



# Cardiorespiratory Fitness and Performance Adaptations to High-Intensity Interval Training: Are There Differences Between Men and Women? A Systematic Review with Meta-Analyses

Merilyn Lock<sup>1</sup> · Ibtisam Yousef<sup>1,2</sup> · Bridget McFadden<sup>3,4</sup> · Hend Mansoor<sup>1,5</sup> · Nathan Townsend<sup>1</sup>

Accepted: 14 August 2023 / Published online: 7 September 2023  
© The Author(s) 2023

## Abstract

**Background** It is important to consider biological sex as a variable that might influence exercise adaptation in order to optimize exercise prescription for men and women.

**Objective** The aim of this study was to quantify the impact of biological sex on maximal oxygen uptake ( $\dot{V}O_{2\max}$ ) and performance outcomes after high-intensity interval training (HIIT).

**Methods** A systematic search and review was conducted by two independent reviewers up to 8 September 2022 using MEDLINE, SPORTDiscus, and Sports Medicine & Education Index in ProQuest. Trials including healthy adults were included if they presented data for or compared male and female  $\dot{V}O_{2\max}$  or performance outcomes in response to HIIT. Performance outcomes included measures of exercise performance and concurrently measured physiological adaptations. Where appropriate, a random-effects, pre-post meta-analysis was undertaken. Data were sub-grouped for men and women, baseline training level, mean age, intervention type, and intervention length. Heterogeneity was assessed using  $\text{Chi}^2$ , Cochran's  $Q$ , and Higgins  $I^2$  and sensitivity analyses, where required. Study quality was assessed using the Newcastle–Ottawa Scale and publication bias was assessed through visual inspection of funnel plots.

**Results** Thirty-three references from 28 trials were included in the review ( $n = 965$ ; 462 women and 503 men). Meta-analyses included 19 studies for  $\dot{V}O_{2\max}$ , eight for peak power output from  $\dot{V}O_{2\max}$  testing (PPO), and five for threshold power ( $\text{power}_{\text{AT}}$ ). Meta-analyses revealed similar increases in  $\dot{V}O_{2\max}$  in women ( $g = 0.57$ ; 95% CI 0.44–0.69) and men ( $g = 0.57$ ; 95% CI 0.42–0.72), and  $\text{power}_{\text{AT}}$  in women ( $g = 0.38$ ; 95% CI 0.13–0.64) and men ( $g = 0.38$ ; 95% CI 0.11–0.64). Raw mean differences for change in  $\dot{V}O_{2\max}$  were  $\Delta 0.32 \text{ L}\cdot\text{min}^{-1}$  and  $3.50 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$  in men, versus  $\Delta 0.20 \text{ L}\cdot\text{min}^{-1}$  and  $3.34 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$  for women. No significant sex differences were present for the primary analysis of any outcome. After sub-grouping, significant differences were present for PPO where the effect size was higher for well-trained women ( $g = 0.37$ ) compared with well-trained men ( $g = 0.17$ ), and for  $\dot{V}O_{2\max}$  where interventions with a duration of 4 weeks or less had significantly smaller effect sizes compared with those longer than 4 weeks ( $p < 0.001$ ). Unweighted mean percentage change in  $\dot{V}O_{2\max}$ , PPO, and  $\text{power}_{\text{AT}}$  across studies was  $11.16 \pm 7.39\%$ ,  $11.16 \pm 5.99\%$ , and  $8.07 \pm 6.55\%$  for women, and  $10.90 \pm 5.75\%$ ,  $8.22 \pm 5.09\%$ , and  $7.09 \pm 7.17\%$  for men, respectively. Significant heterogeneity was present for both  $\dot{V}O_{2\max}$  and PPO ( $I^2$ , range: 62.06–78.80%). Sub-grouping by baseline training status and intervention length decreased heterogeneity in most groups. A qualitative synthesis of other outcomes indicated similar improvements in fitness and performance for men and women with some evidence suggesting differences in the mechanisms of adaptation.

**Limitations and Risk of Bias** Publication bias is unlikely to have significantly influenced results for  $\dot{V}O_{2\max}$  or  $\text{power}_{\text{AT}}$ , but the meta-analysis of PPO could have benefitted from additional study data to strengthen results. The overlap in age categories and sensitivity of the analysis limits the accuracy of the results of the sub-grouping by age.

**Conclusions** Findings indicated no sex-specific differences for any fitness or performance outcomes. Baseline training status and intervention length accounted for most variability in outcomes. PROSPERO registration number: CRD42021272615.

## 1 Introduction

High-intensity interval training (HIIT) involves repeated bouts of exercise performed in the vigorous intensity domain interspersed with periods of complete rest or low-intensity exercise. High-intensity intervals can last anywhere from 5 s to 8 min and are generally performed above the second ventilatory threshold which elicits a heart rate response between 80 to 100% of heart rate maximum [1]. Variables such as the interval duration, intensity, and the number of work and recovery bouts can be manipulated to develop a myriad of different HIIT protocols. When the work rate increases towards the upper bound of the typical intensity range for HIIT, a specific form of HIIT referred to as ‘sprint interval training’ (SIT) occurs. This involves very brief work bouts, usually around 8–30 s, that are repeated and performed at supra-maximal intensity (greater than that associated with 100% of  $\dot{V}O_{2max}$ ) or ‘all-out’ efforts [2]. These relatively short intervals are also interspersed with either long or short recovery periods [3].

Previous studies have shown that HIIT can elicit equal or greater improvements in maximal oxygen uptake ( $\dot{V}O_{2max}$ ) compared with continuous training [4–7], particularly with the use of higher intensities and longer work intervals [8, 9]. Currently, HIIT is increasingly prescribed as a potential therapeutic intervention to address a variety of chronic illnesses including cardiovascular disease, cancer, and metabolic syndrome, due to the robust evidence showing significantly enhanced cardiorespiratory fitness [6, 10–13]. The approach appears to be a viable strategy for fostering mental, psychological, and cognitive health and may reduce the severity of anxiety and depression [14–19]. Lastly, HIIT can be undertaken without the need for expensive gym equipment or access to commercial exercise training facilities. Overall, HIIT appears to be a feasible alternative to traditional endurance training for improving cardiorespiratory fitness, and may facilitate these changes with a surprisingly low training volume [20, 21].

Anatomical and physiological differences between men and women are believed to underlie differences in  $\dot{V}O_{2max}$  and endurance performance [22]; however, differences in the adaptation to chronic training are less well known. Studies in which biological sex was treated as an independent variable have been considered crucial towards improving the understanding of overall human health, and also for enabling more personalized, sex-specific training regimens [23]. Compared with men, the absolute aerobic capacity of trained women is 10–25% lower, however, when maximal oxygen uptake ( $\dot{V}O_{2max}$ ) is adjusted relative to body weight, the difference can be reduced to around 5–10% [24–26]. After normalization of body weight, the remaining difference could be due to lower blood

hemoglobin concentration, cardiac dimensions, and total blood volume [24, 27–29]. For example, women’s hearts and major blood vessels are typically smaller than those of men of the same body weight, ethnicity, and chronological age [22, 30–33]. Similarly, various studies have identified respiratory system limitations in women. Compared with men, height- and weight-matched women appear to have smaller lung sizes [34–36]. Furthermore, the diameter of the conducting airways is lower and the number of alveoli is less than in men, both of which negatively affect airflow and efficiency of gas exchange during heavy exercise [36–39]. Although some research indicates that performance could be impacted by sex differences in lung volume, but not airway anatomy and mechanics [22, 40, 41], such differences likely still contribute to physiological limitations to oxygen transport and thus would tend to exert a negative impact on exercise performance in women compared with age-, height- and/or weight-matched men.

Another physiological sex difference that has the potential to influence exercise response is that the less fatigable type I muscle fibers tend to be more abundant in women [42]. As such, there is evidence that for the same period of high-intensity exercise, women tend to experience less peripheral muscle fatigue-related contractile dysfunction than men, which translates to greater fatigue resistance and faster recovery [43]. Due to the differences in muscle fiber type percentages, women oxidize more fat and less protein and carbohydrate at matched relative intensity compared with men [44], whereas men possess higher glycolytic capacity [45, 46], which would therefore tend to alter intracellular homeostasis to a greater extent in men versus women at an equivalent relative intensity. Moreover, in response to HIIT or SIT, some studies have reported that females present with lower blood lactate levels [47]. Other findings have demonstrated that anaerobic capacity, estimated by energetic equivalents of the phosphagen and glycolytic pathways, may be lower in women when compared with men after a supramaximal effort [48]. Collectively, these studies suggest that women are less prone to peripheral muscle fatigue and have a greater tendency towards more aerobic metabolism than men.

To date, many reviews investigating the role of sex differences on acute exercise responses and chronic adaptation have been narrative in nature [2, 49]. One review concluded that attenuated blood lactate accumulation, lower protein synthesis, and mitochondrial biogenesis occur in women relative to men following SIT [2]. One systematic review, which also included a meta-regression [50], examined the effects of low-volume HIIT on cardiorespiratory fitness in adults and found moderate improvements in the  $\dot{V}O_{2max}$  of active and sedentary participants, without presenting a conclusion regarding a sex-specific response to HIIT. A more recent meta-analysis concluded that HIIT is an efficient

method of decreasing total abdominal and visceral fat mass without differences between men and women, but it did not investigate cardiorespiratory fitness as an outcome [51]. A meta-analysis by Diaz-Canestro and Montero [52] found significantly larger increases in both absolute and relative  $\dot{V}O_{2\max}$  after moderate-intensity endurance training in men compared with women; however, this review did not investigate the effects of HIIT as an intervention. Overall, the findings of these reviews indicate the potential for sex to impact health outcomes and cardiorespiratory fitness adaptations to exercise training, yet to date, no definitive conclusions can be drawn regarding how sex differences influence the adaptation to HIIT. Therefore, the objective of this systematic review with meta-analyses was to examine the influence of biological sex on the relative magnitude of adaptations in cardiorespiratory fitness and performance, following either HIIT or SIT interventions.

## 2 Methods

### 2.1 Development of the Research Question

To address the objective of the review, the research question was formulated using the Population, Intervention, Comparison, Outcome (PICO) framework as follows:

Is the relative magnitude of adaptation of maximal cardiorespiratory fitness and measures of performance (*outcome*) in response to HIIT or SIT (*intervention*) in healthy adults (*population*) different between men and women (*comparison*)?

### 2.2 Literature Search and Screening

This systematic review has been registered on the PROSPERO International Prospective Register of Systematic Reviews (registration number: CRD42021272615). Additionally, this review has been conducted and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines [53].

A search of three major electronic databases (MEDLINE, Sports Medicine & Education Index in ProQuest, and SPORTDiscus) was undertaken through to 8 September 2022. The keywords used during the search were ‘high intensity’, ‘high-intensity’, ‘HIIT’, ‘HIT’, ‘intervals’, ‘intermittent’, ‘sprint’, ‘HIIE’, ‘vigorous’, ‘maximal’, ‘exercise’, ‘workout’, ‘intervention’, ‘physical activity’, ‘activity’, ‘training’, ‘gender’, ‘sex’, ‘male’, ‘males’, ‘man’, or ‘men’ and ‘female’, ‘females’, ‘woman’, or ‘women’. Subject (MeSH) headings were used for ‘exercise’, ‘exercise training’, ‘exercise adaptation’, and ‘physical activity’ in MEDLINE and SPORTDiscus. The search strategy was recreated in the Sports Medicine & Education Index in ProQuest without the

option of subject headings. The ProQuest search was set to search ‘everything except full text’, including title, abstract, and keywords. The full search strategy as it was undertaken in MEDLINE is outlined in Supplementary Online Resource 2 (see electronic supplementary material [ESM]). In addition, the reference lists of previous reviews relevant to HIIT were manually screened to identify any relevant references that were not included in the electronic search. All references captured in the search and identified from reference lists were exported into Zotero reference management software (version 5.0.96.2, USA), and subsequently imported into Covidence online review management program (Australia) for the study selection phase of the review.

Title, abstract screening, and full-text screening were conducted through the Covidence website by two independent reviewers (IY and ML). Any conflicts during the screening process were resolved via consultation between the two reviewers to confirm the reasons underlying inclusion or exclusion. A third reviewer was available for any conflicts that could not be resolved between the first two reviewers.

For inclusion in the review, studies were required to have implemented a HIIT or SIT protocol intervention in a cohort of adults including male and female participants; to have measured cardiorespiratory fitness or performance outcomes; to have presented separate outcomes for men and women, or individual data including sex, and/or presented results of a sex  $\times$  HIIT analysis for the outcomes of interest. This approach to study inclusion was taken in order to control for confounding from different HIIT/SIT protocols across studies, and to ensure that exercise dose is normalized between male and female sub-groups. Peer-reviewed publications or unpublished theses that were written in English were included in the review. Participant groups were excluded from the review if clear pathology was present (i.e., if diagnosed diseases or disorders were a focus of the intervention) or if the intervention included major confounding factors (i.e., dietary supplementation or manipulation, pharmaceutical or herbal intervention, bed rest, or if HIIT/SIT was not the primary cardiorespiratory exercise intervention). Studies and participant groups with risk factors for disease were included if diagnosed disease states were not present. Research designs including reviews, previous meta-analyses, conference abstracts, case studies, and non-scientific articles were excluded from the review. Outcome measures included any measures relating to cardiorespiratory fitness, maximal or sub-maximal exercise performance including power, anaerobic threshold, or speed-related measures (i.e., time trials and sprints). Musculoskeletal performance outcomes such as field tests of muscular power, strength, or endurance were outside of the scope of the current review. No limitations were placed on the type of measure used for fitness or performance outcomes (i.e., measured vs estimated  $\dot{V}O_{2\max}$ ,

or lab-based vs field tests of performance) or whether or not the study achieved a positive effect overall.

### 2.3 Assessing the Risk of Bias Within Studies

The risk of bias for each of the individual included trials was evaluated independently by two authors (IY and ML) using the Newcastle–Ottawa Scale (NOS) for quality assessment of case–control studies [54]. To address the unique research question of the current review, the research team needed to focus on an observational element (biological sex differences in outcomes) within interventional studies, therefore a tool for assessing the risk of bias in observational studies was deemed to be more appropriate than a tool to assess the risk of bias within experimental studies. This approach has been previously used for another meta-analysis with a similar research question [52]. Additionally, to address the risk of bias specific to the research question, the comparison of men and women was applied to the NOS in place of cases versus controls.

