INTRODUCTION

Colorectal cancer (CRC) is the third most diagnosed cancer globally [1]. Despite improvements in surgical techniques and adjunctive therapies, as well as the implementation of enhanced recovery pathways, postoperative morbidity remains a concern [2], particularly for vulnerable groups of patients including frail patients with complex comorbidities [3–5]. While patient age is common-
ly considered when assessing surgical fitness, there has been an increasing shift in attention towards frailty as a more accurate reflection of a patient's physiological and functional tolerance for surgery, particularly as frailty is potentially modifiable [5].

Prehabilitation (PH) refers to the preoperative process of holistic care that seeks to improve one's physical and psychological health to minimize risks of postoperative morbidity and accelerate recovery [6]. The 3 domains of PH are enhancing physical activity, improving nutrition, and conducting psychological interventions. PH has been purported to improve preoperative baseline functional capacity and physiological reserves, resulting in better postoperative outcomes [2, 7–9].

Hitherto, while PH has been regarded as safe and feasible, and it has been widely adopted for CRC patients planned for surgery [10, 11], the benefit of PH on postoperative clinical outcomes remains unclear. Although some authors have demonstrated reduced postoperative morbidity and shortened hospital stays with PH [12, 13], others have shown no improvement [14, 15]. The lack of strong evidence is highlighted in the 2018 Enhanced Recovery After Surgery (ERAS) guidelines, which recommend further research before establishing PH as a compulsory item in an ERAS protocol [16]. Of the components of PH, only a preoperative nutritional assessment is strongly recommended in the existing ERAS guidelines.

The implementation of a routine trimodal PH program for CRC patients can pose a substantial strain on healthcare resources, particularly the need for additional healthcare professionals such as physiotherapists, psychologists, and dietitians [17]. Compliance with complex PH regimes may pose a challenge, even among younger CRC patients. Few studies to date have evaluated patient and caregiver satisfaction, as well as stress levels in response to PH programs [18].

To date, there has been no systematic review and meta-analysis specifically demonstrating the clinical value of preoperative PH programs. It is hence timely to synthesize the available evidence comparing postoperative outcomes between PH and standard of care among CRC patients awaiting surgery.

**METHODS**

**Search process**
This study strictly adhered to the *Cochrane Handbook for Systematic Reviews of Interventions* [19], and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement guidelines [20]. An electronic search was comprehensively performed on August 23, 2021 and subsequently repeated on August 1, 2022 in the following major databases: MEDLINE (via PubMed), the Cochrane databases including CENTRAL (Cochrane Central Register of Controlled Trials), Embase, Scopus, CINAHL, and ClinicalTrials.gov to identify all published studies and abstracts comparing PH versus standard preoperative care for CRC patients awaiting surgery. In consult with a research librarian, a repetitive and exhaustive combination of the following Medical Subject Headings (MeSH) search terms were used: “prehabilitation,” “colorectal cancer,” “colon cancer,” and “rectal cancer.” A manual search of the reference lists of relevant studies was also done to identify additional studies. An updated search was performed between September 2022 and January 2023, which did not yield additional articles for inclusion.

**Inclusion and exclusion criteria**
Both randomized controlled trials (RCTs) and non-RCTs were included if comparative outcomes were reported for CRC patients receiving either PH or standard preoperative care before surgery. Studies that included noncolorectal surgery were excluded. Following the formal definition of PH, only programs that incorporated all 3 domains of exercise, nutrition, and anxiety management were included. This also reduced clinical and, consequently, statistical heterogeneity. There were no restrictions on the types of exercise programs (home-based vs. supervised, moderate vs. high intensity). Given the lack of an available translator, non-English studies were excluded. Other studies of the following designs were also excluded: surveys, trial protocols, review articles, and opinion pieces. Abstracts with no extractable data were also excluded.

**Selection of studies and data extraction**
As shown in Fig. 1, study selection was performed in 2 stages by 2 reviewers in an independent fashion. Studies were first screened for inclusion by their titles and abstracts, and the full texts of these preliminarily included studies were then reviewed in their entirety to confirm their inclusion in the final analysis. The senior author served as the arbiter to resolve differences of opinion regarding the studies’ eligibility.

