2018). Nevertheless, it is important to acknowledge that most trials were not powered primarily to detect rare but potentially catastrophic events, and absolute event rates were low. Furthermore, HIIT in these studies was delivered within structured, supervised rehabilitation settings with heart rate monitoring and gradual progression protocols.

The Generation 100 substudy involving older adults with cardiovascular disease reported no acute cardiovascular events attributable to HIIT over five years of follow-up [9]. Meta-analytic data similarly suggested no increased risk associated with HIIT compared with MICT [2,20]. Overall, both modalities appear safe when delivered to stable, appropriately screened patients under supervision.

3.4 Risk of Bias and Study Quality Considerations

Assessment of methodological quality revealed variability across included studies. Early single-center randomized trials frequently provided limited detail regarding allocation concealment procedures and did not implement participant or personnel blinding, which is inherently challenging in exercise-based interventions. However, the use of objective physiological outcomes such as peak oxygen uptake (peak VO_2) measured via cardiopulmonary exercise testing and left ventricular ejection fraction (LVEF) assessed through standardized imaging reduces susceptibility to measurement bias relative to subjective endpoints. Several smaller trials were also limited by modest sample sizes and were likely underpowered to detect clinically meaningful between-group differences (Benda et al., 2015).

In contrast, larger multicenter trials demonstrated stronger methodological rigor with clearly defined randomization procedures, centralized outcome assessment protocols, and predefined statistical analysis plans. Nevertheless, pragmatic implementation introduced other challenges. In SMARTEX-HF, substantial deviations from prescribed intensity targets were observed in both the HIIT and MICT arms, resulting in bidirectional intensity contamination and reduced intervention contrast (Ellingsen et al., 2017). Such deviations represent a potential source of performance bias and complicate interpretation of true treatment effects in real-world settings.

The overall pattern of findings suggests the possibility of small-study effects. Larger effect sizes were frequently reported in smaller single-center trials, whereas attenuation of between-group differences was observed in larger, methodologically rigorous multicenter studies. This discrepancy raises concern regarding potential overestimation of treatment effects in early trials. Funnel plot asymmetry, as reported in certain meta-analyses, further suggested the possibility of publication bias or selective reporting, particularly where small trials disproportionately contributed to pooled estimates (Trejos-Montoya et al., 2019).

Taken together, these methodological considerations underscore the importance of interpreting pooled superiority findings with

caution and highlight the need to weigh large, pragmatic multicenter evidence more heavily in determining the comparative effectiveness of HIIT and MICT in heart failure populations.

3.5 Subgroup and Sensitivity Findings

Pre-specified subgroup analyses revealed phenotype-specific differences in treatment effects. Pooled superiority of high-intensity interval training (HIIT) over moderate-intensity continuous training (MICT) was more consistently observed in heart failure with reduced ejection fraction (HFrEF) populations, particularly in smaller single-center trials and meta-analytic syntheses (Haykowsky et al., 2013; Xie et al., 2017; Yang et al., 2024). In contrast, among patients with heart failure with preserved ejection fraction (HFpEF), both modalities produced modest improvements in exercise capacity without clear evidence of superiority of HIIT. Findings from the OptimEx-Clin trial and its mechanistic subanalysis demonstrated no significant between-group difference in peak VO_2 once total energy expenditure was considered, reinforcing the hypothesis that cumulative workload may be more influential than peak interval intensity in HFpEF populations (Mueller et al., 2021; Mueller et al., 2022).

Analyses stratified by intervention duration suggested that programs lasting approximately 7–12 weeks were associated with larger short-term improvements in peak VO_2, consistent with the time frames used in most randomized trials (Hannan et al., 2018). However, longer-term follow-up data, such as those from the Generation 100 study, indicated that sustained differences between high- and moderate-intensity training were less pronounced over extended periods (Stensvold et al., 2024). These findings suggest that early gains may diminish over time or that long-term adherence and cumulative training exposure may outweigh initial intensity differences.

Sensitivity analyses excluding smaller trials with limited sample sizes resulted in attenuation of pooled effect estimates, supporting the presence of small-study effects and reinforcing the influence of larger, methodologically rigorous trials on overall conclusions.

In summary, the aggregate evidence indicates that HIIT may confer modest improvements in peak VO_2 and possibly LVEF in selected HFrEF populations. However, these advantages are not consistently reproduced in large multicenter trials and appear attenuated when methodological rigor, adherence fidelity, and real-world implementation factors are considered. In HFpEF populations, current evidence does not demonstrate superiority of HIIT over MICT, and physiological adaptation appears more closely related to total energy expenditure than intensity alone. When delivered under appropriate supervision in clinically stable patients, both exercise modalities demonstrate favorable safety profiles.

4. Results

4.1 Study Selection and Characteristics

The systematic search identified 500 records, of which 16 comparative randomized controlled trials met the predefined inclusion criteria for quantitative and qualitative synthesis. The included studies primarily enrolled patients with heart failure with reduced ejection fraction (HFrEF), while a smaller subset specifically investigated heart failure with preserved ejection fraction (HFpEF) populations (Ellingsen et al., 2017; Mueller et al., 2021). Sample sizes ranged from small single-center trials, typically including approximately 20–40 participants per study (Angadi et al., 2015; Benda et al., 2015), to large multicenter randomized trials, most notably SMARTEX-HF (n = 261) (Ellingsen et al., 2017) and OptimEx-Clin (n = 180) (Mueller et al., 2021).

Intervention duration in most trials ranged from 7 to 12 weeks, reflecting short-term supervised cardiac rehabilitation programs. However, extended follow-up data were available from the Generation 100 study, which examined long-term outcomes over a five-year period in older adults

with cardiovascular disease (Stensvold et al., 2024).

HIIT protocols most commonly employed the 4×4-minute interval model performed at 85–95% of peak heart rate, interspersed with active recovery phases at approximately 60–70% of peak heart rate (Wisløff et al., 2007; Ellingsen et al., 2017). In contrast, MICT protocols typically prescribed continuous aerobic exercise at 60–75% of peak heart rate for durations of 30–47 minutes per session, consistent with conventional cardiac rehabilitation guidelines. Across studies, peak oxygen uptake (peak VO_2) was assessed using objective cardiopulmonary exercise testing (CPET), and left ventricular ejection fraction (LVEF) was measured using standardized echocardiographic or other cardiac imaging modalities, minimizing measurement bias for primary physiological outcomes.

4.2 Primary Outcome: Peak Oxygen Uptake (Peak VO_2)

Across pooled meta-analyses, HIIT demonstrated a modest but statistically significant improvement in peak VO_2 compared with MICT in HFrEF populations (Table 1).