Due to the risk of low inter-rater reliability associated with the subjective interpretation of the NOS and the previously highlighted need for more detailed guidance around the application of the scale [54–56], additional directions were developed by the research team to apply the NOS to the specific objectives of the current review (see Supplementary Online Resource 3 in the ESM). Studies were scored on a scale of nine in accordance with the NOS scoring system. Any conflicts in the quality rating scores of individual items within each study were resolved through discussion between the researchers. Inter-rater reliability for individual items of the NOS was calculated as the number of trials with the same score from both reviewers before conflict resolution as a proportion of the total number of trials. Domain scores were used to categorize studies into *good*, *fair*, and *poor* quality using the thresholds outlined by the Agency for Healthcare Research and Quality (AHRQ) [57].

### 2.4 Data Extraction and Meta-Analysis

Data including reference identification information, details of the participant characteristics such as age and target population, details of the intervention (intervention length, HIIT protocol, frequency, intensity, and exercise mode), methods of fitness or performance outcome testing, and results were extracted using a customized Microsoft Excel spreadsheet and Microsoft Word tables. Additionally, any concurrently measured, potentially influential physiological variables such as those relating to cardiac, muscular, or cellular metabolic adaptations, and measures of blood lactate accumulation or lactate clearance were also extracted. In cases where raw data necessary for meta-analysis was not

directly reported but could be determined from the available information, it was calculated according to Cochrane recommendations [58]. Similarly, where individual trials included multiple intervention groups that met inclusion criteria, the groups were pooled according to Cochrane recommendations [58]. In cases where outcomes were only presented in figure format, the necessary data was extracted using Web-PlotDigitizer software (version 4.5, Ankit Rohatgi, United States of America).

Where possible, pre-and post-intervention outcome data were meta-analyzed using the Meta-Essentials package (Erasmus University, the Netherlands) for Microsoft Excel [59] using differences for dependent groups and continuous data [60]. The dependent measures meta-analysis used a random-effects model and was based on Hedges' *g*. The magnitude of the effect was inferred based on the exercise science-specific thresholds of small (0.20), moderate (0.60), and large (1.20) standardized mean differences, as outlined by Hopkins and colleagues [61]. Correlation coefficients for individual studies were calculated using mean outcome and change data by applying Follmann's equation [62]. A mean of the calculated correlation coefficients from studies that provided the necessary data was used to impute a correlation coefficient for all other studies. Standardized mean differences were used to account for variation in the units of measure that were presented across different studies, however, wherever possible the most frequently reported unit of measure for each given outcome was included in the meta-analysis to maximize consistency. In cases where the first and second anaerobic thresholds were reported, the second threshold was used for inclusion in the meta-analysis.

### 2.5 Sensitivity Analyses and Methods for Exploring Heterogeneity

The primary research question was addressed by subgrouping male and female data. Sensitivity analysis was then undertaken to check the effect of study quality and any observed outlying studies on total effect size and heterogeneity within the meta-analysis. Outlying studies were removed from the analysis if they were observed to have a poor fit with the remaining studies and an underlying methodological explanation for this was identified. Studies that were categorized as poor on the NOS were excluded for sensitivity analysis to assess the effect of study quality and subsequently excluded from all meta-analyses if the effect was deemed significant. An additional sensitivity analysis was undertaken on the primary analysis of  $\dot{V}O_{2\max}$  to check the effects of the meta-analytical approach and to estimate raw mean differences for the pooled data. This analysis was undertaken in order to provide more practical estimates of baseline values and effects, as well as checking whether the use of raw mean differences over

standardized mean differences would result in any changes in the overall findings. Data were pooled in Review Manager (RevMan) version 5.4.1 (*The Cochrane Collaboration*) using a random effects model and a raw mean difference directly comparing baseline absolute and relative  $\dot{V}O_{2\max}$  for men and women, as well as pre-post measures for  $\dot{V}O_{2\max}$  (absolute and relative).

Potential sources of heterogeneity were explored through additional sub-grouping by the pre-determined population characteristics of baseline training status and mean group age, as well as by intervention type and length. Baseline training status was categorized as *untrained*, *moderately trained*, and *well trained* based on the population description by the authors of the primary studies. The grouping of studies into each training status category was confirmed by cross-checking the category against the mean baseline  $\dot{V}O_{2\max}$ , where available. In cases where the description of baseline training status was not clear within the primary study, grouping was informed by the baseline  $\dot{V}O_{2\max}$  and the homogeneity with other studies in the grouping. Untrained populations were identified as previously sedentary or those not currently participating in regular exercise at baseline. Moderately trained populations included recreationally active individuals. Well-trained populations included those described as ‘well trained’, and elite, or semi-elite athletes. Sub-grouping by age was undertaken using the mean sample group age and categorized as (a) adults under 30 years; (b) participants aged 30–45 years, and (c) participants over 45 years of age. Sub-grouping into intervention type involved categorizing study data into HIIT (interventions using sub-maximal intensities) or SIT (supra-maximal/all-out intensities). Sub-grouping by intervention length involved categorizing study data into interventions  $\leq 4$  weeks, 5–9 weeks, and  $\geq 10$  weeks in duration.

## 2.6 Qualitative Synthesis

For outcomes where only a small number of trials reported data, and meta-analytical methods were deemed to be inappropriate, results were synthesized qualitatively by grouping various measures and reporting the relevant results of each study.

## 2.7 Assessing the Risk of Bias Across Studies

The risk of bias across studies was assessed using visual inspection of the funnel plots for the primary meta-analyses [63, 64]. The presence of publication bias was assumed if notable asymmetry was present within the funnel plot. Adjusted effect sizes were reported where relevant.

## 3 Results

### 3.1 Search Results and Study Characteristics

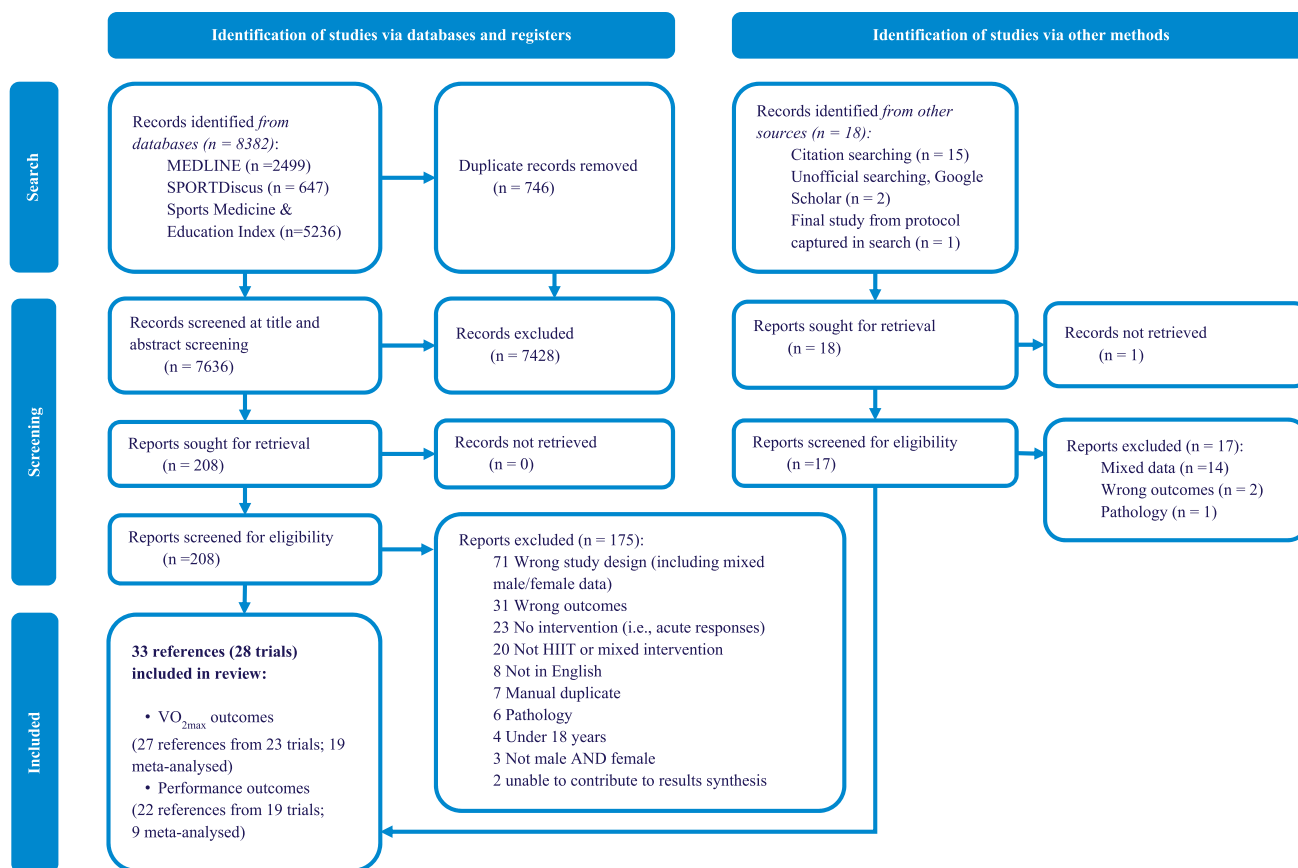
A total of 33 references from 28 individual trials including 965 participants (462 women and 503 men) were included in the review. One study was initially included in the review but was subsequently excluded due to the inclusion of only two female participants, resulting in the inability to compute an effect size in the meta-analysis [65]. A sex  $\times$  HIIT analysis was not undertaken by the study authors for the same reason and therefore despite meeting all inclusion criteria the reference could not contribute to the results of the review. The flow of references through the search and screening process is shown in the PRISMA diagram in Fig. 1. The study and population characteristics are shown in Table 1. A summary of the training protocols used is shown in Table 2. Results of the primary meta-analyses for all outcomes and the sub-grouping by participant characteristics are shown in Table 3. Results of sub-groupings by intervention characteristics are shown in Table 4. Individual study results for all fitness and performance outcomes and measures of physiological adaptation are shown in Tables 5, 6, 7 and 8. The PRISMA checklist for the reporting of review methods and results can be found in Supplementary Online Resource 4 in the ESM.

### 3.2 Risk of Bias Within Studies

Quality appraisal scores for individual studies ranged from five to eight out of a maximum possible score of nine. The majority of studies ( $n = 17$ ) were classified as good quality, indicating a low risk of bias for the current review. Eight trials were classified as fair and three trials were classified as poor. The mean inter-rater reliability for individual items of the NOS was 82.54% ( $\pm 12.17\%$ ; range 64.28–100%). The lowest inter-rater reliability scores were for the *selection of men compared to women* item and the *withdrawals and non-adherers* item. All included studies applied an equivalent intervention for males and females as indicated by question 7 on the NOS, demonstrating that prescribed exercise protocols were dose-matched between men and women within each meta-analysis. A detailed breakdown of scoring for each study can be seen in Supplementary Online Resource 5 in the ESM.

### 3.3 Correlation Coefficients

Seven studies [66–72] contributed data to the imputed correlation coefficients for  $\dot{V}O_{2\max}$ , which were calculated as 0.79 ( $\pm 0.16$ ) for women and 0.81 ( $\pm 0.10$ ) for men. Four studies [67, 71, 73, 74] contributed data for peak power output from incremental exercise testing (PPO), which were calculated



**Fig. 1** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart of the search and screening process. HIIT high-intensity interval training,  $\dot{V}O_{2max}$  maximal oxygen uptake

as  $0.84 (\pm 0.07)$  for women and  $0.81 (\pm 0.08)$  for men. Two studies [73, 75] contributed data for threshold power (power output at lactate or ventilatory threshold;  $power_{AT}$ ), which were calculated as  $0.53 (\pm 0.22)$  for women and  $0.47 (\pm 0.16)$  for men.

### 3.4 Cardiorespiratory Fitness Outcomes

#### 3.4.1 Maximal Oxygen Uptake: Study Characteristics and Primary Analysis

Twenty-eight references from 24 individual trials [66–72, 74, 76–95] presented  $\dot{V}O_{2max}$  outcomes, with 19 trials presenting sufficient data for meta-analysis. A summary of all outcomes and results relating to  $\dot{V}O_{2max}$  is shown in Table 5. All except three of the included studies measured  $\dot{V}O_{2max}$  using direct calorimetry [82, 90, 95]. Of the three studies that did not measure  $\dot{V}O_{2max}$  using direct calorimetry, only two presented sufficient data for meta-analysis [82, 90]. Effect sizes from both studies appeared to be consistent with other studies in all analyses. Out of 32 sex  $\times$  HIIT/SIT interaction analyses for  $\dot{V}O_{2max}$  that were

reported in the primary studies, 26 were not significant. Of the six that were significant, two favored females (one study, relative and absolute  $\dot{V}O_{2max}$ ) while four favored males (four studies, all absolute  $\dot{V}O_{2max}$ ). The primary meta-analysis was undertaken to address the research question by sub-grouping study data by sex only (see Fig. 2). One study [79] demonstrated a large outlying effect size for both men and women on the initial forest plot and was excluded from all subsequent analyses (see Supplementary Online Resource 6 in the ESM). It is possible this outlying study was due to the use of an arm ergometer for measurement of  $\dot{V}O_{2max}$  whereas all other studies employed cycle ergometry or running protocols. This study also used an upper limb-specific HIIT training protocol using ‘battling ropes’, compared with lower limb-specific exercise modes used in all other studies. After excluding this outlying study, heterogeneity decreased from  $I^2 = 77.26$ – $62.14\%$  for women and  $I^2 = 85.97$ – $78.80\%$  for men. The meta-analysis demonstrated near-moderate effect sizes for increasing  $\dot{V}O_{2max}$  for both men ( $g = 0.57$ ;  $p < 0.001$ ) and women ( $g = 0.57$ ;  $p < 0.001$ ) with no between-group differences ( $p = 0.97$ ). Significant levels of heterogeneity were still