In line with the aim of this study, the primary outcomes of interest included the following postoperative outcomes: complications stratified based on the Clavien-Dindo (CD) classification, length of hospital stay, 30-day readmission, and length of operative time. Secondary outcomes of interest included functional outcomes: the 6-minute walk test (6MWT) and Hospital Anxiety and Depression Scale (HADS) scores. Importantly, these were the commonly reported outcomes amenable to formal meta-analysis. In addition, the following data were abstracted from each study: first author, year, study design, duration of PH, ERAS protocols, age, sex, American Society of Anesthesiologists (ASA) physical
status (PS) classification, the proportion of patients who received neoadjuvant therapy, the site of cancer, and the proportion of patients who underwent laparoscopic surgery.

**Statistical analysis**

The RevMan ver. 5.4 (The Nordic Cochrane Centre) was used to perform all statistical analyses. Pooled weighted mean differences or standardized mean differences were used as the summary statistics for continuous variables, while the risk ratio (RR) was employed for dichotomous variables. The $I^2$ value was computed to estimate statistical heterogeneity, and a random-effects model was chosen when the value was greater than 50%. The results were reported with 95% confidence intervals (CIs), and a $P$-value of less than 0.05 was regarded as statistically significant. Variables reported as median (range) were converted to the respective mean and standard deviation using the methods described by Hozo et al. [21]. Sensitivity analyses were conducted when deemed appropriate.

**Assessment of bias**

The Cochrane Risk of Bias tool was utilized to evaluate the risk of bias for RCTs based on domains of selection, performance, detection, attrition, reporting, and other bias. For non-RCTs, the Newcastle-Ottawa scale (NOS) [22] was employed to evaluate aspects of patient selection, the comparability of study groups, and outcomes assessment. Since there were fewer than 10 studies, it was statistically not feasible to assess publication bias using either funnel plots or the Egger regression test [19].

**RESULTS**

**Systematic search**

Initially, 515 citations were retrieved from a systematic search across different databases, and 13 remained after removing duplicates, as well as irrelevant articles identified after title and abstract review. Common reasons for exclusion are presented in Fig. 1. The full texts of these 13 articles were reviewed in their entirety, and 6 were excluded based on the reasons stated in Fig. 1. The remaining 7 articles were 4 RCTs [12, 14, 15, 23], 1 prospective study [24], and 2 retrospective studies [13, 25].

**Study characteristics**

In total, 1,042 CRC patients awaiting surgery were included, of whom 382 underwent trimodal PH and 660 received standard perioperative care under ERAS pathways. There were no significant variations in PH programs among studies, as most employed a mix of moderate aerobic with resistance exercise 3 to 4 times a week. Only 1 study employed high-intensity training [25]. The noncompliance rates ranged from 2% to 55%. The median duration of PH ranged from 20 to 40 days. The mean age of patients ranged from 65.7 to 80.0 and 66.0 to 80.8 years in the PH and control arms, respectively, considering 3 studies [13, 15, 25] that specifically included high-risk or frail patients. Other pertinent parameters, including the ASA PS classification, the proportion of patients who received neoadjuvant therapy, and the proportion of patients who underwent laparoscopic surgery, were comparable between both arms. The detailed baseline characteristics can be found in Table 1 [12–15, 23–25].

**Study quality**

All 4 RCTs had high risk of performance and detection bias, since the participants, personnel, and outcome assessors were not blinded, understandably so given the nature of the interventions. Selection bias was low in all studies given adequate concealment and randomization measures (Fig. 2) [12–15, 23–25]. The 3 other non-RCTs scored 6 and above out of 9, based on the NOS tool, and were deemed to be robust methodologically.