Table 1. Pooled Effect of HIIT vs MICT on Peak VO_2 in HFrEF

Study (Meta-Analysis)	No. of RCTs	Total Participants	Pooled MD (mL/kg/min)	95% CI	P-value
Haykowsky et al. 2013 [18]	7	~200	2.14	0.66 to 3.63	<0.01
Xie et al. 2017 [3]	21 (cardiac; HF subset)	736	1.76	1.06 to 2.46	<0.001
Trejos-Montoya et al. 2019 [6]	12	418	Favored HIIT		<0.05
Yang et al. 2024 [1]	13	513	1.78	0.80 to 2.76	0.0004
Fauzi 2025 [2]	13	411	1.78	0.80 to 2.76	<0.01

In patients with heart failure with reduced ejection fraction (HFrEF), pooled evidence from multiple meta-analyses suggests that high-intensity interval training (HIIT) confers a modest but statistically significant improvement in peak oxygen uptake (peak VO_2) compared with moderate-intensity continuous training (MICT). Across syntheses, the estimated mean difference consistently ranges between approximately 1.7 and 2.1 mL/kg/min in favor of HIIT.

Haykowsky et al. (2013), in one of the earliest focused meta-analyses of aerobic interval training in HFrEF, reported a pooled weighted mean difference of 2.14 mL/kg/min (95% CI 0.66 to 3.63; $p < .01$) across seven randomized trials comprising approximately 200 participants. This analysis suggested a clinically meaningful superiority of interval-based training for improving aerobic capacity and provided early quantitative support for the HIIT paradigm.

Xie et al. (2017), in a broader meta-analysis including cardiac populations with a heart failure subset, reported a pooled mean difference of 1.76 mL/kg/min (95% CI 1.06 to 2.46; $p < .001$). Although this synthesis included a larger total sample (n = 736 across 21 trials), the HFrEF subgroup remained largely influenced by smaller single-center studies. Similarly, Trejos-Montoya et al. (2019) analyzed 12 trials (n = 418) and reported a statistically significant pooled effect favoring HIIT, although precise mean differences varied across models.

More recent meta-analyses have yielded highly consistent pooled estimates. Yang et al. (2024), analyzing 13 trials (n = 513), reported a pooled mean difference of 1.78 mL/kg/min (95% CI 0.80 to 2.76; $p = .0004$), while Fauzi (2025), using a similar trial base (n = 411), reported an identical pooled estimate of 1.78 mL/kg/min (95% CI 0.80 to 2.76; $p < .01$). The consistency of these pooled point

estimates across independent analyses strengthens confidence in the direction of effect.

From a clinical standpoint, an absolute improvement of approximately 1.7–2.1 mL/kg/min in peak VO_2 represents a modest increment in aerobic capacity. Given that peak VO_2 is a powerful prognostic marker in HFrEF, even small improvements may carry incremental prognostic implications. However, the magnitude of benefit observed in pooled analyses is substantially smaller than the dramatic relative increases reported in early mechanistic trials, highlighting an attenuation of effect size when evidence is aggregated across broader populations and settings.

Importantly, the largest and most methodologically rigorous multicenter randomized controlled trial SMARTEX-HF (n = 261) did not demonstrate a statistically significant between-group difference in peak VO_2. The reported mean difference was −0.4 mL/kg/min (95% CI −1.7 to 0.8; p = .70) in favor of MICT, effectively crossing the null and indicating no superiority of HIIT (Ellingsen et al., 2017). This neutral result contrasts sharply with pooled meta-analytic estimates and represents a pivotal finding in the evidence base.

The discrepancy between pooled meta-analyses and the large pragmatic trial likely reflects several interacting factors. First, many pooled analyses are disproportionately influenced by smaller single-center studies reporting larger effect sizes. Second, adherence variability and intensity contamination observed in SMARTEX-HF may have reduced effective between-group separation, thereby attenuating measurable differences. Third, multicenter pragmatic designs introduce greater heterogeneity in patient characteristics, supervision intensity, and implementation fidelity compared with tightly controlled mechanistic trials.

This divergence between small-trial superiority and large-trial neutrality contributed directly to statistical heterogeneity in pooled models and informed downgrading decisions in the GRADE assessment for inconsistency and potential small-study effects. Taken together, while aggregated data suggest a modest physiological advantage of HIIT in HFrEF populations, the robustness and generalizability of this superiority remain uncertain when higher-weight evidence from large multicenter trials is considered.

In contrast to findings observed in heart failure with reduced ejection fraction (HFrEF), no statistically significant superiority of high-intensity interval training (HIIT) over moderate-intensity continuous training (MICT) was demonstrated in patients with heart failure with preserved ejection fraction (HFpEF). The most robust evidence in this phenotype derives from the OptimEx-Clin randomized clinical trial, which enrolled 180 patients with HFpEF and directly compared HIIT with MICT and guideline-based physical activity advice (Mueller et al., 2021). At three months, the between-group mean difference in peak oxygen uptake (peak VO_2) between HIIT and MICT was −0.4 mL/kg/min (95% CI −1.4 to 0.6), indicating no statistically significant or clinically meaningful superiority of HIIT. Notably, neither intervention achieved the prespecified minimal clinically important difference of 2.5 mL/kg/min relative to usual care, further underscoring the modest magnitude of training effects in this population.

A subsequent mechanistic subanalysis of the OptimEx-Clin cohort explored the relationship between achieved training characteristics and physiological adaptation (Mueller et al., 2022). After adjusting for total energy expenditure and actual training volume, the adjusted difference in peak VO_2 between HIIT and MICT was only 0.09 mL/kg/min (p = .98), effectively demonstrating equivalence between modalities. This finding is particularly important, as it suggests that cumulative workload rather than interval intensity per se may be the principal driver of aerobic adaptation in HFpEF.

These results are physiologically plausible given the pathophysiological characteristics of HFpEF. Unlike HFrEF, where systolic dysfunction may

respond to intensity-driven central adaptations, HFpEF is characterized by impaired ventricular compliance, microvascular dysfunction, skeletal muscle abnormalities, and reduced peripheral oxygen extraction. In such a context, improvements in aerobic capacity may depend more heavily on total training exposure, peripheral conditioning, and metabolic adaptations than on peak interval intensity.

The absence of superiority in HFpEF across both primary and adjusted analyses strengthens confidence in the conclusion that HIIT does not confer incremental benefit over MICT in this phenotype. Consequently, phenotype-specific stratification becomes essential when interpreting pooled exercise data. In HFpEF populations, exercise prescription may reasonably prioritize feasibility, adherence, and sustainable volume accumulation rather than high-intensity stimulus alone.

4.3 Secondary Outcome: Left Ventricular Ejection Fraction (LVEF)

In HFrEF populations, pooled analyses suggested small improvements in LVEF with HIIT, though results were inconsistent (Table 3).