Table 1 Study and population characteristics for included studies

Study	Study type	Group	Females		Males		Combined		Population	NOS (Score/9)	Categorization AHRQ
			n	Age, y (±SD)	n	Age, y (±SD)	n	Age, y (±SD)			
Astorino 2011 [76] + Astorino 2012 [66]	NRCT	SIT	9	25.2 (3.1)	11	25.3 (5.5)			Recreationally active men and women	7	Good
		CON	4	22.6 (3.1)	5	23.0 (2.7)					
Bagley 2016 [77] + Bagley 2021 [78]	NCT	SIT	17	41 (SEM 3.2)	24	38 (SEM 2.7)			Healthy adults (unhealthy subjects, or those involved in competitive sports or cycled for > 15 min/d for ≥ 3 d/wk were excluded)	6	Fair
Bornath 2022 [79]	NCT	HIIT	15	23 (1)	18	23 (2)			Recreationally active adults	7	Good
		SIT	9		6		15	21 (2)	Healthy untrained adults	7	Good
Bostad 2021 [80]	NCT	HIIT	9	18.4 (1.3)	10	20.8 (3.8)			Recreationally and moderately trained individuals involved in a variety of sports	8	Good
		SMIT	14	18.6 (1)	6	20.2 (3.1)					
Cicioni-Kolsky 2013 [98]	CON	CON	9	20.9 (4.4)	7	19.7 (1.4)					
Dalzell 2014 [82]	NCT (retrospective)	HIIT					55	51.3 (8.7)	Metabolically healthy obese <i>NOTE: study included a metabolically unhealthy obese participant group (not extracted)</i>	6	Poor
Esbjörnsson Liljedahl 1996 [97]	NRCT	SIT	10	25 (range 21–29)	6	26 (range 23–30)			Recreationally active students at a College for Sports and Recreation Instructors	7	Good
		CON	5	24 (range 22–27)	5	24 (range 23–26)					
Fisher 2017 [96] + Hoffmann 2021 [73]	NCT	HIIT	8	30 (4.8) (range 26–34)	9	29 (6.3) (range 24–34)			Trained to well-trained competitive cyclists and triathletes, aged 18–40 y	8	Good
Gillen 2014 [83]	NCT	SIT	7	30 (10)	7	29 (9)			Overweight and obese but otherwise healthy adults	7	Good
Hiam 2021 [84]	NCT	HIIT	22	34.8 (7.0)	74	31.2 (8.2)			Healthy males and healthy pre-menopausal females	7	Good
Hirsch 2021 [72]	NCT	HIIT	10	36.80 (5.59)	9	36.67 (5.96)			Overweight and obese but otherwise healthy adults (non-smokers, < 150 min/wk of moderate exercise, < 2 d/wk of resistance training, and no HIIT in the previous 12 wk)	8	Fair
Lepretre 2009 [85]	NCT	HIIT	19	65.5 (5.4)	16	64.6 (3.7)	35	65.4 (4.9)	Healthy older adults	7	Fair

Table 1 (continued)

Study	Study type	Group	Females		Males		Combined		Population	NOS (Score/9)	Cat- egorization AHRQ
			n	Age, y ( $\pm$ SD)	n	Age, y ( $\pm$ SD)	n	Age, y ( $\pm$ SD)			
Liu 2021 [86]	RCT	SIT	8	20.5 (1.4)	8	20 (1.3)			Elite badminton players	8	Good
Marterer 2020 [67]	NCT	CON	8	19.4 (1.5)	8	21.5 (2.2)			Healthy trained adults	6	Fair
Menz 2015 [87]	RCT	HIIT	11	24 (3)	9	27 (5)	20	26 (4)	Well-trained athletes ( $\dot{V}O_{2\max}$ $63.7 \pm 7.7$ mL·kg <sup>-1</sup> ·min <sup>-1</sup> , Normal exercise load: $9.5 \pm 5.0$ h/wk)	7	Good
		HIIT	5	25 (1)	14	28 (3)	19	27 (3)			
		CON	3	23 (1)	13	25 (1)	16	24 (2)			
Metcalf 2012 [88]	RCT	REHIT	8	24 (3)	7	26 (3)			Sedentary but healthy adults	8	Good
		CON	8	21 (1)	6	19 (1)					
Metcalf 2016 [68]	NCT	REHIT	18	36 (9)	17	33 (9)			Sedentary but healthy adults	8	Good
Molina-Hidalgo 2020 [69]	QRCT	HIIT (0.0% alcohol beer)	7	23.4 (4)	8	25.4 (8.1)			Healthy adults NOTE: study included HIIT + beer and HIIT + water and vodka groups (not extracted). Participants self-selected to training or non-training condition	8	Good
		HIIT (water)	7	20.9 (2.7)	9	26.9 (7.6)					
		CON	7	19.9 (2.3)	7	20.1 (2.9)					
Mucci 2004 [70]	NCT	HIIT	10	18.5 (0.7)	12	18.4 (0.7)	22	18.5 (0.5)	Non-elite sportsmen and sportswomen	8	Good
Phillips 2017 [89]	NRCT	5-by-1 HIIT					136	36 (9)	Sedentary overweight and obese but otherwise healthy adults	6	Poor
		7-by-1 HIIT					40	37 (10)			
		CON					13	31 (11)			
Sawashita 2009 [90]	NCT	High-intensity interval walking	17	60.3 (SEM: 1.6)	6	65.2 (SEM: 1.8)	23	61.6 (SEM: 1.3)	Middle-aged and older (40–69 y), overweight and obese (BMI $\geq$ 23.6 kg/m <sup>2</sup> ), Japanese adults NOTE: study included a high-interval walking and mild calorie restriction group (not extracted)	5	Poor
Scalzo 2014 [91]	NCT	SIT	10	22 (SEM: 1)	11	23 (SEM: 1)			Healthy recreationally active adults	7	Good



Table 1 (continued)

Study	Study type	Group	Females		Males		Combined		Population	NOS (Score/9)	Cat- egorization AHRQ
			n	Age, y ( $\pm$ SD)	n	Age, y ( $\pm$ SD)	n	Age, y ( $\pm$ SD)			
Schmitz 2019 [75]	NCT	Unmodified 4 × 30 s SIT	7		10		17	22.0 (1.6)	Moderately trained university students (Physical Education)	6	Fair
			8		9		17	21.1 (1.8)			
Schmitz 2020 [95]	NCT	Progressive 4–7 × 30 s SIT	14	23.1 (1.4)	7	22.1 (1.8)			Young healthy moderately trained university students (Physical Education)	5	Fair
			12	23.4 (4.4)	5	23.6 (1.3)			Moderately active (> 120 min/wk of PA), healthy adults	6	Fair
Schubert 2017a [71] + Schubert 2017b [92]	QRCT	HIIT	7		5						
			7		5						
			3		3						
Sjøgaard 2018 [93] + Chrøis 2020 [81]	NCT	HIIT	11	63 (SEM 2)	11	63 (SEM 1)			Older sedentary adults (inclusion criteria were age 55–75 years)	7	Good
			22		72		94	48.6 (18.2)	Healthy male and female matched for pre-test $\dot{V}O_{2max}$ relative to age: typically, weekly low to moderate aerobic activity (0–2 h/wk)	6	Fair
Weber 2002 [74]	NCT	HIIT	7	22.7 (SEM 2.6)	7	23.7 (SEM 1.6)			Untrained males and females	8	Good

All values are reported as mean (standard deviation) unless otherwise stated

AHRQ categorization according to Agency for Healthcare Research and Quality criteria, BMI body mass index, CON control, HIIT high-intensity interval training, NCT non-controlled trial, NOS Newcastle–Ottawa Scale, NRCT non-randomized controlled trial, PA physical activity, QRCT quasi-randomized controlled trial (some self-selected groups), RCT randomized controlled trial, REHIT reduced-exertion HIIT, SEM standard error of the mean, SIT sprint interval training, SMIT supra-maximal interval training

**Table 2** Summary of training protocols used within the included studies

Study	Session type	Exercise modality	Intervention length	Sessions/week or total	Interval training protocol	Session duration	Dose delivered vs dose received (fidelity of the intervention)
Astorino 2011 [76] + Astorino 2012 [66]	SIT	Cycle ergo	2–3 wk	6 × in total, separated by ≥ 2 d	Intervals: 4 × 30-s sprints Intensity: all-out 'Win-gate' interval Recovery: 5-min passive	Not stated	All participants completed all requirements of the protocol
Bagley 2016 [77] + Bagley 2021 [78]	CON SIT	N/A Cycle ergo	Not stated 12 wk	3 ×/wk	Intervals: 4 × 20-s sprints Intensity: ≈ 175% $\dot{V}O_{2\max}$ Recovery: 2 min	≈ 10 min	Participants maintained a training log to record workloads during training sessions. No further details
Bornath 2022 [79]	SIT	Battling ropes	6 wk	3 ×/wk	Intervals: 10 × 30-s Intensity: > 85% $HR_{\max}$ Recovery: 60 s	Not stated	Number of repetitions plus exercise intensity monitored during sessions (HR, RPE, and BLA)
Bostad 2021 [80]	SIT	Cycle ergo	12 wk	3 ×/wk, separated by ≥ 1 d	Intervals: 2 × 20-s sprints Intensity: all-out Recovery: 2 min	≈ 10 min	28 ± 3 supervised sessions completed. No further details
Cicioni-Kolsky 2013 [98]	HIIT	Running on 200-m track	6 wk	2 ×/wk	Intervals: 4 × 4 min → 5 × 4 min → 6 × 4 min (add interval every 2nd week) Intensity: 3000 m TT average pace Recovery: 1:1 work to rest ratio	Not stated	Only data of participants who completed all components of the training study were analyzed
	SMIT	Running on 200-m track	6 wk	3 ×/wk	Intervals: 7 × 30-s sprints adding 1 × sprint interval each week Intensity: 130% of 3000 m TT average pace Recovery: 150 s passive recovery	Not stated	
Dalzell 2014 [82]	HIIT	Cycle ergo	9 mo	2–3 ×/wk	Intervals: 15–30 s Intensity: 80% of maximal aerobic power Recovery: 15–30 s Sets: 2 × 10 min with 4 min passive recovery	≈ 60 min	Adherence to exercise training sessions was 3.0 ± 1.1 and 2.7 ± 1.0 sessions/week for metabolically healthy and unhealthy obese groups, respectively

Table 2 (continued)

Study	Session type	Exercise modality	Intervention length	Sessions/week or total	Interval training protocol	Session duration	Dose delivered vs dose received (fidelity of the intervention)
Esbjörnsson Liljedahl 1996 [97]	SIT	Cycle ergo	4 wk	3 ×/wk	Intervals: 3 × 30-s sprints Intensity: all-out 'Win-gate' interval Recovery: 20 min	Not stated	There was no significant difference between the sexes in relative training intensity. No further details Training diaries were completed to ensure compliance and to compare between sexes. No between-sex differences were found for end HR, RPE or peak BL <sub>a</sub> measures across the HIIT sessions
Fisher 2017 [96] + Hoffmann 2021 [73]	HIIT	Cycle ergo	6 wk	1–3 ×/wk 10 sessions total	Intervals: 10 × 90 s Intensity: 100% PPO Recovery: 60 s active recovery	Not stated	Adherence to the training sessions was 100%. Relative peak and mean power output during first and last session did not differ between men and women
Gillen 2014 [83]	SIT	Cycle ergo	6 wk	3 ×/wk	Intervals: 3 × 20 s sprints Intensity: all-out Recovery: 2 min	10 min	All participants trained under supervision. Intensity was individually determined from baseline graded exercise test. No further details
Hiam 2021 [84]	HIIT	Cycle ergo	4 wk	3 ×/wk	Intervals: pro-gression from 2 × 2 min → 16 × 2 min Intensity: 40–70% delta PPO—power at LT Recovery: 1 min at 60 W	Not stated	All training sessions were supervised one-on-one. Mean compliance to the HIIT protocol was 96 ± 9% for the entire 8 wk. Completion of at least 13 sessions was considered compliant
Hirsch 2021 [72]	HIIT	Cycle ergo	8 wk	2 ×/wk	Intervals: Week 1: 6 × 1 min Week 2: 7 × 1 min Week 3: 8 × 1 min Week 4: 9 × 1 min Weeks 5–8: 10 × 1 min Intensity: 90% of max Wattage Recovery: 1-min rest	12–20 min	Details not reported
Lepretre 2009 [85]	HIIT	Cycle ergo	9 wk	2 ×/wk	Intervals: 6 × 1 min Intensity: 2nd ventilatory threshold Recovery: 4 min at 1st ventilatory threshold	30 min	

Table 2 (continued)

Study	Session type	Exercise modality	Intervention length	Sessions/week or total	Interval training protocol	Session duration	Dose delivered vs dose received (fidelity of the intervention)
Lui 2021 [86]	SIT	Cycle ergo	8 wk	3 ×/wk	Intervals: 5 × 30 s sprints Intensity: all-out Recovery: 5 min	Not stated	HR was monitored during training to obtain HR <sub>max</sub> , time spent in %HR <sub>max</sub> zones and training impulse. Mean and peak HR, total and effective training times were similar for men and women
Marterer 2020 [67]	HIIT	Running	8 wk	Not stated	Continuous 'Fartlek' running at 65–79% HR <sub>max</sub>	40 min	Session one was supervised to reinforce the protocol. HR was recorded during sessions and monitored by the test coordinator. Participants who did not complete > 30 of 31 HIIT sessions were excluded from analysis
Menz 2015 [87]	HIIT	Running on 400-m track	3 wk	Week 1: 3 ×/wk Week 2–3: 4 ×/wk	Intervals: 4 × 4 min Intensity: 90–95% HR <sub>max</sub> Recovery: 4 min active recovery	Not stated	Sessions were supervised, and training data were recorded in a logbook. HR and RPE were monitored during sessions. All participants completed 11 sessions, except for one who completed 10
Metcalfe 2012 [88]	SIT	Cycle ergo	6 wk	3 ×/wk	Intervals: Session 1: 1 × 10 s Sessions 2–3: 2 × 10 s Sessions 4–9: 2 × 15 s Sessions 10–18: 2 × 20 s Intensity: all-out Recovery: 3–4 min low intensity	10 min in total	All training sessions were supervised. Mean adherence to the training program was 97%. Mean RPE from the whole training program was higher for women than men
Metcalfe 2016 [68]	SIT	Cycle ergo	6 wk	3 ×/wk	Intervals: Session 1: 1 × 10 s Sessions 2–3: 2 × 10 s Sessions 4–9: 2 × 15 s Sessions 10–18: 2 × 20 s Intensity: all-out Recovery: 3–4 min low intensity	10 min in total	All training sessions were supervised. Mean adherence to the training programme was 98.5%. Mean RPE was 14 ± 2 with no significant differences between men and women

Table 2 (continued)