**Postoperative outcomes**

No significant difference was found in the length of operative time (2 studies, n = 197) between both arms (mean difference [MD], −10.54 minutes; 95% CI, −26.23 to 5.15 minutes; $P = 0.19$)
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Comparison</th>
<th>Exercise program</th>
<th>ERAS program</th>
<th>Noncompliance (%)</th>
<th>Duration of PH (day)</th>
<th>No. of patients</th>
<th>Age (yr)</th>
<th>Male sex</th>
<th>ASA PS classification</th>
<th>Neoadjuvant therapy</th>
<th>Site of cancer</th>
<th>Laparoscopic approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li et al. [14] (2013)</td>
<td>PT</td>
<td>Multimodal PH vs. ERAS alone</td>
<td>Supervised aerobic and muscular resistance exercises for 30 min 3 times/wk</td>
<td>Preoperative education for premedication and exercise, with early mobilization</td>
<td>53.0</td>
<td>33 (21–46)</td>
<td>87</td>
<td>42</td>
<td>45</td>
<td>67.4 ± 11.0</td>
<td>66.4 ± 12.0</td>
<td>PH</td>
<td>Control</td>
</tr>
<tr>
<td>Bousquet-Dion et al. [21] (2018)</td>
<td>RCT</td>
<td>Multimodal PH vs. ERAS alone</td>
<td>Supervised moderate exercise 1 time/wk and home-based aerobic and resistance training 3 times/wk</td>
<td>Preoperative education for premedication and exercise, with early mobilization</td>
<td>2.0</td>
<td>32 (25–48)</td>
<td>63</td>
<td>37</td>
<td>26</td>
<td>73.4 ± 3.0</td>
<td>67.8 ± 5.7</td>
<td>PH</td>
<td>Control</td>
</tr>
<tr>
<td>Carli et al. [15] (2020)</td>
<td>RCT</td>
<td>Multimodal PH vs. ERAS alone</td>
<td>Supervised moderate exercise 1 time/wk and home-based aerobic and resistance training 3 times/wk</td>
<td>Preoperative education for premedication and exercise, with early mobilization</td>
<td>20.0</td>
<td>40 (28–51)</td>
<td>110</td>
<td>55</td>
<td>35</td>
<td>77.5 ± 2.9</td>
<td>80.8 ± 2.6</td>
<td>PH</td>
<td>Control</td>
</tr>
<tr>
<td>de Kleer et al. [25] (2021)</td>
<td>RCT</td>
<td>Multimodal PH vs. ERAS alone</td>
<td>Supervised high intensity training 3 times/wk and aerobic training 4 times/wk</td>
<td>Preoperative education for premedication and exercise, with early mobilization</td>
<td>10.0</td>
<td>35 (18–54)</td>
<td>351</td>
<td>76</td>
<td>275</td>
<td>74.0 ± 6.9</td>
<td>685 ± 15.0</td>
<td>PH</td>
<td>Control</td>
</tr>
<tr>
<td>Lopez-Rodriguez et al. [12] (2021)</td>
<td>RCT</td>
<td>Multimodal PH vs. ERAS alone</td>
<td>Home-based aerobic and muscular resistance training 30–45 min for 30-day before surgery</td>
<td>Preoperative education for premedication and exercise, with early mobilization</td>
<td>NA</td>
<td>29 ± 2.8</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>65.5 ± 0.12</td>
<td>66.0 ± 0.8</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>van der Helst et al. [13] (2021)</td>
<td>RCT</td>
<td>Multimodal PH vs. ERAS alone</td>
<td>Supervised exercise program 30–45 min with local personal trainer 2 times/wk for at &gt; 4 wk and at-home workouts</td>
<td>Preoperative education for premedication and exercise, with early mobilization</td>
<td>30.6</td>
<td>NA</td>
<td>334</td>
<td>124</td>
<td>210</td>
<td>80.0 ± 2.3</td>
<td>74.5 ± 1.7</td>
<td>66</td>
<td>123</td>
</tr>
</tbody>
</table>

Values are presented as number only, median (range), or mean ± standard deviation.
ERAS, Enhanced Recovery After Surgery; PH, prehabilitation; ASA, American Society Anesthesiologists; PS, physical status; PT, prospective study; NA, not applicable; RCT, randomized controlled trial.
Fig. 2. Risk of bias. (A) Overall assessment. (B) Summary of the risk of bias of the randomized controlled trials. Green circles indicate low risk of bias, empty fields indicate unclear risk of bias, and red circles indicate high risk of bias.