Table 2. Pooled Effect of HIIT Versus MICT on Peak VO_2 in HfpEF

Study	Total Participants	Effect Measure	Pooled Effect	95% CI	Interpretation
OptimEx-Clin 2021 [5]	180	Mean Difference	−0.4 mL/kg/min	−1.4 to 0.6	Not significant
Mueller et al. 2022 (adjusted) [8]	91	Adjusted Difference	0.09 mL/kg/min		No difference

Table 3. Pooled Effect of HIIT vs MICT on LVEF (HFrEF)

Study (Meta-Analysis)	No. of RCTs	Total Participants	Effect Measure	Pooled Effect (MD, %)	95% CI	P-value
Haykowsky et al. 2013 [18]	7		Mean Difference	3.29	−0.7 to 7.28	NS
Xie et al. 2017 [3]			Mean Difference	2.18	−0.54 to 4.90	0.12
Yang et al. 2024 [1]	13	513	Mean Difference	3.13	1.25 to 5.02	0.001
Fauzi 2025 [2]	13	411	Mean Difference	3.13	1.25 to 5.02	<0.01
Trejos-Montoya et al. 2019 [6]	12	418	Within-group improvement	HIIT +6.4%	3.7 to 9.1	<0.001

Table 4. Pooled Effect of HIIT Versus MICT on 6-Minute Walk Distance

Study	Effect (MD, meters)	95% CI	P-value
Yang et al. 2024 [1]	28.13 m	14.56 to 41.70	<0.01
Fauzi 2025 [2]	28.13 m	14.56 to 41.70	<0.01

In patients with heart failure with reduced ejection fraction (HFrEF), pooled analyses suggest that high-intensity interval training (HIIT) may confer small improvements in left ventricular ejection fraction (LVEF) compared with moderate-intensity continuous training (MICT); however, results are inconsistent and highly sensitive to trial selection and methodological rigor.

Early mechanistic work provided the initial signal for central remodeling benefits. Wisløff et al. (2007) reported an approximate 10-percentage-point absolute increase in LVEF following 12

weeks of HIIT, accompanied by reverse ventricular remodeling. This large effect size substantially influenced subsequent perceptions regarding the structural superiority of interval-based training.

When evidence was aggregated across randomized trials, the magnitude of pooled LVEF improvement became more modest. Haykowsky et al. (2013) reported a weighted mean difference of 3.29% (95% CI −0.7 to 7.28), which did not reach statistical significance. Similarly, Xie et al. (2017) observed a non-significant pooled mean difference of 2.18% (95% CI −0.54 to 4.90; p = .12), suggesting that although point estimates favored HIIT, confidence intervals crossed the null.

More recent meta-analyses reported statistically significant pooled effects. Yang et al. (2024), analyzing 13 randomized trials (n = 513), reported a pooled mean difference of 3.13% (95% CI 1.25 to 5.02; p = .001) favoring HIIT. Fauzi (2025), evaluating a similar trial base (n = 411), reported an identical pooled mean difference of 3.13% (95% CI 1.25 to 5.02; p < .01). Trejos-Montoya et al. (2019) reported a significant within-group improvement in LVEF in the HIIT arms (+6.4%; 95% CI 3.7 to 9.1; p < .001), although direct between-group contrasts were more variable across included studies.

Despite these pooled findings, the largest multicenter randomized controlled trial SMARTEX-HF (n = 261) found no statistically significant between-group difference in LVEF at 12 weeks (Ellingsen et al., 2017). The absence of superiority in this well-powered pragmatic trial introduces important inconsistency into the evidence base and materially influences overall certainty of evidence.

The discrepancy between early single-center trials demonstrating large structural improvements and later multicenter trials showing neutrality likely reflects several factors, including small-study effects, heterogeneity in training fidelity, differences in imaging methodology, and variation in baseline systolic function. Moreover, absolute LVEF changes of approximately 2–3% even when statistically significant may represent modest structural effects whose clinical implications remain uncertain.

In heart failure with preserved ejection fraction (HFpEF) populations, LVEF remained largely unchanged across trials, as expected given preserved baseline systolic function. Neither the OptimEx-Clin trial nor smaller HFpEF studies demonstrated superiority of HIIT over MICT in altering LVEF (Mueller et al., 2021; Angadi et al., 2015).

Overall, while pooled analyses suggest small improvements in LVEF with HIIT in HFrEF populations, these findings are inconsistent and attenuated in large multicenter trials. The variability in magnitude and statistical significance across analyses contributed to downgrading for inconsistency and imprecision in the GRADE assessment and supports cautious interpretation of structural superiority claims.

Beyond peak oxygen uptake, functional capacity assessed by the six-minute walk test (6MWT) was reported in several trials and incorporated into pooled analyses. The 6MWT reflects submaximal exercise performance and has direct clinical relevance, as it correlates with daily functional ability and prognosis in heart failure populations.

Pooled meta-analytic data suggest that high-intensity interval training (HIIT) confers a modest improvement in six-minute walk distance compared with moderate-intensity continuous training (MICT). Yang et al. (2024) reported a pooled mean difference of 28.13 meters (95% CI 14.56 to 41.70; p < .01) favoring HIIT. A similar pooled estimate was reported by Fauzi (2025), with an identical mean difference of 28.13 meters (95% CI 14.56 to 41.70; p < .01). These findings indicate statistical superiority of HIIT in improving submaximal functional capacity.

From a clinical perspective, the magnitude of improvement approximately 28 meters approaches commonly cited minimal clinically important difference (MCID) thresholds for heart failure populations, which are generally estimated

in the range of 25–45 meters. However, the upper and lower bounds of the confidence intervals suggest variability in the precision of this estimate. Furthermore, as with peak VO_2 outcomes, pooled improvements in 6MWT are influenced by smaller single-center trials, and heterogeneity in baseline functional status, intervention duration, and supervision intensity may contribute to variation in effect size.

Importantly, both HIIT and MICT were associated with improvements relative to baseline in most trials, reinforcing that structured aerobic training irrespective of intensity enhances functional capacity in clinically stable heart failure patients. The incremental benefit of HIIT over MICT in 6MWT performance, while statistically significant in pooled analyses, remains modest in absolute terms and should be interpreted within the broader context of adherence feasibility and phenotype-specific responsiveness.

4.5 Safety Outcomes

Across the 16 comparative randomized studies included in this review, serious adverse events were infrequent and broadly comparable between high-intensity interval training (HIIT) and moderate-intensity continuous training (MICT). No consistent signal of increased cardiovascular risk associated with HIIT was observed during supervised intervention phases.

In the large multicenter SMARTEX-HF trial, no statistically significant excess of serious cardiovascular events was detected in the HIIT group compared with MICT during the 12-week intervention period (Ellingsen et al., 2017). Although a numerical trend toward more cardiovascular events in the HIIT arm was noted during extended follow-up, this difference did not reach statistical significance and was interpreted within the context of adherence variability and intensity contamination.

Similarly, the OptimEx-Clin trial in heart failure with preserved ejection fraction (HFpEF) reported comparable rates of acute coronary syndromes and other serious adverse events across HIIT, MICT, and guideline-based physical activity advice groups (Mueller et al., 2021). These findings suggest that, within structured and monitored rehabilitation settings, higher-intensity interval protocols do not confer excess short-term cardiovascular risk relative to moderate-intensity continuous exercise.

Long-term observational data from the Generation 100 study, which included older adults with cardiovascular disease, likewise did not demonstrate excess acute cardiovascular complications attributable to HIIT over a five-year follow-up period (Stensvold et al., 2024). In addition, pooled meta-analytic assessments have reported no statistically significant increase in serious adverse events with HIIT compared with MICT in heart failure populations (Fauzi, 2025; Hannan et al., 2018).