Study	Session type	Exercise modality	Intervention length	Sessions/week or total	Interval training protocol	Session duration	Dose delivered vs dose received (fidelity of the intervention)
Molina-Hidalgo 2020 [69]	HIIT	Circuit training (e.g. horizontal rows, squats, push ups)	10 wk	2 ×/wk Separated by >48 h	Intervals: 2 × sets of 8 × exercises Intensity: 8–10 RPE Recovery (sets): time NER active recovery	40 min → 65 min	Participants that did not attend ≥ 80% of sessions were not included in the analysis. Training sessions were supervised in groups of ≈8 and participants were continuously encouraged to reach the specific target intensities
Mucci 2004 [70]	CON HIIT	N/A Running	8 wk	3 ×/wk	Sessions (each wk): Mon: 2 × sets of 13 × 5-s sprint at $V_{max}/20$ -s recovery Wed: 2 × sets of 15 × 10 s at $vVO_{2max} + 40%$ ( $V_{max} - vVO_{2max}$ )/10-s recovery Fri: 2 × sets of 6 × 30 s at $vVO_{2max} + 40%$ ( $V_{max} - vVO_{2max}$ )/90-s recovery	Not stated	The target running speeds were achieved by setting a covering distance during the work time defined in each training session (i.e., 5, 10, or 30 s). Participants who did not regularly attend sessions were excluded
Phillips 2017 [89]	5-by-1 HIIT	Cycle ergo	6 wk	3 ×/wk	Recovery (sets): 5 min Intervals: 5 × 1 min Intensity: 100% $VO_{2max}$ Recovery: 90 s	Not stated	Sessions were fully supervised. No further details
	7-by-1 HIIT	Cycle ergo	6 wk	3 ×/wk	Intervals: 7 × 1 min Intensity: 100% $VO_{2max}$ increasing 10% every 2 wk Recovery: 90 s	Not stated	
Sawashita 2009 [90]	CON HIIT	N/A Walking	6 wk 16 wk	≥ 4 ×/wk	Intervals: 5 × 3 min Intensity: 70–85% $VO_{2max}$ Recovery: 3 min ≈ 40% $VO_{2max}$	Not stated	Steps and intensity were monitored with an accelerometer with audio cues to signal the change in intensity during training. If targets were not met from accelerometer data, trainers encouraged participants to increase their efforts

Table 2 (continued)

Study	Session type	Exercise modality	Intervention length	Sessions/week or total	Interval training protocol	Session duration	Dose delivered vs dose received (fidelity of the intervention)
Scalzo 2014 [91]	SIT	Cycle ergo	3 wk	9 sessions total Separated by 1–2 d	Intervals: 4–8 × 30 s Intensity: all-out Recovery: 4-min active recovery	Not stated	Details not reported
Schmitz 2019 [75]	HIIT	Running	4 wk	3 ×/wk Separated by ≥ 1 d	Intervals: 4 × 30 s Intensity: all-out Recovery: 30-s active recovery	Not stated	The first and last training session were supervised and running speed was recorded. Reported adherence to training was $92.2 \pm 12.35\%$ for the HIIT group and $91.7 \pm 11.28\%$ for the proHIIT group
	ProHIIT	Running	4 wk	3 ×/wk Separated by ≥ 1 d	Intervals: 4 × 30 s increasing by 1 interval each week → 7 × 30 s Intensity: all-out Recovery: 30-s active recovery	Not stated	
Schmitz 2020 [95]	SIT	Running on 200-m indoor track	4 wk	2 ×/wk Separated by ≥ 2 d	Intervals: 4 × 30 s Intensity: all-out Recovery: 3-min active recovery	Not stated	Participants who completed < 6 of 8 training sessions were excluded from the analysis.
	SIT	Running on 200-m indoor track	4 wk	2 ×/wk Separated by ≥ 2 d	Intervals: 4 × 30 s Intensity: all-out Recovery: 30-s active recovery	Not stated	Training sessions were controlled by at least two experienced trainers
Schubert 2017a [71] + Schubert 2017b [92]	HIIT	Cycle ergo	4 wk	3 ×/wk	Intervals: Wk 1–2: 2 × 90 s; Wk 3–4: 4 × 90 s Intensity: 90% PPO Recovery: 1-min active recovery (10% PPO)	16–20 min	All participants completed all 12 sessions over 4 wk. Average HR during sessions were monitored to determine intensity
	SIT	Cycle ergo	4 wk	3 ×/wk	Intervals: 3 × 20 s Intensity: all-out Recovery: 2-min active recovery (10% PPO)	10–15 min	
Sjøgaard 2018 [93] + Chrøis 2020 [81]	CON	N/A	4 wk				
	HIIT	Cycle ergo	6 wk	3 ×/wk	Intervals: 5 × 1 min Intensity: 85% $\dot{V}O_{2\max}$ increasing by 10% each interval Recovery: 90-s active recovery (25 W)	Not stated	All sessions were completed by 18 participants. Three men missed one session and one woman missed two sessions. Men and women trained at a similar load

Table 2 (continued)

Study	Session type	Exercise modality	Intervention length	Sessions/week or total	Interval training protocol	Session duration	Dose delivered vs dose received (fidelity of the intervention)
Støren 2017 [94]	HIIT	Treadmill or cycle ergo	8 wk	3 ×/wk	Intervals: 4 × 4 min Intensity: 90–95% of $HR_{max}$ Recovery: 3-min active recovery at 70% $HR_{max}$	Not stated	Training sessions were supervised. HR was monitored to verify target intensity. Mean compliance to the intervention was 92% ± 4%, with no significant differences between men and women
Weber 2002 [74]	HIIT	Cycle ergo	8 wk	3 ×/wk	Intervals: 3 × 2 min Intensity: 85% → 100% of the workload = 120% $VO_{2peak}$ Recovery: 6 min	Not stated	Peak HR recorded during training at weeks 1, 4 and 8 was not different between men and women

*BLA* blood lactate concentration, *CON* non-exercise control or comparison group, *HIIT* high-intensity interval training, *HR* heart rate, *HR<sub>max</sub>* maximal heart rate, *LT* lactate threshold, *N/A* not applicable/no intervention, *NER* not explicitly reported, *PPO* peak power output on incremental test, *proHIIT* progressive HIIT, *RPE* rating of perceived exertion, *SIT* sprint interval training, *SMIT* supra-maximal interval training, *TT* time trial, *V<sub>max</sub>* maximal running velocity, *VO<sub>2max</sub>* maximal oxygen uptake, *VO<sub>2peak</sub>* peak oxygen uptake, *vVO<sub>2max</sub>* running velocity at  $VO_{2max}$ , → denotes progressive overload throughout training intervention

present for both men ( $I^2 = 78.80\%$ ,  $Q = 84.91$ ,  $p < 0.001$ ) and women ( $I^2 = 62.14\%$ ,  $Q = 47.54$ ,  $p < 0.001$ ).

Three studies [82, 89, 90] were categorized as poor using the NOS and AHRQ criteria, and only two had sufficient data for meta-analysis [82, 90]. Upon removing these studies, sensitivity analysis revealed small changes in effect sizes for  $\dot{V}O_{2max}$  in both men ( $g = 0.57$ ;  $p < 0.001$ ) and women ( $g = 0.52$ ;  $p < 0.001$ ). Similarly, only small changes in heterogeneity occurred for both men ( $I^2 = 79.91\%$ ,  $Q = 79.64$ ,  $p < 0.001$ ) and women ( $I^2 = 54.59\%$ ,  $Q = 35.23$ ,  $p = 0.004$ ). The impact of removing the data from these studies was considered to be small, and hence the data were retained for subsequent meta-analyses and sub-group analyses.

Initial qualitative analysis of  $\dot{V}O_{2max}$  outcomes revealed conflicting results regarding whether there were sex-specific differences when  $\dot{V}O_{2max}$  was considered in relative or absolute terms. When absolute and relative  $\dot{V}O_{2max}$  outcomes were retrospectively meta-analyzed separately, effect sizes were found to be similar for men and women for both outcomes, without the presence of between-group differences for either measure (Table 3; Supplementary Online Resource 7, see ESM).

The sensitivity analysis for the meta-analytical approach and for estimating raw pooled mean differences found that baseline  $\dot{V}O_{2max}$  was significantly higher in men compared with women for both absolute  $\dot{V}O_{2max}$  (between-group  $\Delta 1.06 \text{ L}\cdot\text{min}^{-1}$ ;  $p < 0.001$ ) and relative  $\dot{V}O_{2max}$  (between-group  $\Delta 5.88 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ;  $p < 0.001$ ). Heterogeneity was significant for both baseline absolute and baseline relative  $\dot{V}O_{2max}$  ( $I^2 = 70\%$  and  $60\%$  respectively;  $p = 0.001$  for both). Pre-post response to HIIT/SIT interventions measured using a raw mean difference was similar between men and women with no significant between-group differences for change in either absolute (men,  $\Delta 0.32 \text{ L}\cdot\text{min}^{-1}$  vs women,  $\Delta 0.20 \text{ L}\cdot\text{min}^{-1}$ ;  $p = 0.38$ ) or relative  $\dot{V}O_{2max}$  (men,  $\Delta 3.50 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$  versus women,  $\Delta 3.34 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ;  $p = 0.88$ ). Heterogeneity was low for the pre-post analysis of both absolute and relative  $\dot{V}O_{2max}$  ( $I^2 = 0\%$  for both;  $p = 0.97$  and  $p = 1.0$ , respectively). Forest plots for this sensitivity analysis are shown in Supplementary Online Resource 8a–d (see ESM).

### 3.4.2 Maximal Oxygen Uptake: Sub-Groupings

When  $\dot{V}O_{2max}$  data were stratified for training status (Fig. 3a), baseline training status accounted for significant levels of heterogeneity for moderately trained and well-trained men and women ( $I^2$  range 0.00–34.48;  $p$  range 1.00–0.19). Mean baseline  $\dot{V}O_{2max}$  in the moderately trained groups tended to sit between 35 and 48  $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$  in men, and 28 and 40  $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$  in women, whereas untrained groups and well-trained groups tended to sit below and above those ranges, respectively. The results of this meta-analysis showed significant differences overall with smaller effect

**Table 3** Summary of meta-analyses of  $\dot{V}O_{2max}$ , peak power output from incremental testing, and work at anaerobic threshold, primary analysis, relative and absolute  $\dot{V}O_{2max}$ , and sub-groupings by participant characteristics

Outcome	Sub-group	Within-group effects						Heterogeneity			Between-group difference			
		<i>n</i> <sub>trials</sub>	<i>n</i> <sub>participants</sub>	% $\Delta$ ( $\sigma$ )	<i>g</i>	95% CI	<i>p</i>	<i>I</i> <sup>2</sup> (%)	<i>Q</i>	<i>p</i> ( <i>Q</i> )	<i>Q</i>	<i>df</i>	<i>p</i>	
Men vs women, total														
$\dot{V}O_{2max}$	Women	19	265	11.16 (7.39)	0.57	(0.44–0.69)	<b>0.000</b>	62.14	47.54	<b>0.000</b>				
	Men	19	273	10.90 (5.75)	0.57	(0.42–0.72)	<b>0.000</b>	78.80	84.91	<b>0.000</b>	0.00	1	0.965	
PPO	Women	8	99	11.16 (5.99)	0.56	(0.32–0.80)	<b>0.000</b>	75.84	28.97	<b>0.000</b>				
	Men	8	143	8.22 (5.09)	0.41	(0.22–0.59)	<b>0.000</b>	62.06	18.45	<b>0.010</b>	1.05	1	0.304	
Work <sub>AT</sub>	Women	5	79	8.07 (6.55)	0.38	(0.13–0.64)	<b>0.001</b>	31.26	5.82	0.213				
	Men	5	123	7.09 (7.17)	0.38	(0.11–0.64)	<b>0.002</b>	29.85	5.70	0.223	0.00	1	0.964	
Men vs women, relative $\dot{V}O_{2max}$														
$\dot{V}O_{2max}$	Women	16	205	11.35 (7.70)	0.57	(0.43–0.71)	<b>0.000</b>	63.27	40.84	<b>0.000</b>				
	Men	16	249	9.88 (4.92)	0.55	(0.38–0.71)	<b>0.000</b>	80.73	77.85	<b>0.000</b>	0.04	1	0.845	
Men vs women, absolute $\dot{V}O_{2max}$														
$\dot{V}O_{2max}$	Women	7	87	10.33 (6.30)	0.56	(0.28–0.83)	<b>0.000</b>	80.33	30.50	<b>0.000</b>				
	Men	7	91	11.34 (6.28)	0.51	(0.32–0.70)	<b>0.000</b>	57.35	14.07	<b>0.029</b>	0.06	1	0.805	
Men vs women by baseline training status														
$\dot{V}O_{2max}$	Untrained	Women	11	164	13.22 (8.49)	0.63	(0.48–0.78)	<b>0.000</b>	56.81	23.15	<b>0.010</b>			
		Men	11	114	13.86 (5.23)	0.66	(0.48–0.85)	<b>0.000</b>	68.34	31.59	<b>0.000</b>			
	Moderately trained	Women	5	72	8.40 (6.12)	0.33	(0.15–0.51)	<b>0.000</b>	34.48	6.11	0.191			
		Men	5	130	4.37 (1.68)	0.18	(0.09–0.28)	<b>0.000</b>	0.00	3.39	0.494			
	Well trained	Women	3	29	8.23 (1.20)	0.72	(0.41–1.04)	<b>0.000</b>	23.73	2.62	0.269			
		Men	3	29	10.92 (1.14)	0.86	(0.86–0.87)	<b>0.000</b>	0.00	0.00	1.00			
	All groups										48.74	5	<b>0.000</b>	
PPO	Untrained	Women	4	44	13.99 (6.73)	0.77	(0.36–1.18)	<b>0.000</b>	79.57	14.68	<b>0.002</b>			
		Men	4	41	12.23 (3.61)	0.60	(0.31–0.89)	<b>0.000</b>	54.58	6.61	<b>0.086</b>			
	Moderately trained	Women	2	36	9.96 (5.25)	0.35	(0.01–0.68)	<b>0.036</b>	54.31	2.19	0.139			
		Men	2	84	5.86 (0.65)	0.27	(0.14–0.41)	<b>0.000</b>	0.00	1.00	0.317			
	Well trained	Women	2	19	6.71 (3.67)	0.37	(0.26–0.48)	<b>0.000</b>	0.00	0.25	0.618			
		Men	2	18	2.56 (1.52)	0.17	(0.15–0.19)	<b>0.000</b>	0.00	0.01	0.927			
	All groups										11.04	5	<b>0.051</b>	
Men versus women by group mean age														
$\dot{V}O_{2max}$	Under 30 y	Women	10	93	9.49 (3.82)	0.63	(0.47–0.79)	<b>0.000</b>	35.93	14.05	0.121			
		Men	10	95	11.54 (5.65)	0.71	(0.54–0.89)	<b>0.000</b>	49.67	17.88	<b>0.037</b>			
	30–45 y	Women	5	81	10.63 (6.74)	0.39	(0.20–0.57)	<b>0.000</b>	52.12	8.35	<b>0.079</b>			
		Men	5	134	8.04 (6.78)	0.40	(0.04–0.75)	<b>0.000</b>	88.73	35.51	<b>0.000</b>			
	Over 45 y	Women	4	91	15.99 (13.52)	0.68	(0.35–1.00)	<b>0.000</b>	72.45	10.89	<b>0.012</b>			
		Men	4	44	12.88 (4.64)	0.49	(0.21–0.77)	<b>0.000</b>	52.66	6.34	<b>0.096</b>			
	All groups										8.60	5	0.126	

*p*(*Q*) significance set at <0.10; effect size (*g*) significance set at *p*=0.05. Bolded *p* values indicate statistical significance

% $\Delta$  percentage change (unweighted mean across studies),  $\sigma$  standard deviation (of % $\Delta$ ), *CI* confidence interval, *df* degrees of freedom, *g* effect size (Hedges' *g*), *I*<sup>2</sup> Higgins *I*<sup>2</sup>, *PPO* peak power output from incremental exercise testing, *Q* Cochran's *Q* (sum of squares),  $\dot{V}O_{2max}$  maximal oxygen uptake, *Work*<sub>AT</sub> threshold power (inclusive of lactate and ventilatory thresholds)

sizes present in the moderately trained groups, but effect sizes for men and women were similar (Table 3).