(Fig. 3B) [15, 24]. Postoperatively, the risk of CD grades I and II complications (7 studies, n = 1,042) was comparable between both groups (RR, 0.82; 95% CI, 0.62 to 1.07; P = 0.15) (Fig. 4A) [12–15, 23–25]. A similar finding was noted in the risk of severe (≥ CD grades III) complications (7 studies, n = 1,042; RR, 1.02; 95% CI, 0.72 to 1.44; P = 0.92) (Fig. 4B) [12–15, 23–25].

No statistically significant differences were found in the length of hospital stay (6 studies, n = 708; MD, –0.48 days; 95% CI, –1.21 to 0.26 days; P = 0.21) (Fig. 3A) [12, 14, 15, 23–25] between both groups, or in the risk of 30-day readmission (4 studies, n = 337; RR, 0.74; 95% CI, 0.18 to 3.07; P = 0.68) (Fig. 3C) [14, 15, 23, 24].

The findings above were consistent after sensitivity analyses, in which non-RCTs and studies that specifically included frail patients were excluded from the analysis. In addition, a sensitivity analysis of studies that included only frail or high-risk patients showed corroborative findings. Another sensitivity analysis was performed that excluded the study by Lopez-Rodriguez-Arias et al. [12], which had a small sample size of 20, and the results remained consistent.

Secondary outcomes
A meta-analysis of functional outcomes was not conducted given heterogeneity in reporting. Li et al. [24] demonstrated better 6MWT scores in the PH group at both 4 weeks (407 ± 111 m vs. 356 ± 71 m; P = 0.01) and 8 weeks postoperatively (459 ± 101 m vs. 375 ± 58 m; P < 0.01). In addition, patients in the PH group remained above baseline values in walking capacity at 8 weeks, while the control group went below the preoperative level (P < 0.01). These findings corroborated those of Gillis et al. [14], who showed that patients in the PH arm were above baseline at 8 weeks as compared to the control group (+23.4 ± 8.2 m vs. –21.8 ± 80.7 m; P = 0.01). In contrast, Bousquet-Dion et al. [23] showed no difference in walking capacity between both groups. This was recapitulated by Carli et al. [15], who demonstrated no statistically significant difference in the proportion of patients who returned to their preoperative 6MWT scores (P = 0.26).

In terms of anxiety scores, which were reported by 2 studies (n = 187) [14, 15], there were no statistically significant differences in both HADS Anxiety (MD, 0.41; 95% CI, –0.21 to 1.02; P = 0.19) (Fig. 5A) and HADS Depression scores (MD, –0.13; 95% CI, –0.61 to 0.34; P = 0.58) (Fig. 5B) 4 weeks postoperatively between both arms.

DISCUSSION
To our knowledge, this is the first systematic review and meta-analysis specifically investigating postoperative outcomes of CRC patients who received trimodal PH while awaiting surgery. While the objective of PH was to improve short-term clinical outcomes, our analyses revealed no differences in these outcomes of interest, namely postoperative complication rates, length of hospital stays, and readmission rates. Several explanations are possible for these findings.

The first relates to the duration of PH. A program with a short-
er duration of 4 to 5 weeks may not be sufficient to increase physiological reserves preoperatively [15], as significant improvements in lean body mass, muscle strength, and functional capacity only occur after 12 weeks [26]. A longer duration of PH, however, may decrease patient compliance and potentially delay surgical or adjuvant therapy.

High dropout or suboptimal compliance rates in the PH arm may also explain the lack of apparent benefit. While we did consider performing meta-regression analyses against compliance rates, this was statistically not plausible given the lack of studies. Nonetheless, noncompliance could be related to the duration and complexity of PH regimes, availability of resources, as well as patient motivation and family support [17]. Li et al. [24] reported a high PH noncompliance rate of 55% from weekly phone call logs, positing that patients, especially frail and older adults, could be daunted by the physical demands of certain exercises. Moreover, the ongoing symptoms related to the physical presence of the tumor may further impede compliance. Another important consideration involves medical practitioners’ perspectives as the main advocates and executors of PH programs. A recent study suggested that considerable healthcare worker deviation or noncompliance rates for recommended components within a newly implemented ERAS program resulted from resistance to change, as well as personal beliefs and practices [27].