Nevertheless, several important caveats must be acknowledged. In most trials, safety outcomes were secondary endpoints rather than primary objectives, and event rates were low. Consequently, studies were not powered to detect rare but clinically significant adverse events. Moreover, HIIT interventions were delivered under structured supervision, with heart rate monitoring and gradual progression protocols. Therefore, while available evidence supports the short-term safety of HIIT in clinically stable, appropriately screened patients, caution is warranted when extrapolating these findings to unsupervised or less-controlled settings.

4.6 Risk of Bias Summary

Risk-of-bias assessment, conducted using the Cochrane Risk of Bias 2 (RoB 2) framework, demonstrated variability in methodological quality across included trials. Overall, most studies were judged to have either low risk of bias or “some concerns,” with one large multicenter trial rated as high risk in the domain of deviations from intended interventions.

Large multicenter trials, particularly SMARTEX-HF, demonstrated robust methodological features, including centralized randomization procedures,

clearly defined allocation processes, prespecified outcomes, and objective outcome assessment using cardiopulmonary exercise testing (CPET) and standardized echocardiographic imaging (Ellingsen et al., 2017). These strengths reduced the likelihood of bias arising from the randomization process and outcome measurement. However, substantial deviations from prescribed intensity targets were documented in both intervention arms. A significant proportion of participants assigned to HIIT failed to achieve target high-intensity thresholds, while many in the MICT arm exceeded their prescribed intensity. This bidirectional intensity contamination narrowed the effective contrast between groups and led to a judgment of high risk of bias in the domain of deviations from intended interventions for this trial.

In contrast, smaller single-center trials often demonstrated adequate reporting of outcome measurement but provided limited detail regarding allocation concealment and sequence generation procedures. In several studies, the absence of explicit reporting of concealment mechanisms resulted in ratings of "some concerns" in the randomization domain. Blinding of participants and personnel was inherently not feasible in exercise interventions; however, the primary outcomes peak VO_2 and LVEF were objective and instrument-based, mitigating concerns regarding detection bias.

Missing outcome data were generally low across trials, with most studies reporting high completion rates. Selective reporting bias was not strongly evident, although smaller trials with limited protocols occasionally lacked prospective registration details, contributing to some uncertainty.

Collectively, the overall risk-of-bias profile suggests that while methodological rigor was generally acceptable particularly in larger multicenter trials performance-related issues such as adherence variability and intensity contamination represent important threats to internal validity. These factors were incorporated into GRADE downgrading decisions, particularly for outcomes in which large pragmatic trials diverged from pooled small-study estimates.

4.7 GRADE Summary of Findings

The integrated evidence suggests that high-intensity interval training (HIIT) may confer modest improvements in peak oxygen uptake (peak VO_2) in patients with heart failure with reduced ejection fraction (HFrEF). However, these effects are attenuated when the largest, methodologically rigorous multicenter trial is considered, which demonstrated no statistically significant superiority over moderate-intensity continuous training (MICT). The discordance between small single-center trials and large pragmatic trials led to downgrading for inconsistency and risk of bias, resulting in low certainty for peak VO_2 superiority in HFrEF.

For heart failure with preserved ejection fraction (HFpEF), randomized evidence consistently demonstrates no clinically meaningful superiority of HIIT over MICT. Adjusted analyses accounting for total energy expenditure further support equivalence between modalities. Although confidence intervals cross minimal clinically important difference thresholds, the direction of effect is consistent across datasets. Consequently, certainty was rated as moderate for no important difference.

Evidence regarding improvements in left ventricular ejection fraction (LVEF) in HFrEF populations was rated as very low certainty. While pooled meta-analyses report small mean improvements (~3%) favoring HIIT, these estimates are highly sensitive to inclusion of small trials with large effect sizes. The largest multicenter trial did not confirm superiority, and confidence intervals frequently cross the null. Downgrading occurred for inconsistency, imprecision, and suspected publication bias.

For six-minute walk distance (6MWT), pooled analyses suggest an improvement of approximately 28 meters favoring HIIT. Although statistically significant, the clinical magnitude

approaches but does not robustly exceed commonly cited minimal clinically important difference thresholds. Certainty was rated as low to moderate, primarily due to heterogeneity in baseline functional capacity and small-study influence.

Table 5. GRADE Evidence Profile

Outcome	Participants (Studies)	Effect (HIIT vs MICT)	Certainty of Evidence (GRADE)	Reasons for Downgrading
Peak VO_2 (HFrEF)	16 comparative studies (predominantly HFrEF; includes large multicenter RCT n=261)	Pooled small–moderate improvement (~+1.7 to +2.1 mL/kg/min) in meta-analyses; largest multicenter RCT shows no significant difference	●●○○ **Low**	Risk of bias (lack of blinding, protocol deviations), inconsistency (small trials vs large RCT discordance), suspected publication bias (small-study dominance)
Peak VO_2 (HFpEF)	2 HFpEF-focused RCT datasets (including OptimEx-Clin n=180)	No statistically significant superiority; adjusted difference ≈ 0 mL/kg/min	●●●○ **Moderate**(for no important difference)	Imprecision (wide CIs crossing MCID), variability in adherence and energy expenditure confounding
LVEF (HFrEF)	Subset of 16 studies reporting EF	Small pooled improvement (~3%) in some meta-analyses; largest RCT shows no significant difference	●○○○ **Very Low**	Risk of bias (small single-center trials), inconsistency (heterogeneous direction and magnitude), imprecision (CIs cross null in large trial), suspected publication bias
6-Minute Walk Distance	Subset of pooled RCTs	~+28 meters favoring HIIT (borderline clinical relevance)	●●○○ **Low–Moderate**	Inconsistency (heterogeneous functional status), risk of bias in smaller studies
Serious Adverse Events	All 16 comparative datasets	No consistent increase in adverse events with HIIT; event rates low	●●●○ **Moderate**	Imprecision (low event counts), safety underpowered for rare outcomes

Legend: ●●●● High certainty; ●●●○ Moderate certainty; ●●○○ Low certainty; ●○○○ Very low certainty

Safety outcomes were reassuring. Across all 16 comparative datasets, no consistent increase in serious adverse events was observed with HIIT when delivered under supervised conditions.

However, event counts were low and studies were not powered to detect rare harms. Certainty was therefore rated as moderate, downgraded for imprecision.

Overall, certainty of evidence ranges from low to moderate across outcomes. Confidence is highest for the conclusion that HIIT does not provide clinically meaningful superiority over MICT in HFpEF populations and does not increase short-term serious adverse events in stable patients. In contrast, certainty supporting superiority of HIIT in HFrEF remains limited due to heterogeneity between small and large trials and concerns regarding intervention fidelity.

Taken together, current evidence indicates that while HIIT may offer modest physiological advantages in selected HFrEF populations, these benefits are not robustly reproducible across large multicenter trials and do not extend to HFpEF. Both HIIT and MICT appear safe under structured supervision. Exercise prescription should therefore emphasize sustainability, adherence, phenotype-specific considerations, and clinical context rather than presumed intensity-driven superiority alone.