Sub-group analysis for men and women by the mean age of the participant group (Fig. 3b) did not account for significant levels of heterogeneity in any of the sub-groups, with

the exception of women under 30 years (*I*<sup>2</sup>=35.93; *p*=0.12). All significant heterogeneity in the men's 30–45-year-old group was accounted to one study [72]. The exclusion of this study resulted in a substantially lowered effect size for this group (*g*=0.18) and a significant between-group



**Table 4** Summary of meta-analyses of  $\dot{V}O_{2max}$  and peak power output from incremental testing, sub-groupings by intervention characteristics

Outcome	Sub-group	Within-group effects					Heterogeneity					Between-group difference		
		$n_{trials}$	$n_{participants}$	$\% \Delta (\sigma)$	$g$	95% CI	$p$	$I^2$ (%)	$Q$	$p(Q)$	$Q$	$df$	$p$	
Men vs women, HIIT versus SIT														
$\dot{V}O_{2max}$	HIIT	Women	11	177	11.00 (9.29)	0.52	(0.34– 0.71)	<b>0.000</b>	76.06	41.77	<b>0.000</b>			
		Men	11	172	11.20 (5.54)	0.60	(0.37– 0.82)	<b>0.000</b>	85.98	71.32	<b>0.000</b>			
SIT		Women	9	96	10.86 (4.19)	0.59	(0.46– 0.71)	<b>0.000</b>	0.00	6.77	0.562			
		Men	9	93	9.82 (6.32)	0.46	(0.27– 0.65)	<b>0.000</b>	52.58	16.87	<b>0.040</b>			
All groups														
PPO	HIIT	Women	7	131	8.26 (5.50)	0.48	(0.28– 0.68)	<b>0.000</b>	69.41	19.61	<b>0.003</b>	1.36	3	0.716
		Men	7	85	10.10 (5.69)	0.41	(0.20– 0.62)	<b>0.000</b>	66.08	17.69	<b>0.007</b>			
SIT		Women	2	14	7.16 (1.14)	0.90	(0.14– 1.66)	<b>0.011</b>	73.79	3.82	<b>0.051</b>			
		Men	2	12	16.09 (2.71)	0.33	(–0.01 to 0.68)	<b>0.033</b>	9.94	1.11	0.292			
All groups														
Men vs women by intervention length														
$\dot{V}O_{2max}$	$\leq 4$ weeks	Women	4	55	5.82 (2.37)	0.25	(0.11– 0.39)	<b>0.000</b>	0.00	2.20	0.532			
		Men	4	106	3.97 (1.63)	0.16	(0.05– 0.27)	<b>0.003</b>	0.00	2.51	0.474			
5–9 weeks		Women	10	109	9.39 (4.23)	0.55	(0.40– 0.70)	<b>0.000</b>	29.97	12.85	0.169			
		Men	10	103	11.24 (3.64)	0.67	(0.50– 0.85)	<b>0.000</b>	58.70	21.79	<b>0.010</b>			
$\geq 10$ weeks		Women	5	101	18.98 (9.42)	0.79	(0.63– 0.94)	<b>0.000</b>	19.55	4.97	0.290			
		Men	5	64	15.76 (6.27)	0.66	(0.32– 1.00)	<b>0.000</b>	80.41	20.42	<b>0.000</b>	59.35	5	<b>0.000</b>
All groups														

$p(Q)$  significance set at  $< 0.10$ ; effect size ( $g$ ) significance set at  $p=0.05$ . Bolded  $p$  values indicate statistical significance

$\% \Delta$  percentage change (unweighted mean across studies),  $\sigma$  standard deviation (of  $\% \Delta$ ),  $CI$  confidence interval,  $df$  degrees of freedom,  $g$  effect size (Hedges'  $g$ ),  $I^2$  Higgins  $I^2$ ,  $PPO$  peak power output from incremental exercise testing,  $Q$  Cochran's  $Q$  (sum of squares),  $\dot{V}O_{2max}$  maximal oxygen uptake,  $Work_{AT}$  threshold power (inclusive of lactate and ventilatory thresholds)

Table 5 Summary of outcomes and results of included studies: maximal oxygen uptake

References	Measurement	Outcome (units)	Sex, baseline	$\Delta M$	$\Delta F$	Interaction
<b>Maximal oxygen uptake</b>						
Astorino 2011 [76] + Astorino 2012 [66]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup> $\dot{V}O_{2max}$ (L/min)	NS M > F ( $p < 0.05$ )	+ + (5.9 ± 3.9%)	+ + (6.8 ± 7.1%)	NS NS ( $p = 0.58$ )
Bagley 2016 [77] + Bagley 2021 [78]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup> $\dot{V}O_{2max}$ (L/min) <sup>a</sup>	( $p = 0.005$ ) ( $p < 0.001$ )	+ (6.0%) +	+ (18.7%) +	F > M ( $p = 0.049$ ) F > M ( $p = 0.009$ )
Bornath 2022 [79]	Arm cycle test, indirect calorimetry (Astrand protocol)	$\dot{V}O_{2max}$ —arms (mL/kg/min), 3 wk $\dot{V}O_{2max}$ —arms (mL/kg/min), 6 wk	NER NER	+ (ES: <0.45; $p < 0.004$ ) + (ES: 0.66; $p < 0.001$ )	+ (ES: 0.67; $p < 0.001$ ) + (ES: 0.48; $p = 0.001$ )	NER NER
Bostad 2021 [80]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2peak}$ (L/min), 2 wk $\dot{V}O_{2peak}$ (L/min), 6 wk	NER NER	+ ( $p = 0.03$ ) + ( $p < 0.01$ )	NS + ( $p < 0.01$ )	M > F ( $p < 0.01$ ) NS
Dalzell 2014 [82]	Incremental treadmill test	$\dot{V}O_{2peak}$ (L/min), 12 wk <sup>a</sup>	M > F ( $p < 0.001$ )	+ ( $p < 0.001$ )	+ ( $p < 0.0001$ )	NS
Gillien 2014 [83]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2peak}$ (mL/kg/min) <sup>a</sup> $\dot{V}O_{2peak}$ (L/min; baseline only)	NS ( $p \leq 0.05$ )	+ (12%; $p < 0.001$ )	+ (12%; $p < 0.001$ )	NS
Hiam 2021 [84]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2peak}$ (mL/kg/min) <sup>a</sup>	M > F ( $p = 0.04$ )	NER	NER	NS ( $p = 0.48$ )
Hirsch 2021 [72]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2peak}$ (mL/kg/min), 0–4 wk $\dot{V}O_{2peak}$ (mL/kg/min), 4–8 wk	M > F ( $p = 0.03$ )	NS (6.8%) NS (12.7%)	NS (−0.5%) + (16.6%; $p < 0.05$ )	NS NS
Lepretre 2009 [85]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2peak}$ (mL/kg/min), 0–8 wk <sup>a</sup>	M > F ( $p < 0.05$ )	+ (19.5%; $p < 0.05$ )	+ (16.1%; $p < 0.05$ )	NS
Liu 2021 [86]	Treadmill test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup>	NER	+ (ES: 1.14; $p < 0.05$ )	+ (14.5%; $p < 0.001$ )	NS
Marterer 2020 [67]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2max}$ —legs (mL/kg/min) <sup>a</sup> $\dot{V}O_{2max}$ —legs (mL/min) <sup>a</sup>	NER NER	+ ( $p < 0.001$ ) + (9.96%; $p < 0.001$ )	+ ( $p = 0.002$ ) + (7.95%; $p = 0.002$ )	NS ( $p = 0.088$ ) M > F ( $p = 0.007$ )
Menz 2015 [87]	Arm cycle test, indirect calorimetry	$\dot{V}O_{2max}$ —arms (mL/kg/min) $\dot{V}O_{2max}$ —arms (mL/min)	NER NER	NS + ( $p = 0.031$ )	NS NER	NS ( $p = 0.267$ ) NS ( $p = 0.134$ )
Metcalf 2012 [88]	Treadmill test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min) $\dot{V}O_{2max}$ (mL/min)	NER NER	NER NER	NER NER	NS, data pooled NS, data pooled
	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2peak}$ (mL/kg/min) <sup>a</sup>	M > F ( $p < 0.001$ )	+ (15.0%; $p < 0.01$ )	+ (12.0%; $p < 0.01$ )	NS

Table 5 (continued)

References	Measurement	Outcome (units)	Sex, baseline	$\Delta M$	$\Delta F$	Interaction
Metcalf 2016 [68]	Cycle ergo test, Douglas bag method	$\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup>	M > F ( $p = 0.010$ )	+ ( $p < 0.001$ )	+ ( $p < 0.001$ )	NS ( $p = 0.926$ )
Molina-Hidalgo 2020 [69]	Treadmill test, indirect calorimetry (modified Balke)	$\dot{V}O_{2max}$ (L/min) <sup>a</sup> $\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup> $\dot{V}O_{2max}$ (mL/min) <sup>a</sup>	M > F ( $p < 0.001$ ) NER NER	+ ( $p < 0.001$ ) + ( $p < 0.05$ ) + ( $p < 0.05$ )	+ ( $p < 0.001$ ) + ( $p < 0.05$ ) + ( $p < 0.01$ )	NS ( $p = 0.402$ ) NER NER
Mucci 2004 [70]	Multi-stage running test, portable gas analysis	$\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup>	M > F ( $p < 0.05$ )	+ (10.0%; $p < 0.001$ )	+ (7.8%; $p < 0.01$ )	NS
Phillips 2017 [89]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min; $\Delta\%$ only)	NER	NER	NER	NS
Sawashita 2009 [90]	Estimated from graded walking test, triaxial accelerometer and ECG	$\dot{V}O_{2max}$ (L/min; 5-by-1 HIIT)	NER	+ (0.32 $\pm$ 0.3 L/min)	+ (0.19 $\pm$ 0.2 L/min)	M > F ( $p < 0.001$ )
Scalzo 2014 [91]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2peak}$ (mL/kg/min) <sup>a</sup>	NER	+ ( $p < 0.001$ )	+ ( $p < 0.001$ )	NS
Schmitz 2020 [95]	Estimated from maximal running speed (Léger and Boucher equation)	$\dot{V}O_{2max}$ (mL/kg/min)	NER	NER	NER	NER, baseline only
Schubert 2017a [71] + Schubert 2017b [92]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup> $\dot{V}O_{2max}$ (L/min; pooled only)	NER	NER	NER	NER, individual data presented
Søgaard 2018 [93] + Chrøst 2020 [81]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min) <sup>a</sup> $\dot{V}O_{2max}$ (mL/min) <sup>a</sup>	M > F ( $p < 0.012$ ) M > F ( $p < 0.001$ )	+ ( $p = 0.024$ ) + ( $p = 0.030$ )	+ ( $p = 0.024$ ) + ( $p = 0.030$ )	NS ( $p = 0.277$ ) NS ( $p = 0.473$ )
Støren 2017 [94]	Treadmill or cycle ergometer test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min; pooled only) $\dot{V}O_{2max}$ (L/min; pooled only)	NER NER	+ (4.1 $\pm$ 2.5 mL/kg/min; $p < 0.001$ ) + (10.8 $\pm$ 8.1%; $p < 0.001$ )	+ (4.2 $\pm$ 2.5 mL/kg/min; $p < 0.001$ ) + (13.8% $\pm$ 10.0%; $p < 0.001$ )	NS ( $p = 0.300$ ) NER
Weber 2002 [74]	Cycle ergo test, indirect calorimetry	$\dot{V}O_{2max}$ (mL/kg/min, base-line only) $\dot{V}O_{2peak}$ (L/min) <sup>a</sup>	M > F ( $p < 0.05$ ) M > F ( $p < 0.001$ )	NER + (7.9 $\pm$ 2.0%; $p < 0.01$ )	NER NS (2.9 $\pm$ 1.5%)	NER M > F ( $p < 0.05$ )

$\Delta$  indicates change (i.e., pre-post), + denotes a significant increase; ES effect size, F women/females, M men/males, MET metabolic equivalent, NER not explicitly reported, NS not significant,  $\dot{V}O_{2max}$  maximal oxygen uptake,  $\dot{V}O_{2peak}$  peak oxygen uptake

<sup>a</sup>Included in meta-analyses

**Table 6** Summary of outcomes and results of included studies, peak power output from incremental testing and threshold power