PH may not have had a significant impact on healthy and fit patients. The majority of patients (n=714, 68.5%), in this meta-analysis were ASA PS grade 1 or 2, evenly distributed between those who underwent PH and those who did not. Interestingly, a randomized trial from 2020 [15] also did not demonstrate the benefit of PH on early postsurgical outcomes among high-risk, frail, and older adults. It may be important to recognize that the group of patients who in theory would benefit the most from PH (i.e., frail or debilitated patients) often present at an advanced stage of disease and may require more urgent interventions. These patients, therefore, may not have the time to follow through with an exhaustive PH process. To this end, risk scoring or stratification systems may be a possible direction of future research to identify patients who are most likely to benefit from streamlined

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>PH</th>
<th>Control</th>
<th>Mean difference IV, Random, 95% CI</th>
<th>Mean difference IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD Total</td>
<td>Mean</td>
<td>SD Total</td>
</tr>
<tr>
<td>Bousquet-Dion et al. [23] (2018)</td>
<td>3.3</td>
<td>0.4</td>
<td>37</td>
<td>2.5</td>
</tr>
<tr>
<td>Carli et al. [15] (2020)</td>
<td>4.6</td>
<td>1.5</td>
<td>55</td>
<td>4.8</td>
</tr>
<tr>
<td>de Klerk et al. [25] (2021)</td>
<td>9.3</td>
<td>7.3</td>
<td>76</td>
<td>13.3</td>
</tr>
<tr>
<td>Gillis et al. [14] (2014)</td>
<td>4.0</td>
<td>0.6</td>
<td>38</td>
<td>4.5</td>
</tr>
<tr>
<td>Li et al. [24] (2013)</td>
<td>4.3</td>
<td>0.9</td>
<td>42</td>
<td>4.3</td>
</tr>
<tr>
<td>Lopez-Rodríguez-Arias et al. [12] (2021)</td>
<td>4.5</td>
<td>1.0</td>
<td>10</td>
<td>7.2</td>
</tr>
</tbody>
</table>

Total (95% CI) 258 450 100.0% -0.48 [-1.21, 0.26]

Heterogeneity: Tau^2 = 0.63, Chi^2 = 55.55, df = 5 (P < 0.00001); I^2 = 91%

Test for overall effect: Z = 1.32 (P = 0.19)

PH, prehabilitation; SD, standard deviation; IV, interval variable; Random, random-effects model; CI, confidence interval; M-H, Mantel-Haenszel test; Fixed, fixed-effect model.

Fig. 3. Forest plots. (A) Length of hospital stay in days. (B) Operation time in minutes. (C) 30-Day readmission. PH, prehabilitation; SD, standard deviation; IV, interval variable; Random, random-effects model; CI, confidence interval; M-H, Mantel-Haenszel test; Fixed, fixed-effect model.
ence in walking capacity between both groups. This was recapitulated by Bousquet-Dion et al. [23] who showed no difference in walking capacity at 8 weeks, while the control group went below the preoperative level (P < 0.01).

These findings corroborated those of Gillis et al. [14], who showed an improvement in walking capacity of +23.4 ± 4.8 m vs. –21.8 ± 80.7 m; P = 0.01 [14]. In contrast, Bousquet-Dion et al. [23] showed no difference in walking capacity between both groups.

In terms of anxiety scores, which were reported by 2 studies (n = 187) [14, 15], there were no statistically significant differences between the PH and control groups. The mean difference in HADS Anxiety scores was MD, 0.41; 95% CI, –0.21 to 1.02; P = 0.19, and in HADS Depression scores it was MD, –0.13; 95% CI, –0.61 to 0.35; P = 0.58. The reduction in visceral adipose tissue, as well as concomitant inhibition of inflammatory cytokine production, resulting in better functional capacity. This may have led to improved patient tolerance of systemic chemotherapy, with consequently fewer side effects and increased compliance. A meta-analysis of longer-term postoperative functional outcomes after PH, including return to daily activities, indicated a potential benefit of PH over the control group. These results support the benefit of PH in reducing inpatient complications and improving postoperative outcomes in colorectal cancer surgery.