5. Discussion

This systematic review synthesizes comparative evidence from 16 randomized studies evaluating high-intensity interval training (HIIT) versus moderate-intensity continuous training (MICT) in older adults with heart failure. The principal finding is that although pooled meta-analyses suggest modest improvements in peak oxygen uptake (peak VO_2) and possibly left ventricular ejection fraction (LVEF) in heart failure with reduced ejection fraction (HFrEF) (Haykowsky et al., 2013; Xie et al., 2017; Yang et al., 2024; Fauzi, 2025), these advantages are not consistently confirmed in large, methodologically rigorous multicenter trials, most notably SMARTEX-HF (Ellingsen et al., 2017). In heart failure with preserved ejection fraction (HFpEF), contemporary randomized evidence, including OptimEx-Clin and its mechanistic subanalyses, does not demonstrate superiority of HIIT over MICT (Mueller et al., 2021; Mueller et al., 2022). Across included trials and pooled safety analyses, adverse event rates appear comparable when exercise is delivered under structured and supervised conditions (Ellingsen et al., 2017; Mueller et al., 2021; Fauzi, 2025; Gomes-Neto et al., 2018).

Taken together, the aggregate evidence suggests that exercise intensity alone is unlikely to be the dominant determinant of functional adaptation in older adults with heart failure. Rather, total training volume, adherence fidelity, supervision quality, and phenotype-specific pathophysiology appear to exert substantial influence on observed outcomes.

5.1 Principal Findings

In predominantly HFrEF populations, pooled meta-analyses report an approximate 1.7–2.1 mL/kg/min improvement in peak VO_2 with HIIT compared with MICT (Haykowsky et al., 2013; Xie et al., 2017; Yang et al., 2024; Fauzi, 2025). Early quantitative syntheses by Haykowsky et al. (2013) identified a statistically significant advantage favoring HIIT, and this direction of effect was reinforced in subsequent analyses by Xie et al. (2017), Yang et al. (2024), and Fauzi (2025). Across these pooled datasets, the signal consistently trends toward modest superiority of HIIT when smaller randomized trials are aggregated.

However, the clinical magnitude of this difference warrants careful interpretation. An improvement of 1–2 mL/kg/min in peak VO_2, although statistically significant, represents a modest increment in aerobic capacity. Peak VO_2 is a well-established prognostic marker in HFrEF, yet the absolute magnitude observed in pooled analyses is substantially smaller than the dramatic relative increases reported in early single-center mechanistic investigations, particularly the landmark trial by Wisløff et al. (2007), which also described reverse remodeling and marked improvements in LVEF. The contrast between these early findings and later pooled estimates

suggests that initial effect sizes may reflect optimized research environments, highly selected patient cohorts, and small-study variability rather than broadly generalizable physiological superiority.

The most influential counterweight to pooled superiority claims is the SMARTEX-HF trial (Ellingsen et al., 2017), the largest multicenter randomized comparison of HIIT and MICT in HFrEF (n = 261). In this rigorously conducted study, no statistically significant differences were observed between groups for peak VO_2 or LVEF at 12 weeks. Confidence intervals for between-group comparisons crossed the null, directly challenging the robustness of superiority conclusions derived from smaller trials. When large, pragmatic, methodologically robust multicenter evidence fails to replicate earlier findings, greater evidentiary weight must be accorded to the larger dataset. The discordance between small-trial superiority and large-trial neutrality represents a central interpretative inflection point in this evidence base.

In HFpEF populations, the findings are more internally consistent. The OptimEx-Clin trial (Mueller et al., 2021), the largest HFpEF-specific randomized comparison (n = 180), did not demonstrate superiority of HIIT over MICT for peak VO_2. Furthermore, a mechanistic subanalysis adjusting for total energy expenditure demonstrated near-complete attenuation of between-group differences (Mueller et al., 2022), suggesting that cumulative workload rather than interval intensity may drive physiological adaptation in this phenotype. Given the complex pathophysiology of HFpEF characterized by diastolic dysfunction, microvascular impairment, skeletal muscle abnormalities, and impaired peripheral oxygen extraction these findings are physiologically coherent. In such a multifactorial syndrome, modulation of intensity alone may be insufficient to overcome central and peripheral constraints.

With respect to LVEF in HFrEF populations, early small trials reported notable improvements with HIIT, including evidence of reverse remodeling (Wisløff et al., 2007; Angadi et al., 2015). However, pooled estimates across subsequent analyses demonstrate heterogeneity and imprecision (Haykowsky et al., 2013; Xie et al., 2017; Trejos-Montoya et al., 2019; Yang et al., 2024), and SMARTEX-HF did not confirm significant between-group differences (Ellingsen et al., 2017). The variability in direction and magnitude of LVEF effects ranging from pronounced improvement in small mechanistic studies to neutral findings in large pragmatic trials results in low to very low certainty regarding structural superiority of HIIT.

Safety outcomes across the 16 included comparative studies were generally reassuring. Neither SMARTEX-HF (Ellingsen et al., 2017) nor OptimEx-Clin (Mueller et al., 2021) reported statistically significant excess serious adverse cardiovascular events in HIIT arms. Meta-analytic syntheses similarly found no increased risk associated with HIIT relative to MICT (Fauzi, 2025; Gomes-Neto et al., 2018). Nonetheless, most trials were not powered primarily for safety endpoints, and absolute event counts were low. Therefore, although no harm signal is evident under supervised conditions, definitive equivalence for rare adverse outcomes cannot be conclusively established.

5.2 Reconciling Small-Trial Superiority With Large-Trial Neutrality

A defining feature of this evidence base is the divergence between early small single-center trials reporting marked superiority of HIIT and later multicenter trials demonstrating neutral findings. Several complementary explanations likely account for this discrepancy.

First, small-study effects and potential publication bias may have inflated early estimates of benefit. The seminal trial by Wisløff et al. (2007) was conducted in a highly specialized academic environment with intensive supervision, strict intensity monitoring, and exceptional protocol fidelity. Under such optimized conditions, adherence to prescribed high-intensity targets is

more readily achieved, thereby maximizing physiological contrast between HIIT and MICT. Smaller samples are also inherently more susceptible to random variability and overestimation of effect sizes. Subsequent meta-analyses (Haykowsky et al., 2013; Xie et al., 2017; Yang et al., 2024; Fauzi, 2025), which incorporate multiple small single-center trials, may therefore amplify early pronounced findings when pooled, particularly if larger neutral trials are fewer in number.

Second, intensity contamination represents a major methodological challenge in pragmatic exercise trials. In SMARTEX-HF (Ellingsen et al., 2017), a substantial proportion of participants randomized to HIIT did not consistently reach prescribed high-intensity thresholds, whereas many participants assigned to MICT exceeded their moderate-intensity targets. This bidirectional contamination narrowed the effective exposure differential between groups. When the contrast in physiological stimulus is attenuated, the capacity to detect meaningful between-group differences logically diminishes. In real-world cardiac rehabilitation settings, where supervision intensity and patient monitoring vary, maintaining strict separation between prescribed intensity domains is inherently difficult. The SMARTEX-HF experience suggests that superiority claims may depend on the feasibility of achieving sustained intensity separation in practice.