References	Measurement	Outcome (units)	Sex, baseline	$\Delta$ M	$\Delta$ F	Interaction
Peak power output from incremental testing						
Bostad 2021 [80]	Cycle ergo test, indirect calorimetry	$W_{\text{peak}}$ (W; for ExRx only)	NER	NER	NER	NER
Fisher 2017 [96] + Hoffmann 2021 [73]	Cycle ergo test, earlobe capillary blood sampling	Absolute PPO (W) <sup>a</sup> Relative PPO (W·kg <sup>-1</sup> and W·kg <sup>-0.32</sup> )	M > F ( $p < 0.001$ ) M > F ( $p < 0.001$ )	+( $p = 0.05$ ) +( $p = 0.04$ )	+( $p = 0.05$ ) +( $p = 0.04$ )	NS NS
Gillen 2014 [83]	Cycle ergo test, indirect calorimetry	Maximal workload (W; baseline)	M > F ( $p \leq 0.05$ )	+(14%; $p < 0.001$ )	+(14%; $p < 0.001$ )	NS
		Relative PPO (W/kg FFM) <sup>a</sup>	NER	+(0.9 W/kg FFM; $p \leq 0.05$ )	+(1.8 W/kg FFM; $p \leq 0.05$ )	NER
		Relative MPO (W/kg FFM)	NER	+(1.6 W/kg FFM; $p \leq 0.05$ )	+(3.0 W/kg FFM; $p \leq 0.05$ )	NER
Hiam 2021 [84]	Cycle ergo test, indirect calorimetry	PPO (W/kg) <sup>a</sup>	M > F ( $p = 0.030$ )	NER	NER	NS ( $p = 0.650$ )
Lepretre 2009 [85]	Cycle ergo test, indirect calorimetry	MTP (W) <sup>a</sup>	M > F ( $p \leq 0.05$ )	+( $p < 0.001$ )	+( $p < 0.001$ )	NS
Marterer 2020 [67]	Cycle ergo or arm cycle test, earlobe capillary blood sampling	PPO—legs (W) <sup>a</sup>	NER	+( $p = 0.033$ )	+( $p = 0.026$ )	NS ( $p = 0.562$ )
		PPO—arms (W)	NER	NS ( $p = 0.176$ )	NS ( $p = 0.498$ )	NS ( $p = 0.563$ )
Phillips 2017 [89]	Incremental cycle ergo test	PPO/ $W_{\text{max}}$ (W; pooled only)	NER	NER	NER	NER
Schubert 2017a [71] + Schubert 2017b [92]	Cycle ergo test, highest workload completed for 60 s	PPO (W) <sup>a</sup>	NER	NER	NER	NER, individual data presented
Søgaard 2018 [93] + Chrøis 2020 [81]	Cycle ergo test, indirect calorimetry	Maximum workload (W) <sup>a</sup>	M > F ( $p < 0.001$ )	+(31 W; $p < 0.001$ )	NS (9 W)	M > F ( $p = 0.004$ )
Støren 2017 [94]	Treadmill or cycle ergometer test, indirect calorimetry	Work performance (W)	NER	+(24.5 ± 34.4%; $p < 0.001$ )	+(23.8 ± 43.7%; $p < 0.001$ )	NS ( $p = 0.980$ )
Weber 2002 [74]	Cycle ergo test, indirect calorimetry	PPO (W) <sup>a</sup>	M > F ( $p < 0.001$ )	+(10.7 ± 2.0%; $p < 0.01$ )	+(11.2 ± 1.1%; $p < 0.01$ )	NS
Threshold power						
Fisher 2017 [96] + Hoffmann 2021 [73]	Cycle ergo test, earlobe capillary blood sampling	LT <sub>2</sub> power output (W) <sup>a</sup>	M > F ( $p < 0.001$ )	NS	NS	NS
Hiam 2021 [84]	Cycle ergo test, indirect calorimetry	Power output at LT (W/kg) <sup>a</sup>	M > F ( $p = 0.040$ )	NER	NER	NS ( $p = 0.410$ )
Lepretre 2009 [85]	Cycle ergo test, indirect calorimetry	Power at VT <sub>1</sub> (W)	M > F ( $p \leq 0.05$ )	+(29.0%; $p < 0.001$ )	+(32.5% $p < 0.001$ )	NS
		Power at VT <sub>1</sub> (%MTP)	NS	+(12.0%; $p < 0.001$ )	+(9.0%; $p < 0.001$ )	NS
		Power at VT <sub>2</sub> (W) <sup>a</sup>	M > F ( $p \leq 0.05$ )	+( $p < 0.001$ )	+( $p < 0.001$ )	NS
		Power at VT <sub>2</sub> (%MTP)	M > F ( $p \leq 0.05$ )	NS	NS	NS

**Table 6** (continued)

References	Measurement	Outcome (units)	Sex, baseline	$\Delta$ M	$\Delta$ F	Interaction
Marterer 2020 [67]	Cycle ergo or arm cycle test, earlobe capillary blood sampling <i>LT defined as 4 mmol/L of BLa</i>	Power output at LT—legs (W/kg)	NER	NS ( $p=0.291$ )	NS ( $p=0.059$ )	NS ( $p=0.878$ )
		Power output, at LT—legs (W) <sup>a</sup>	NER	NS ( $p=0.301$ )	+ ( $p=0.054$ )	NS ( $p=0.962$ )
		Power output at LT—arms (W/kg)	NER	NS ( $p=0.714$ )	NS ( $p=0.195$ )	NS ( $p=0.384$ )
		Power output, at LT—arms (W)	NER	NS ( $p=0.580$ )	NS ( $p=0.182$ )	NS ( $p=0.178$ )
Schmitz 2019 [75]	Incremental running test, earlobe capillary blood sampling	Speed at LT (m/s) <sup>a</sup> , pooled HIIT + proHIIT	NER	+ (0.13 m/s)	+ (0.10 m/s)	NS ( $p=0.09$ )
Schmitz 2020 [95]	Incremental running test, capillary blood sampled rest + immediate post	Speed at LT (km/h; baseline only)	M > F ( $p < 0.001$ )	NER	NER	NER

+ denotes a significant increase;  $\Delta$  denotes change (i.e., pre-post); *BLa* blood lactate concentration, *ExRx* exercise prescription, *F* women/females, *FFM* fat free mass, *LT* lactate threshold, *LT<sub>2</sub>* second lactate threshold, *M* men/males, *MTP* maximal tolerated power, *NER* not explicitly reported, *NS* not significant, *PPO* peak power output, *VT<sub>1</sub>* first ventilatory threshold, *VT<sub>2</sub>* second ventilatory threshold, *HIIT* high-intensity interval training, *proHIIT* progressive HIIT, *W* Watts, *W<sub>max</sub>* maximum Watts, *W<sub>peak</sub>* peak Watts

<sup>a</sup>Included in meta-analyses

difference overall ( $p < 0.001$ ), indicating high sensitivity and a general lack of robustness within this particular analysis. Despite the application of the age categories, participants over 45 years actually only consisted of studies with a mean age of  $\geq 59$  years.

Sub-grouping by intervention type comparing HIIT and SIT protocols demonstrated no significant between-group differences ( $p = 0.72$ ; Fig. 4a), whereas sub-grouping by intervention length demonstrated a significant between-group difference ( $p < 0.001$ ) with significantly smaller effect sizes present in both men and women for interventions with a duration of 4 weeks or less compared to those with a longer duration (Fig. 4b).

### 3.5 Performance Outcomes

A multitude of different performance outcomes were presented in the included studies. Fourteen references from 11 individual trials presented measures of PPO [67, 71, 73, 74, 80, 81, 83–85, 89, 92–94, 96], and seven references from six individual trials presented measures of power<sub>AT</sub> ( $n = 6$ ) [67, 73, 75, 84, 85, 95, 96]. A summary of the outcomes and results for PPO and power<sub>AT</sub> is outlined in Table 6. Forest plots for the meta-analyses of PPO, and PPO sub-grouped by baseline training status and intervention type, are shown in Figs. 5, 6a, b, respectively. The forest plot for power<sub>AT</sub> is

shown in Fig. 7. Power and  $\dot{V}O_2$ -based outcomes that were not meta-analyzed included peak power output during training sessions ( $n = 2$ ) [80, 91], threshold  $\dot{V}O_2$  ( $n = 4$ ) [67, 72, 85, 86], Wingate outcomes ( $n = 2$ ) [66, 76, 97], and relative and absolute power output (peak and at lactate threshold) from an arm cycle protocol ( $n = 1$ ) [67]. Additional performance measures included time trials for running ( $n = 1$ ) [98] and cycling ( $n = 3$ ) [73, 80, 91], maximal speed ( $n = 2$ ) [70, 95], 40-m sprint ability ( $n = 1$ ) [98], repeated sprint ability ( $n = 2$ ) [86, 98], fatigue ( $n = 4$ ) [73, 76, 80, 93], and speed decrement ( $n = 1$ ) [95].

#### 3.5.1 Peak Power Output from Incremental Testing

Eight trials reporting peak power output from incremental testing were meta-analyzed [67, 74, 83–85, 92, 93, 96]. All trials tested PPO using a cycle ergometer protocol. Results demonstrated significant increases in PPO for all female and male subgroups. Women consistently demonstrated larger percent increases (6.71–13.99%) and effect sizes ( $g$ , range 0.35–0.77) for PPO compared to men (2.56–12.23%;  $g$ , range: 0.17–0.60), and the between-group difference reached the threshold for statistical significance in the sub-grouping by baseline training status, due to the larger

Table 7 Summary of outcomes and results of included studies, additional cardiorespiratory and performance outcomes and measures of fatigue

References	Measurement	Outcome (units)	Sex, baseline	$\Delta M$	$\Delta F$	Interaction
<b>Threshold oxygen uptake</b>						
Hirsch 2021 [72]	Cycle ergo test, indirect calorimetry	VT (L/min), 0–4 weeks VT (L/min), 4–8 weeks VT (L/min), 0–8 weeks	Main effect ( $p=0.004$ )	NS NS +( $p<0.05$ )	NS +( $p<0.05$ ) NS	NS NS NS
Lepretre 2009 [85]	Cycle ergo test, indirect calorimetry	VT <sub>1</sub> (% $\dot{V}O_{2peak}$ ) VT <sub>2</sub> (% $\dot{V}O_{2peak}$ )	M > F ( $p=0.054$ ) NS ( $p=0.362$ )	+( $p=0.011$ ) NS ( $p=0.362$ )	+( $p=0.011$ ) NS ( $p=0.362$ )	NS NS
Liu 2021 [86]	Treadmill test, indirect calorimetry	VT (% $\dot{V}O_{2max}$ )	NER (M: 74.0%; F: 75.0%)	+(11.0%; $p<0.05$ )	+(7.0%; $p<0.05$ )	NER
Marterer 2020 [67]	Cycle ergo test, earlobe capillary blood sampled end of each stage	VT (mL/min) $\dot{V}O_2$ at LT—legs (mL/min) $\dot{V}O_2$ at LT—legs (mL/min/kg)	NER NER NER	+(19.0%; $p<0.05$ ) NS ( $p=0.097$ ) NS ( $p=0.088$ )	+(11.0%; $p<0.05$ ) NS ( $p=0.058$ ) NS ( $p=0.249$ )	NER NS ( $p=0.427$ ) NS ( $p=0.960$ )
	Arm cycle test, earlobe capillary blood sampled end of each stage	$\dot{V}O_2$ at LT—arms (mL/min) $\dot{V}O_2$ at LT—arms (mL/min/kg)	NER NER	NS ( $p=0.859$ ) NS ( $p=0.844$ )	NS ( $p=0.191$ ) NS ( $p=0.296$ )	NS ( $p=0.296$ ) NS ( $p=0.465$ )
<b>Wingate power</b>						
Astorino 2011 [76] + Astorino 2012 [66]	Wingate test (repeated; two tests performed with 5 min of active recovery in between)	Maximal power (W/kg; pooled) Mean power (W/kg; pooled) Minimum power (W/kg; pooled)	M > F ( $p<0.05$ ) M > F ( $p<0.05$ ) NS	+(10.5%; $p<0.05$ ) +(10.4%; $p<0.05$ ) +(14.9%; $p<0.05$ )	+(9.1%; $p<0.05$ ) +(10.9%; $p<0.05$ ) +(9.5%; $p<0.05$ )	NS NS NS
Esbjörnsson Liljedahl 1996 [97]	Wingate test (repeated; three tests performed with a 20-min rest in between)	Maximal power (W), 1st test Maximal power (W), 2nd test Maximal power (W), 3rd test Mean power (W), 1st test Mean power (W), 2nd test Mean power (W), 3rd test Absolute power difference between M and F, mean of 3 tests	NER NER NER NER NER NER NER M > F (47.0%)	NS NS NS NS NS NS NS N/A	NS NS NS +(10.0%; $p<0.01$ ) +(6.0%; $p<0.05$ ) +(4.0%; $p<0.05$ ) N/A	NS NS NS F > M ( $p<0.05$ ) F > M ( $p<0.05$ ) NS Post: M > F (38.0%; $p<0.03$ )

Table 7 (continued)

References	Measurement	Outcome (units)	Sex, baseline	$\Delta$ M	$\Delta$ F	Interaction
<b>Fatigue and speed decrement</b>						
Astorino 2011 [76] + Astorino 2012 [66]	Wingate test (repeated; two tests performed with 5 min of active recovery in between)	Fatigue index (%; pooled)	M > F (M: 55.2%; F: 47.3%; $p < 0.05$ )	NS (-3.3%)	NS (0.0%)	NS
Scalzo 2014 [91]	The mean power from sprints during sessions 1 and 9	Fatigue index (%; session 1–9)	NS (M: 28.0%; F: 13.0%; $p < 0.0066$ )	-(12%)	-(3%)	NS ( $p = 0.066$ )
Schmitz 2020 [95]	Incremental running test	Maximal speed (km/h; baseline only)	M > F ( $p < 0.001$ )	NER	NER	NER
		Repeated sprint speed decrement, 4 × 30:30	NER	NS	-(3.4%; $p = 0.0038$ )	F > M ( $p = 0.0038$ )
		Repeated sprint speed decrement, 4 × 30:180	NER	NS	NS	NS
Fisher 2017 [96] + Hoffmann 2021 [73]	Cycle ergo test	Incremental time to fatigue (min)	F > M ( $p < 0.01$ )	+(1.8%; $p = 0.01$ )	+(5.1%; $p = 0.01$ )	NS
Weber 2002 [74]	Timed cycling test to fatigue at 120% $VO_{2peak}$	Time to fatigue (s), $\Delta$ pre-mid training	NS	+(30.6 ± 3.4%; $p < 0.01$ )	+(29.0 ± 3.2%; $p < 0.01$ )	NER
		Time to fatigue (s), $\Delta$ mid-post training	NS	+(12.4 ± 3.7%; $p < 0.05$ )	+(9.6 ± 3.9%; $p < 0.05$ )	NER
<b>Time trials, maximal performance, and sprint ability</b>						
Bostad 2021 [80]	Average power output from all sprints and all training sessions	Sprint power output (% $W_{peak}$ )	N/A	N/A	N/A	NS (M: 237% $W_{peak}$ ; F: 222% $W_{peak}$ ; $p = 0.25$ )
	2 kJ/kg body weight cycle ergo time trial (distance NER)	Time trial completion, 2 wk	NER	-( $p < 0.01$ )	-( $p < 0.01$ )	NER, pooled only
		Time trial completion, 6 wk		-( $p = 0.14$ )	-( $p = 0.14$ )	
		Time trial completion, 12 wk		-( $p < 0.01$ )	-( $p < 0.01$ )	
Cictoni-Kolsky 2013 [98]	Maximal, 640 m repeated sprint ability test (sum of six sprints)	40-m maximal sprint (s), HIIT	NER	NS (1.28%; $p = 0.091$ )	-(1.95%; $p = 0.013$ )	NS (HIIT + SMIT, $p = 0.680$ )
		40-m maximal sprint (s), SMIT	NER	-(2.99%; $p = 0.010$ )	-(3.73%; $p < 0.001$ )	
	2 kJ/kg body weight cycle ergometer time trial	Sum of six sprints (s), HIIT	NER	-(2.86%; $p = 0.020$ )	-(4.90%; $p = 0.001$ )	NS (HIIT + SMIT, $p = 0.593$ )
		Sum of six sprints (s), SMIT	NER	-(5.16%; $p < 0.001$ )	-(7.89%; $p < 0.001$ )	F > M (HIIT + SMIT, $p = 0.012$ )