**Tables and Figures**

**Table 1:** Meta-analysis of postoperative outcomes after prehabilitation for colorectal cancer patients undergoing surgery: a systematic review and meta-analysis of randomized and nonrandomized studies.

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>PH</th>
<th>Control</th>
<th>Mean difference IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carli et al. [15] (2020)</td>
<td>4.6 ± 1.4 55</td>
<td>4.0 ± 1.7 55</td>
<td>0.06 [-0.20, 0.30]</td>
</tr>
<tr>
<td>Li et al. [24] (2013)</td>
<td>4.9 ± 3.9 38</td>
<td>3.6 ± 4.1 39</td>
<td>0.60 [-1.06, 2.26]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>93 94 100.0%</td>
<td>0.41 [-0.21, 1.02]</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4:** Forest plots depicting the risk of complications. (A) Clavien-Dindo grades I and II complications. (B) Clavien-Dindo grades ≥III complications. PH, prehabilitation; M-H, Mantel-Haenszel test; Fixed, fixed-effects model; CI, confidence interval.

**Fig. 5:** Forest plots of Hospital Anxiety and Depression Scale (HADS). (A) HADS Anxiety scores. (B) HADS Depression scores. PH, prehabilitation; SD, standard deviation; IV, interval variable; Fixed, fixed-effects model; CI, confidence interval.
work, exercise tolerance, and performance of activities of daily living, was not possible in the current study due to a lack of data and heterogeneous reporting. These outcomes should be homogeneously reported in follow-up studies after PH to establish long-term functional effects.

Our findings should be interpreted in the context of known limitations. Given the small number of studies, it was deemed necessary to capture all available real-world evidence. However, this meant that the power of our overall analysis was impaired given the small sample size resulting from the lack of previous studies. Furthermore, the inclusion of non-RCTs inevitably gave rise to both statistical and qualitative heterogeneity. Selection bias from retrospective studies was possible as the patients in the PH arm may have been physically fitter or more motivated for the PH program [13, 25]. Sensitivity analyses were performed to reduce such risks. Other potential study confounders included noncompliance rates and differences in neoadjuvant therapies and surgical approaches among patients in nonrandomized studies. The small number of these studies did not permit formal meta-regression or subgroup analyses to be conducted. However, even after the exclusion of nonrandomized studies, no significant differences in short-term postoperative outcomes were demonstrated among patients who were randomized to PH and those who were not.

In a rapidly aging society globally, PH is an intriguing subject that has garnered significant interest. Incorporating PH into healthcare protocols could enhance health adjustments before surgery, resulting in potential benefits that surpass existing ERAS programs regarding recovery. While our study demonstrated that protocolized multimodal PH programs did not appear to enhance short-term postoperative outcomes following CRC surgery, the quality of evidence is impaired by the small number of included studies and heterogeneity. Therefore, further large-scale prospective trials should be done to draw definitive conclusions, establish the long-term effects of PH, and determine the specific patient groups that are most likely to benefit from PH.

**ARTICLE INFORMATION**

**Conflict of interest**
No potential conflict of interest relevant to this article was reported.

**Funding**
None.

**Author contributions**
Conceptualization: IJYW, ISE, EKWT; Data curation: all authors; Formal analysis: IJYW; Investigation: IJYW, ISE; Methodology: all authors; Project administration: all authors; Resources: all authors; Software: IJYW; Supervision: ISE, AYC, ES, CHK, WL, CM, EKWT; Validation: ISE, AYC, ES, CHK, WL, CM, EKWT; Visualization: ISE, AYC, ES, CHK, WL, CM, EKWT; Writing–original draft: IJYW, ISE, EKWT; Writing–review & editing: all authors. All authors read and approved the final manuscript.

**REFERENCES**