Third, selection bias and regression to the mean may partially explain early favorable findings. Small single-center trials often recruit highly motivated patients with fewer comorbidities, higher baseline physiological reserve, and strong adherence capacity. Such populations may exhibit greater responsiveness to high-intensity stimuli. In contrast, large multicenter trials recruit broader, more heterogeneous cohorts reflective of routine clinical practice, including older adults with multimorbidity, frailty, and variable training compliance. Differences in baseline risk, functional limitation, and adaptive reserve may materially influence observed responses to exercise intensity.

Fourth, intervention duration may constrain the detection of structural remodeling. Most comparative trials were limited to approximately 12 weeks of intervention. While improvements in peak VO_2 can occur within this timeframe, sustained structural cardiac remodeling particularly changes in left ventricular geometry and function may require longer exposure. Although long-term data from the Generation 100 study (Stensvold et al., 2017) provide insight into sustained high-intensity training in older adults, these data do not directly evaluate comparative HIIT versus MICT effects specifically in heart failure populations. Thus, the possibility that longer-duration, tightly monitored HIIT might yield incremental structural benefits cannot be definitively excluded.

Finally, phenotype heterogeneity likely contributes to conflicting results. Aggregating heart failure with reduced ejection fraction (HFrEF) and heart failure with preserved ejection fraction (HFpEF) without careful stratification risks obscuring genuine phenotype-specific differences. Contemporary evidence from OptimEx-Clin and its mechanistic analyses (Mueller et al., 2021; Mueller et al., 2022) demonstrates that in HFpEF, superiority of HIIT is not observed and that total energy expenditure may account for most of the observed adaptation. Early pooled analyses that did not sufficiently stratify by phenotype may therefore have overgeneralized conclusions derived primarily from HFrEF cohorts.

Collectively, these considerations suggest that the apparent superiority of HIIT is context-dependent rather than universal. In tightly controlled research environments among selected HFrEF patients with high adherence and clear intensity separation, HIIT may confer modest incremental benefit. However, in pragmatic multicenter settings and in HFpEF populations, this advantage appears attenuated or absent. The totality of evidence therefore supports a nuanced interpretation in which delivery context, adherence fidelity, phenotype, and cumulative

workload materially shape outcomes, rather than intensity alone determining clinical superiority.

5.3 Clinical and Translational Implications

The clinical implication of this synthesis is not that high-intensity interval training (HIIT) lacks efficacy, but rather that its superiority over moderate-intensity continuous training (MICT) is not consistently established across heart failure phenotypes or delivery contexts. Both modalities reliably improve exercise capacity relative to baseline in patients with heart failure (Wisløff et al., 2007; Ellingsen et al., 2017; Haykowsky et al., 2013; Xie et al., 2017; Angadi et al., 2015). The consistent within-group improvements observed across trials suggest that structured aerobic exercise regardless of intensity format remains a cornerstone of non-pharmacological management.

Accordingly, the more clinically relevant question may not be which modality is superior in absolute physiological terms, but which approach optimizes adherence, sustainability, and safety for a given patient. In real-world cardiac rehabilitation settings, long-term engagement and behavioral consistency likely exert greater influence on outcomes than marginal differences in prescribed intensity. The neutral findings of large pragmatic trials such as SMARTEX-HF (Ellingsen et al., 2017) and OptimEx-Clin (Mueller et al., 2021) reinforce the importance of contextualizing intensity within implementation realities.

Cardiac rehabilitation programs should therefore prioritize feasibility, patient preference, symptom tolerance, and long-term adherence. In motivated, clinically stable HFrEF patients under structured supervision and careful monitoring, HIIT may be an appropriate and efficient strategy to achieve improvements in peak VO_2. However, even in HFrEF populations, superiority is modest and sensitive to adherence fidelity. In HFpEF populations where randomized evidence has not demonstrated meaningful superiority of HIIT (Mueller et al., 2021; Mueller et al., 2022) exercise prescription should emphasize tolerability, sustainability, and cumulative workload rather than presumed intensity-driven advantage.

The modest magnitude of pooled differences further underscores that factors beyond intensity may be equally or more influential determinants of adaptation. Total training volume, progressive overload, integration of resistance training, and incorporation of behavioral support strategies likely contribute substantially to functional improvement. Mechanistic data indicating attenuation of between-group differences after adjustment for total energy expenditure in HFpEF (Mueller et al., 2022) highlight the importance of cumulative workload rather than interval intensity alone.

From a translational perspective, these findings suggest that individualized exercise prescription guided by phenotype, comorbidity burden, patient goals, and capacity for sustained adherence may yield greater clinical value than uniform adoption of high-intensity protocols. In contemporary heart failure care, the optimization of long-term participation in structured exercise may be more consequential than marginal gains in short-term peak VO_2 attributable to interval intensity.

5.4 Safety Considerations and Practical Implementation

Although the safety profile of high-intensity interval training (HIIT) appears reassuring across randomized trials, interpretation must remain contextualized. In the included studies, HIIT was delivered under controlled, supervised conditions with structured warm-up phases, individualized heart rate monitoring, and clearly defined progression protocols (Ellingsen et al., 2017; Mueller et al., 2021). These safeguards likely contributed to the low incidence of serious adverse cardiovascular events reported in both HIIT and moderate-intensity continuous training (MICT) arms (Fauzi, 2025; Gomes-Neto et al., 2018). However, the degree to which these findings generalize to unsupervised, home-based, or community settings remains uncertain. The

fidelity of intensity prescription and real-time monitoring in research environments may not be reproducible in routine practice.

Older adults with heart failure frequently present with multimorbidity, frailty, sarcopenia, renal dysfunction, arrhythmias, and polypharmacy. Such factors may modify physiological response to high-intensity exercise and potentially increase vulnerability to hemodynamic instability or musculoskeletal injury. Although no consistent excess of acute adverse events was observed in major trials such as SMARTEX-HF (Ellingsen et al., 2017) or OptimEx-Clin (Mueller et al., 2021), these studies were not powered to detect rare catastrophic outcomes such as malignant arrhythmias or sudden cardiac events. Consequently, absence of evidence of harm should not be interpreted as definitive proof of equivalence for rare but clinically serious events.

Clinical implementation of HIIT should therefore incorporate comprehensive pre-participation assessment, including functional evaluation, review of comorbidities, medication reconciliation, and risk stratification. Gradual progression of interval intensity, particularly during the initial weeks of training, may mitigate hemodynamic stress and enhance tolerability. Structured supervision at least during early phases appears prudent, especially for patients with advanced age, significant left ventricular dysfunction, or arrhythmic risk.

Adherence fidelity remains central to both efficacy and safety. As demonstrated in SMARTEX-HF (Ellingsen et al., 2017), deviation from prescribed intensity targets can attenuate between-group contrast and potentially obscure true physiological effects. Programs integrating behavioral counseling, motivational reinforcement, wearable heart rate monitoring, and remote feedback systems may enhance compliance with prescribed intensity zones while simultaneously improving safety oversight. Remote monitoring technologies, when combined with structured coaching, may help maintain protocol fidelity in pragmatic settings while reducing contamination between intensity domains.