Table 7 (continued)

References	Measurement	Outcome (units)	Sex, baseline	$\Delta M$	$\Delta F$	Interaction
Fisher 2017 [96] + Hoffmann 2021 [73]	3000-m time-trial on a HIIT	3000-m time-trial (s), HIIT	NER	- (7.35%; $p = 0.022$ )	- (7.73%; $p = 0.001$ )	
		3000-m time-trial (s), SMIT	NER	- (5.88%; $p < 0.035$ )	- (9.50%; $p < 0.001$ )	
	40km time trial on a cycle ergometer	Time trial time (min)	F > M ( $p < 0.01$ )	- (1.70%; $p < 0.01$ )	- (2.69%; $p < 0.01$ )	NS ( $p = 0.10$ )
		Time trial power output (W)	M > F ( $p < 0.01$ )	+ (4.96%; $p < 0.01$ )	+ (7.49%; $p < 0.01$ )	NS ( $p = 0.14$ )
Liu 2021 [86]	Incremental cycle test $LT_2$ determined using the $D_{max}$ method	Correlational analysis for $\Delta LT_2$ power output	N/A	NS ( $\Delta$ time trial performance; $r^2 = 0.07$ ; $p = 0.49$ )	+ (related to $\Delta$ time trial performance $r^2 = 0.77$ ; $p = 0.004$ )	N/A
		Correlational analysis and $\Delta LT_2$ , BLA, $\Delta$ absolute + relative PPO, $\Delta$ incremental time to fatigue	N/A	NS ( $\Delta$ time trial performance for all)	NS ( $\Delta$ time trial performance for all)	N/A
Mucci 2004 [70]	60-m running speed field test, 20-m intervals with photocells at every 20 m	YO-YO IR2, or distance (m)	NER	+ (ES: 0.68; $p < 0.05$ )	+ (ES: 0.68; $p < 0.05$ )	NER
		Maximal running speed, $V_{max}$ (m/s; max from 3 tests)	M > F ( $p < 0.001$ )	NS	NS	NER
Scalzo 2014 [91]	The mean power from repeated sprints during sessions 1 and 9	Speed at $\dot{V}O_{2max}$ , $v\dot{V}O_{2max}$ (m/s)	M > F ( $p < 0.05$ )	+ ( $p < 0.01$ )	+ ( $p < 0.001$ )	NER
		Absolute mean power from 4 sprints, sessions 1 and 9 (W)	M > F (all $p < 0.001$ except 4th sprint, session 1, NS)	+ 3rd sprint session 9 vs 1 only ( $p = 0.004$ )	NS for all	M > F ( $p = 0.049$ )
Schmitz 2019 [75]	40-km time trial on a stationary cycle ergometer	Relative mean power from sprints, sessions 1 and 9 (W/kg)	NS ( $p = 0.301$ )	NS	NS	NS ( $p = 0.120$ )
		Time to cycle 40 km (min)	F > M ( $p = 0.004$ )	NS	NS	NS ( $p = 0.269$ )
Schmitz 2020 [95]	Incremental running test	Maximal speed (m/s)	NER	NER	NER	NER, pooled only
		Maximal speed (km/h; baseline only)	M > F ( $p < 0.001$ )	NER	NER	NER, baseline only

$\Delta$  indicates change (i.e., pre-post); + denotes a significant increase; - denotes a significant decrease; BLA blood lactate concentration, ES effect size, F women/females, HIIT high-intensity interval training, LT lactate threshold,  $LT_2$  second lactate threshold, M men/males, N/A not applicable, NER not explicitly reported, NS not significant, PPO peak power output, SMIT supra-maximal interval training,  $D_{max}$  maximum-deviation,  $V_{max}$  maximal running velocity,  $\% \dot{V}O_{2max}$  percent of maximal oxygen uptake,  $\% \dot{V}O_{2peak}$  percent of peak oxygen uptake, VT ventilatory threshold,  $VT_1$  first ventilatory threshold,  $VT_2$  second ventilatory threshold,  $v\dot{V}O_{2max}$  running velocity at  $\dot{V}O_{2max}$ ,  $\% W_{peak}$  percent of peak Watts



effect size for well-trained women ( $g = 0.37$ ) compared with well-trained men ( $g = 0.17$ ;  $p = 0.05$ ). Baseline training status accounted for all significant heterogeneity in PPO in moderately trained and well-trained men and women ( $I^2$ , range: 0.00–54.31%;  $p$ , range: 0.14–0.93); however, significant levels of heterogeneity were present for the total sample and untrained sub-groups ( $I^2$ , range: 54.58–79.57%;  $p$ , range: 0.00–0.09). Due to the smaller number of studies, PPO could not be sub-grouped by mean group age or intervention length.

### 3.5.2 Meta-Analysis of Threshold Power

Five trials [67, 73, 75, 84, 85] presented sufficient data to meta-analyze outcomes relating to power<sub>AT</sub>. No differences were demonstrated between men and women ( $p = 0.96$ ). The percent increase in power<sub>AT</sub> for men was  $7.09 \pm 7.17\%$  (small effect size:  $g = 0.38$ ;  $p < 0.01$ ), and that for women was  $8.07 \pm 6.55\%$  (small effect size:  $g = 0.38$ ;  $p < 0.01$ ). Some inconsistency existed in the units presented for these outcomes (e.g., work presented as Watts, W/kg, and speed in m/s) and the measures of anaerobic thresholds (lactate thresholds and ventilatory thresholds both included); however, despite this, heterogeneity was not significant, and the grouping of these outcomes appeared to be appropriate. Results demonstrated small increases in power<sub>AT</sub> for men and women with low heterogeneity (men:  $I^2 = 29.85\%$ ,  $Q = 5.70$ ,  $p = 0.22$ ; women:  $I^2 = 31.26\%$ ,  $Q = 5.82$ ,  $p = 0.21$ ). Due to the small number of studies presenting relevant data, outcomes for power<sub>AT</sub> could not be further sub-grouped.

### 3.5.3 Additional Performance Outcomes

A summary of results for additional performance outcomes and measures of fatigue is shown in Table 7. Most performance outcomes showed no significant differences in the magnitude of improvement between men and women. In cases where significant sex  $\times$  HIIT interactions existed, these included a greater improvement in mean and maximal Wingate power output [97], repeated sprint speed decrement [95], and a 3000-m cycling time trial [98] for women compared to men. Additionally, one study demonstrated a significant correlation between the change in power output at the second lactate threshold (LT<sub>2</sub>) and the change in 40-km time trial performance for women ( $r^2 = 0.77$ ;  $p < 0.01$ ) [73], while no relationships with any of the measured variables were present for men ( $r^2 = 0.01$ –0.21;  $p$  all  $< 0.05$ ). One study demonstrated a greater improvement in men for the mean power of the third of four repeated sprints within SIT

sessions (pertaining to less power decrement over repeated sprints) [91].

## 3.6 Concurrent Measures of Physiological Adaptation

Various physiological adaptations that were measured alongside other fitness and performance outcomes were reported in 15 trials [67, 70, 74, 76, 78–80, 83, 85–87, 91, 93, 96, 97]. These included maximal accumulated oxygen deficit (MAOD;  $n = 1$ ) [74], various blood lactate measures ( $n = 8$ ) [67, 70, 73, 74, 79, 85, 86, 95, 96], cardiac adaptations ( $n = 1$ ) [80], mitochondrial and metabolic adaptations ( $n = 3$ ) [81, 91, 93, 97], muscle fiber types ( $n = 1$ ) [97], and correlational analyses for fitness or performance outcomes ( $n = 4$ ) [67, 73, 80, 96]. A summary of the results relating to these concurrent measures of physiological adaptation is shown in Table 8.

Sex  $\times$  HIIT interactions for most blood lactate and cellular or muscular measures were either not reported or not significant. Cases where significant interactions indicated greater increases in women compared to men after HIIT included maximal [67] or post-test blood lactate [70], type II muscle fiber cross-sectional area [97], and muscle glycogen content [97]. Conversely, significant interactions where men demonstrated greater increases compared to women included total muscle creatine content [97], muscle fiber  $\beta$ -HAD activity and GLUT4 protein content [83], coupled and uncoupled mitochondrial respiratory capacity [81], muscle mitochondrial biogenesis [91], and muscle protein synthesis [91]. All significant sex  $\times$  HIIT interactions relating to central cardiorespiratory measures other than  $\dot{V}O_{2\max}$  indicated greater changes for men compared to women following HIIT. These included significant interactions for increases in maximum carbon dioxide output [76], peak cardiac output [80], peak stroke volume [80], peak cardiac index [80], maximal minute ventilation [70], oxygen pulse [67], and accumulated oxygen uptake [74], and decreases in accumulated oxygen deficit [74].

## 3.7 Assessment of Publication Bias

The funnel plots for the meta-analyses of  $\dot{V}O_{2\max}$ , PPO, and power<sub>AT</sub> are shown in Fig. 8a–c. Some asymmetry can be seen on the funnel plots for  $\dot{V}O_{2\max}$  and PPO where a lack of data points can be seen at the negative effect size area of the plot. The combined male and female adjusted effect size for  $\dot{V}O_{2\max}$  was calculated as 0.32 (95% confidence interval [CI] 0.27–0.38) compared with the observed effect size of 0.49 (95% CI 0.43–0.55). The combined adjusted effect size for PPO was calculated as 0.25 (95% CI 0.19–0.32) compared with the observed effect size of 0.41 (95% CI 0.33–0.49).

Table 8 Summary of outcomes and results of included studies, concurrent measures of physiological adaptation

References	Measurement	Outcome (units)	Sex, baseline	$\Delta$ M	$\Delta$ F	Interaction
<b>Blood lactate measures</b>						
Bornath 2022 [79]	BLa (pre- and post-) for 9th and 18th HIIT sessions $\Delta =$ within session change	BLa, $\Delta$ from rest-imm post, 9th BLa, $\Delta$ from rest-imm post, 18th BLa, $\Delta$ from imm-5 min post, 9th BLa, $\Delta$ from imm-5 min post, 18th	NS for rest NS for imm post NS for 5 min post N/A	+(ES: 0.95; $p < 0.001$ ) +(ES: 0.95; $p < 0.001$ ) -(ES: 0.21; $p = 0.048$ ) -(ES: 0.29; $p = 0.017$ )	+(ES: 0.84; $p < 0.001$ ) +(ES: 0.88; $p < 0.001$ ) NS -(ES: 0.53; $p < 0.001$ )	NER NER NER NER
Fisher 2017 [96] + Hoffmann 2021 [73]	Incremental cycle test, ear-lobe blood sampling $LT_2$ determined with $D_{max}$ method	$LT_2$ BLa (mmol/L)	NS	NS	NS	NS ( $p = 0.19$ )
Lepretre 2009 [85]	Maximal incremental cycle ergo test, capillary blood sampling	BLa (mmol 100 mL <sup>-1</sup> ) same absolute intensity (MTP pre-) BLa (mmol 100 mL <sup>-1</sup> ) same relative intensity (MTP pre- vs post-)	M > F ( $p < 0.05$ ) M > F ( $p < 0.05$ )	-(45.03%; $p < 0.05$ ) NS	-(21.24%; $p < 0.05$ ) NS	NER NER
Liu 2021 [86]	Treadmill test, BLa tested at rest, 0-, 1-, 3-, 5-, 7-, and 10-min post-test	Lactate clearance rate (%)	NER	+(ES: 2.26; $p < 0.05$ )	+(ES: 1.10; $p < 0.05$ )	NER
Marterer 2020 [67]	Cycle or arm cycle ergo test, earlobe blood sampled end of each stage $LT$ defined as 4 mmol/L BLa	Maximal BLa—legs (mmol/L) Maximal BLa—arms (mmol/L)	NER NER	NS (-6.56%; $p = 0.083$ ) NS ( $p = 0.214$ )	+(13.33%; $p = 0.017$ ) NS ( $p = 0.920$ )	F > M ( $p = 0.003$ ) NS ( $p = 0.470$ )
Mucci 2004 [70]	Multi-stage running test, fingertip blood sampled pre- and post	Post-test BLa (mmol/L)	NS	NS	+( $p < 0.05$ )	F > M ( $p < 0.05$ )
Weber 2002 [74]	BLa during session 1, wk 1 [baseline] and session 3, wks 4 [mid] and 8 [post]	BLa, resting (mmol/L) BLa, 3-min post-session (mmol/L) BLa, 3-min post-timed cycle (mmol/L)	M > F M > F ( $p < 0.01$ ) M > F ( $p < 0.01$ )	NER +(9.6 $\pm$ 3.0%; $p < 0.05$ ) -( $p < 0.05$ )	NER +9.6 $\pm$ 4.6%; $p < 0.05$ ) -( $p < 0.05$ )	NS NS NER
<b>Cellular and muscular adaptations</b>						
Bagley 2016 [77] + Bagley 2021 [78]	Magnetic resonance imaging and isokinetic dynamometer	Quadriceps CSA (cm <sup>2</sup> ) Absolute MVC torque (Nm) Relative MVC torque (Nm·cm <sup>-2</sup> ) Quadriceps fatigue index (%)	M > F ( $p < 0.001$ ) M > F ( $p < 0.001$ ) NS ( $p = 0.248$ ) NS ( $p = 0.282$ )	+(4.06%; $p = 0.023$ ) +(2.76%; $p = 0.079$ ) NS ( $p = 0.712$ ) +(7.67%; $p = 0.048$ )	+(5.82%; $p = 0.023$ ) +(5.62%; $p = 0.079$ ) NS ( $p = 0.712$ ) +(15.88%; $p = 0.048$ )	NS ( $p = 0.140$ ) NS ( $p = 0.385$ ) NS ( $p = 0.705$ ) NS ( $p = 0.127$ )

Table 8 (continued)