Ultimately, while supervised HIIT appears safe in stable, appropriately screened heart failure patients, careful patient selection, individualized progression, and structured monitoring are essential to translate trial findings into routine clinical practice.

5.5 Strengths and Limitations

This review possesses several methodological strengths. First, it incorporates phenotype-specific stratification, explicitly distinguishing between heart failure with reduced ejection fraction (HFrEF) and heart failure with preserved ejection fraction (HFpEF), thereby avoiding inappropriate aggregation of biologically distinct syndromes. Second, it applies a hierarchical interpretative framework that contrasts early small single-center trials with larger, pragmatic multicenter randomized controlled trials, assigning appropriate evidentiary weight to study size and methodological rigor. Third, the use of a structured GRADE assessment provides transparent evaluation of certainty across outcomes, explicitly accounting for risk of bias, inconsistency, imprecision, and suspected publication bias. By integrating pooled quantitative findings with trial-level context and methodological appraisal, this synthesis aims to provide a balanced and clinically meaningful interpretation aligned with contemporary standards of evidence evaluation.

Nonetheless, important limitations must be acknowledged. Intervention protocols varied substantially across trials, including differences in interval duration, intensity targets, total session time, weekly frequency, and supervision models. Such heterogeneity may contribute to between-study variability and complicate interpretation of pooled effects. Adherence rates and achieved intensity were inconsistently reported, and objective verification of compliance was not uniformly performed. As demonstrated in larger

trials, intensity contamination can materially influence observed treatment effects.

Most included studies were short-term, typically spanning 7–12 weeks. While this duration is sufficient to detect improvements in peak oxygen uptake, it may be insufficient to evaluate sustained structural remodeling or long-term clinical impact. Furthermore, the majority of trials relied on surrogate physiological endpoints principally peak VO_2 and left ventricular ejection fraction (LVEF) rather than hard clinical outcomes such as hospitalization, cardiovascular events, or mortality. Although peak VO_2 is strongly prognostic in heart failure, improvements in this parameter do not necessarily translate directly into reductions in clinical events.

Safety endpoints were generally secondary outcomes, and event rates were low. Consequently, trials were underpowered to detect rare but serious adverse events. Finally, potential small-study effects and publication bias cannot be excluded, particularly given the pattern of larger effect sizes observed in smaller single-center trials compared with larger pragmatic studies.

Taken together, while this review provides a comprehensive and methodologically rigorous synthesis, interpretation should remain cognizant of heterogeneity in intervention delivery, limited long-term outcome data, and reliance on surrogate endpoints.

5.6 Future Research Directions

Future investigations should move beyond small, single-center mechanistic trials and prioritize adequately powered, multicenter randomized designs capable of detecting modest but clinically meaningful differences between high-intensity interval training (HIIT) and moderate-intensity continuous training (MICT). Preservation of true intervention contrast will be critical. This requires strict intensity verification using objective heart rate monitoring, cardiopulmonary exercise testing–derived thresholds, or wearable-based workload tracking to minimize contamination between groups, as observed in prior pragmatic trials. Without reliable separation of achieved intensity, conclusions regarding superiority or equivalence remain inherently uncertain.

Standardized reporting of adherence achieved workload, and cumulative training volume should become mandatory in future trials. Reporting should extend beyond prescribed intensity to include actual time spent within target heart rate zones, session completion rates, and progression metrics. Transparent documentation of these parameters would facilitate more accurate interpretation of dose–response relationships and reduce heterogeneity across pooled analyses.

Longer intervention durations and extended follow-up are also necessary. Most existing trials have been limited to approximately 12 weeks, a timeframe sufficient to detect changes in peak oxygen uptake but potentially inadequate to evaluate sustained ventricular remodeling or long-term clinical trajectories. Extended follow-up would clarify whether short-term physiological gains persist, plateau, or translate into reductions in heart failure–related hospitalization and mortality.

Critically, future trials should incorporate hard clinical endpoints. While peak VO_2 and left ventricular ejection fraction are prognostically relevant, modest differences in these surrogate measures do not automatically confer meaningful reductions in adverse clinical events. Trials powered for hospitalization, cardiovascular events, and mortality would determine whether intensity-driven physiological differences yield tangible prognostic benefit.

Phenotype-specific stratification must also be embedded prospectively into trial design. Given the differential pathophysiology of HFrEF and HFpEF, separate powered analyses are required rather than post hoc subgroup interpretation. Integration of mechanistic biomarkers including measures of endothelial function, skeletal muscle oxidative capacity, inflammatory markers, and ventricular remodeling indices may further clarify

why certain phenotypes respond differently to intensity modulation.

Finally, pragmatic integration of wearable technologies and remote monitoring platforms may improve adherence tracking and enhance protocol fidelity. Continuous heart rate telemetry, workload quantification, and automated feedback systems could mitigate intensity drift and contamination while simultaneously enhancing patient engagement. Such approaches may bridge the gap between controlled research environments and real-world cardiac rehabilitation settings.

Conclusions

This systematic review provides a comprehensive comparative evaluation of high-intensity interval training (HIIT) and moderate-intensity continuous training (MICT) in older adults with heart failure. The totality of evidence indicates that HIIT may confer modest improvements in peak oxygen uptake (peak VO_2) in patients with heart failure with reduced ejection fraction (HFrEF), with pooled estimates suggesting an approximate 1.7–2.1 mL/kg/min advantage. However, this apparent superiority is not consistently reproduced in large, methodologically rigorous multicenter trials, which demonstrate neutral between-group differences. In heart failure with preserved ejection fraction (HFpEF), contemporary randomized evidence does not support meaningful superiority of HIIT over MICT, particularly when total energy expenditure is accounted for.

Improvements in left ventricular ejection fraction (LVEF) are heterogeneous and primarily driven by small single-center studies; larger pragmatic trials do not confirm consistent structural advantage. Functional capacity gains, including six-minute walk distance, are modest and approach but do not robustly exceed commonly cited thresholds for clinical relevance. Safety profiles appear comparable between HIIT and MICT when exercise is delivered under supervised and structured conditions, although available trials are underpowered to detect rare adverse events.

Collectively, these findings suggest that exercise intensity alone is unlikely to be the principal determinant of functional adaptation in older adults with heart failure. Rather, cumulative training volume, adherence fidelity, supervision quality, patient phenotype, and long-term sustainability appear to exert equal or greater influence on clinical outcomes. While HIIT represents a safe and potentially time-efficient strategy for selected, clinically stable HFrEF patients under appropriate supervision, its universal superiority over MICT is not established.

Future research should focus on adequately powered multicenter trials incorporating strict intensity verification, longer follow-up, phenotype-specific stratification, and hard clinical endpoints to determine whether modest physiological differences translate into meaningful reductions in hospitalization and mortality. Until such evidence emerges, exercise prescription in heart failure should prioritize individualized, sustainable, and context-appropriate rehabilitation strategies rather than presumed intensity-driven superiority.

Funding

The author/s have not received any funding for this study

Conflicts of Interest

The authors declare no potential conflicts of interest related to the research, writing, or

publication of this work.

Declaration of generative AI and AI-assisted technologies

The authors used generative AI tools solely to improve language clarity and correct grammatical issues during manuscript preparation. All content was carefully reviewed and revised by the authors, who take full responsibility for the accuracy,

integrity, and originality of the final published work.

Reference

Angadi, S., Mookadam, F., Lee, C.-D., Tucker, W. J., Haykowsky, M., & Gaesser, G. A. (2015). High-intensity interval training vs. moderate-intensity continuous exercise training in heart failure with preserved ejection fraction: A pilot study. Journal of Applied Physiology. https://doi.org/10.1152/japplphysiol.00518.2014

Badenhop, D., Long, M., Laurent, C., & Keylock, K. T. (2020). High-intensity interval versus moderate-intensity continuous training in cardiac rehabilitation. Journal of Clinical Exercise Physiology. https://doi.org/10.31189/2165-6193-9.1.10

Beale, L., McIntosh, R., Raju, P., et al. (2013). A comparison of high intensity interval training with circuit training in a short-term cardiac rehabilitation programme for patients with chronic heart failure. https://doi.org/10.4172/2329-9096.1000151

Benda, N., Seeger, J., Stevens, G., et al. (2015). Effects of high-intensity interval training versus continuous training on physical fitness, cardiovascular function and quality of life in heart failure patients. PLOS ONE. https://doi.org/10.1371/journal.pone.0141256

Donelli da Silveira, A., Beust de Lima, J., da Silva Piardi, D., et al. (2020). High-intensity interval training is effective and superior to moderate continuous training in patients with heart failure with preserved ejection fraction: A randomized clinical trial. European Journal of Preventive Cardiology. https://doi.org/10.1177/2047487319901206

Ellingsen, Ø., Halle, M., Conraads, V., et al. (2017). High-intensity interval training in patients with heart failure with reduced ejection fraction. Circulation. https://doi.org/10.1161/CIRCULATIONAHA.116.022924

Fauzi, F. (2025). Elevating cardiac rehabilitation: A literature review of high-intensity interval training's superior efficacy in patients with heart failure. International Journal of Medical Science and Health Research. https://doi.org/10.70070/krhcdt47

Gasser, B., Boesing, M., Schoch, R., et al. (2021). High-intensity interval training for heart failure patients with preserved ejection fraction (HIT-HF): Rationale and design of a prospective, randomized, controlled trial. Frontiers in Physiology. https://doi.org/10.3389/fphys.2021.734111

Hannan, A. L., Hing, W., Simas, V., et al. (2018). High-intensity interval training versus moderate-intensity continuous training within cardiac rehabilitation: A systematic review and meta-analysis. Open Access Journal of Sports Medicine. https://doi.org/10.2147/OAJSM.S150596

Haykowsky, M., Timmons, M. P., Kruger, C., et al. (2013). Meta-analysis of aerobic interval training on exercise capacity and systolic function in patients with heart failure and reduced ejection fractions. American Journal of Cardiology. https://doi.org/10.1016/j.amjcard.2013.01.303

Koufaki, P., Mercer, T., George, K., & Nolan, J. (2014). Low-volume high-intensity interval training vs continuous aerobic cycling in patients with chronic heart failure: A pragmatic randomised clinical trial of feasibility and effectiveness. Journal of Rehabilitation Medicine. https://doi.org/10.2340/16501977-1278

Maiorana, A. (2012). Interval training confers greater gains than continuous training in people with heart failure. Journal of Physiotherapy. https://doi.org/10.1016/S1836-9553(12)70113-3

Mueller, S., Winzer, E., Duvinage, A., et al. (2021). Effect of high-intensity interval training, moderate continuous training, or guideline-based physical activity advice on peak oxygen consumption in patients with heart failure with preserved ejection fraction: A randomized clinical trial. JAMA. https://doi.org/10.1001/jama.2020.26812

Mueller, S., Cervenka, M., Winzer, E., et al. (2022). Associations between training characteristics and change in peak oxygen consumption following exercise training in patients with heart failure with preserved ejection fraction. European Heart Journal. https://doi.org/10.1093/eurheartj/ehac544.2476

Nasser, R., Verelst, F. R., Mueller, S., et al. (2025). Effect of structured exercise training on chronotropic incompetence in patients with heart failure with preserved ejection fraction. European Journal of Preventive Cardiology. https://doi.org/10.1093/eurjpc/zwaf236.365

Papathanasiou, J. (2020). Are the group-based interventions improving the functional exercise capacity and quality of life of frail subjects with chronic heart failure? Journal of Frailty, Sarcopenia and Falls. https://doi.org/10.22540/JFSF-05-102

Papathanasiou, J., Petrov, I., Tokmakova, M., et al. (2020). Group-based cardiac rehabilitation interventions: A challenge for physical and rehabilitation medicine physicians: A randomized controlled trial. European Journal of Physical and Rehabilitation Medicine. https://doi.org/10.23736/S1973-9087.20.06013-X

Papathanasiou, J., Petrov, I., Tsekoura, D., et al. (2021). Does group-based high-intensity aerobic interval training improve the inflammatory status in patients with chronic heart failure? A randomized controlled trial. European Journal of Physical and Rehabilitation Medicine. https://doi.org/10.23736/S1973-9087.21.06894-5

Stensvold, D., Dalen, H. D., Vesterbekkmo, E. K. V., et al. (2024). Is high intensity interval training superior to moderate intensity continuous training in older adults with cardiovascular disease? The Generation 100 study. European Journal of Preventive Cardiology. https://doi.org/10.1093/eurjpc/zwae175.032

Støylen, A., Conraads, V., Halle, M., et al. (2012). Controlled study of myocardial recovery after interval training in heart failure: SMARTEX-HF Rationale and design. European Journal of Preventive Cardiology. https://doi.org/10.1177/1741826711403252

Trejos-Montoya, J., Araya-Ramírez, F., Sánchez-Ureña, B., & Kelley, G. (2019). High intensity interval versus moderate intensity training in heart failure patients: Systematic review and meta-analysis. Medicine & Science in Sports & Exercise. https://doi.org/10.1249/01.mss.0000560698.73462.7e

Wang, C., Xing, J., Zhao, B., et al. (2022). The effects of high-intensity interval training on exercise capacity and prognosis in heart failure and coronary artery disease: A systematic review and meta-analysis. Cardiovascular Therapeutics. https://doi.org/10.1155/2022/4273809

Wisløff, U., Støylen, A., Loennechen, J. P., et al. (2007). Superior cardiovascular effect of aerobic interval training versus moderate continuous training in heart failure patients: A randomized study. Circulation. https://doi.org/10.1161/CIRCULATIONAHA.106.675041

Xie, B., Yan, X., Cai, X., & Li, J. (2017). Effects of high-intensity interval training on aerobic capacity in cardiac patients: A systematic review with meta-analysis. BioMed Research International. https://doi.org/10.1155/2017/5420840

Yang, C., Zhang, L., Cheng, Y., et al. (2024). High intensity interval training vs. moderate intensity continuous training on aerobic capacity and functional capacity in patients with heart failure: A systematic review and meta-analysis. Frontiers in Cardiovascular Medicine. https://doi.org/10.3389/fcvm.2024.1302109