References	Measurement	Outcome (units)	Sex, baseline	$\Delta$ M	$\Delta$ F	Interaction	
Esbjörnsson Lijedahl 1996 [97]	Muscle needle biopsy	Type II fiber CSA ( $\mu\text{m}^2$ )	M > F ( $p < 0.01$ )	NS	+ (17%; $p < 0.05$ )	F > M ( $p < 0.01$ )	
		Type II fiber (%)	NS	NS	NS	NS	
		% and CSA of other fiber types	NS	NS	NS	NS	NS
		Glycogen content	NS	NS	+ ( $p < 0.01$ )	+ ( $p < 0.01$ )	F > M ( $p < 0.05$ )
		Total creatine content	NS	+ ( $p < 0.05$ )	NS	NS	M > F ( $p < 0.05$ )
		Lactate dehydrogenase activity	M > F ( $p < 0.05$ )	+ ( $p < 0.05$ )	+ ( $p < 0.05$ )	NS	NS
		PFK, CS or HAD	NS	NS	NS	NS	NS
		CS activity	NS	+	+	+	NS
		$\beta$ -HAD activity	F > M ( $p < 0.05$ )	+	NS	NS	M > F ( $p < 0.05$ )
		COX IV protein content	NS	+	+	+	NS
Søgaard 2018 [93] + Chrøis 2020 [81]	Muscle biopsy, Western blotting	GLUT4 protein content	NS	+ (138%; $p < 0.01$ )	+ (23%; $p < 0.01$ )	M > F ( $p < 0.05$ )	
		Mitochondrial respiratory capacity:					
		Coupled	NS	+ ( $p < 0.001$ )	NS	NS	M > F ( $p = 0.003$ )
		Uncoupled	NS	+ ( $p < 0.001$ )	NS	NS	M > F ( $p = 0.003$ )
		CS activity	NS	+ ( $p < 0.001$ )	+	+ ( $p = 0.013$ )	NS
		Adenosine diphosphate sensitivity	NS	- ( $p = \text{NER}$ )	-	- ( $p = \text{NER}$ )	NS ( $p = 0.096$ )
		Creatine kinase content	NS	NS	NS	NS	NS
		Muscle glycogen content	NS	+ ( $p = 0.001$ )	+ ( $p = 0.001$ )	+ ( $p = 0.001$ )	NS
		GLUT4	NS	+ ( $p < 0.050$ )	+ ( $p < 0.050$ )	+ ( $p < 0.050$ )	NER
		Glycogen synthase	NS	+ ( $p = 0.001$ )	+ ( $p = 0.001$ )	+ ( $p = 0.001$ )	NER
Scalzo 2014 [91]	Muscle biopsy, D <sub>2</sub> O labelling (heavy water), Western blotting	Hexokinase II	NS	+ ( $p < 0.050$ )	+ ( $p < 0.050$ )	NER	
		Protein expression (multiple)	NS	Varied	Varied	NS	NS
		Muscle mitochondrial biogenesis	NER	+ ( $p = \text{NER}$ )	+	NER	M > F ( $p = 0.056$ )
		Muscle protein synthesis	M > F ( $p < 0.001$ )	+ ( $p = \text{NER}$ )	+	NER	M > F ( $p = 0.056$ )
		CS protein content	NS	+ ( $p < 0.050$ )	+ ( $p < 0.050$ )	+ ( $p < 0.050$ )	NER
		COX IV protein content	F > M	NS	NS	NS	NS
		PGC1 $\alpha$ protein content	NS	+ ( $p < 0.050$ )	+ ( $p < 0.050$ )	+ ( $p < 0.050$ )	NS

Table 8 (continued)

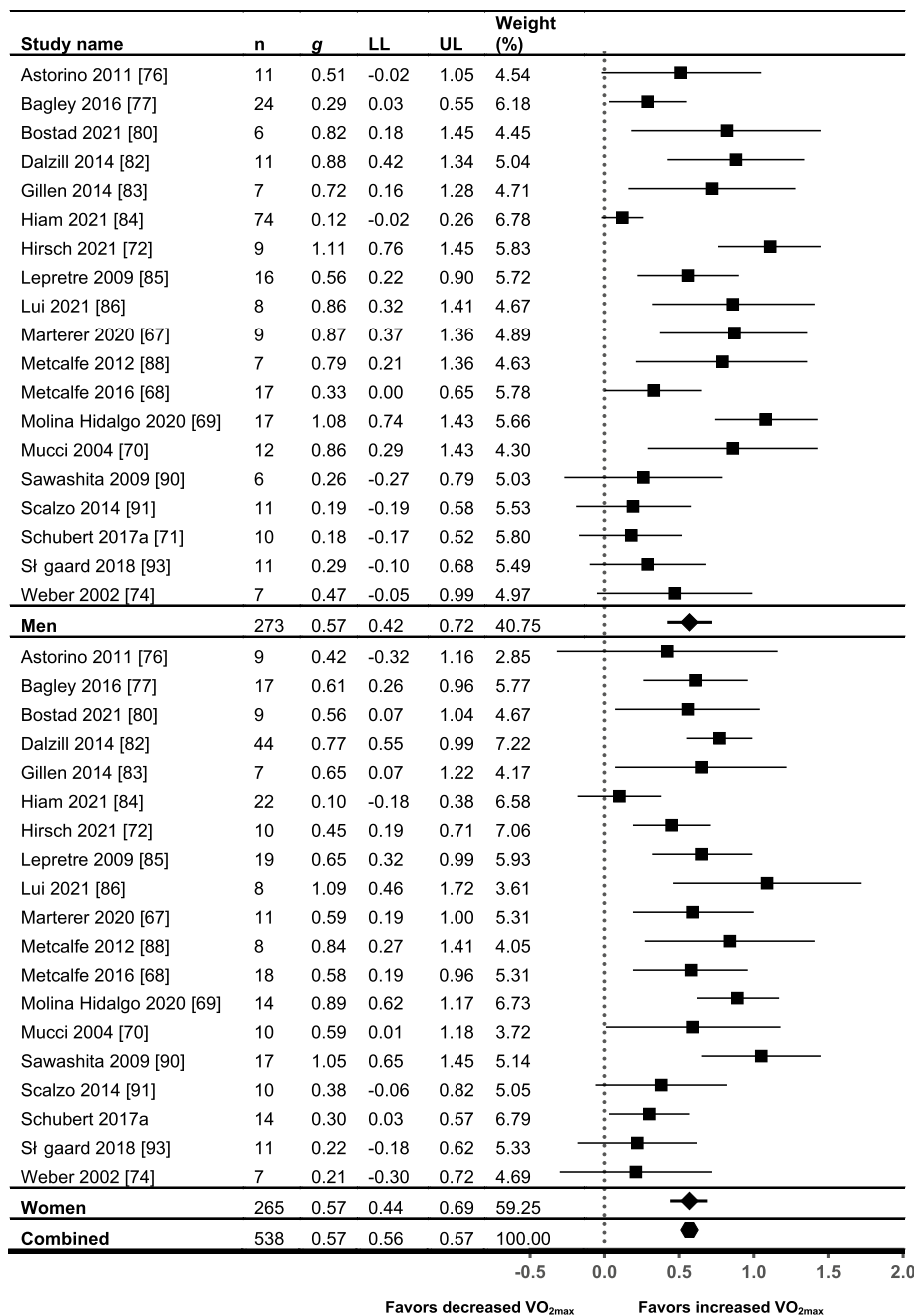
References	Measurement	Outcome (units)	Sex, baseline	$\Delta M$	$\Delta F$	Interaction
Central cardiorespiratory adaptations (other than $\dot{V}O_{2max}$ )						
Astorino 2011 [76]	Cycle ergo test, indirect calorimetry	$VCO_{2max}$ (L/min) $RER_{max}$ $HR_{peak}$ (bpm) $VE_{max}$ (L/min)	M > F ( $p < 0.05$ ) NS NS M > F ( $p < 0.05$ )	+ ( $p < 0.050$ ) NS NS NS	+ ( $p < 0.050$ ) NS NS NS	M > F ( $p < 0.050$ ) NS NS NS
Bostad 2021 [80]	Cycle ergo test, indirect calorimetry Inert gas rebreathing test, approx. 20 min post $\dot{V}O_{2peak}$ test <i>Peak cardiac index calculated as <math>Q_{peak}/body\ surface\ area</math></i>	Oxygen pulse (mL/beat) $Q_{peak}$ , 6 weeks $Q_{peak}$ , 12 weeks $SV_{peak}$ , 6 weeks $SV_{peak}$ , 12 weeks Peak cardiac index, 6 weeks Peak cardiac index, 12 weeks $a-\dot{V}O_{2diff}$ , 6 weeks $a-\dot{V}O_{2diff}$ , 12 weeks $HR_{peak}$ (bpm) Correlational analysis, $\Delta \dot{V}O_{2peak}$ , $Q_{peak}$ , or calculated peak $a-\dot{V}O_{2diff}$	M > F ( $p < 0.05$ ) NER NER NER NER NER NER NER NER NER NER NER NER N/A	+ ( $p < 0.050$ ) + ( $p < 0.01$ ) + (10%; $p < 0.01$ ) + ( $p < 0.01$ ) + ( $p < 0.01$ ) + ( $p < 0.01$ ) + ( $p < 0.01$ ) + ( $p < 0.05$ ) + ( $p < 0.01$ ) NS NS correlation to $\Delta$ time to test completion	+ ( $p < 0.050$ ) NS NS (+0.6%; $p = 0.96$ ) NS NS NS NS NS + ( $p < 0.05$ ) + ( $p < 0.01$ ) NS NS correlation to $\Delta$ time to test completion	NS ( $p = 0.30$ ) NER M > F ( $p < 0.01$ ) NER M > F ( $p < 0.01$ ) NER M > F ( $p < 0.01$ ) NER M > F ( $p < 0.01$ ) NER NS NS ( $p = 0.98$ ) NS ( $p = 0.13$ ) N/A
Lepretre 2009 [85]	Maximal incremental cycle ergo test with indirect calorimetry	$VE_{max}$ (L/min)	M > F ( $p < 0.05$ )	+ (7.75%; $p < 0.05$ )	+ (16.93%; $p < 0.05$ )	NER
Mucci 2004 [70]	Multi-stage running test, portable gas analysis and fingertip blood sampled pre- and post <i>EIH defined as a decrease of &gt; 4% in oxygen saturation during the last stage of exercise testing</i>	$VE_{max}$ (L/min) $HR_{peak}$ (bpm) $VE/VO_2$ $VE/VCO_2$ Developed EIH post-HIIT Did not develop EIH post-HIIT	M > F ( $p < 0.01$ ) NS NS NS NER NER	+ (5.77%; $p < 0.05$ ) - (2.47%; $p < 0.01$ ) - (5.62%; $p < 0.05$ ) - (6.25%; $p < 0.05$ ) $n = 7$ $n = 5$	NS (+3.72%) NS (-0.63%) NS (-3.96%) NS (-4.73%) $n = 2$ $n = 8$	M > F ( $p < 0.001$ ) NS NS NS NER NER

Table 8 (continued)

References	Measurement	Outcome (units)	Sex, baseline	$\Delta M$	$\Delta F$	Interaction
Marterer 2020 [67]	Cycle ergo test, indirect calorimetry	Oxygen pulse—legs (mL/beat)	NER	+ (11.36%; $p < 0.001$ )	+ (8.94%; $p = 0.001$ )	M > F ( $p = 0.008$ )
		HR <sub>peak</sub> —legs	NER	NS ( $p = 0.251$ )	NS ( $p = 0.234$ )	NS ( $p = 0.982$ )
	Arm cycle test, indirect calorimetry	Oxygen pulse—arms (mL/beat)	NER	+ (7.39%; $p = 0.041$ )	+ (9.09%; $p = 0.004$ )	NS ( $p = 0.432$ )
		HR <sub>peak</sub> —arms	NER	NS ( $p = 0.833$ )	- ( $p = 0.013$ )	NS ( $p = 0.096$ )
		RER <sub>max</sub> —legs and arms	NER	NS	NS	NS
	Correlational analysis, $\dot{V}O_{2max}$ —legs and $\dot{V}O_{2max}$ —arms	N/A	NS correlation to blood volume or Hb mass	NS correlation to blood volume or Hb mass	NS	
Menz 2015 [87]	Carbon monoxide rebreathing method and capillary blood samples	Total hemoglobin mass (g and g/kg)	NER	NER	NER	NS, data pooled
		Blood volume (ml/kg)	NER	NER	NER	NS, data pooled
		Plasma volume (ml/kg)	NER	NER	NER	NS, data pooled
		Maximal heart rate (bpm)	NER	NER	NER	NS, data pooled
		Maximal oxygen pulse (ml/beat)	NER	NER	NER	NS, data pooled
		RER <sub>max</sub>	NER	NER	NER	NS, data pooled
		MAOD (L), $\Delta$ pre-mid training	M > F ( $p < 0.01$ )	+ (14.3 $\pm$ 5.2%; $p < 0.05$ )	+ (14.0 $\pm$ 3.0%; $p < 0.01$ )	NS
Weber 2002 [74]	Timed cycling test to fatigue at 120% $\dot{V}O_{2peak}$ , indirect calorimetry	MAOD (L), $\Delta$ mid-post training	M > F ( $p < 0.01$ )	+ (6.6 $\pm$ 1.9%; $p < 0.01$ )	+ (5.1 $\pm$ 2.3%; $p < 0.05$ )	NS
		AO <sub>2</sub> deficit (L)	M > F ( $p < 0.01$ )	- ( $p < 0.01$ )	NS	M > F ( $p = \text{NER}$ )
	AO <sub>2</sub> demand and AO <sub>2</sub> uptake	AO <sub>2</sub> uptake (L)	M > F ( $p < 0.05$ )	+ ( $p < 0.01$ )	NS	M > F ( $p = \text{NER}$ )
		HR <sub>peak</sub> (bpm)	NS	- ( $p < 0.05$ )	- ( $p < 0.05$ )	NER

+ denotes a significant increase; - denotes a significant decrease;  $\Delta$  indicates change (i.e., pre-post); AO<sub>2</sub> accumulated oxygen (deficit, demand, uptake),  $a\text{-}\dot{V}O_{2diff}$  arteriovenous oxygen difference, *BLA* blood lactate concentration, *CS* citrate synthase, *CSA* cross-sectional area, *EIH* exercise-induced hypoxemia, *ES* effect size, *F* women/females, *HAD* 3-hydroxyacyl CoA dehydrogenase,  $\beta\text{-HAD}$  3- $\beta$ -hydroxyacyl CoA dehydrogenase, *Hb* hemoglobin, *HIIT* high-intensity interval training, *HR<sub>peak</sub>* peak (maximal) heart rate, *imm* immediate, *LT<sub>2</sub>* second lactate threshold, *D<sub>max</sub>* maximum-deviation, *M* men/males, *MAOD* maximal accumulated oxygen deficit, *MTP* maximal tolerated power, *MVC* maximal voluntary contraction, *N/A* not applicable, *NER* not explicitly reported, *NS* not significant, *PFK* phosphofructokinase, *Q<sub>peak</sub>* peak cardiac output, *RER<sub>max</sub>* maximal respiratory exchange ratio, *SV<sub>peak</sub>* peak stroke volume, *VCO<sub>2</sub>* carbon dioxide production, *VCO<sub>2max</sub>* maximal carbon dioxide production, *VE<sub>max</sub>* maximal minute ventilation, *VO<sub>2</sub>* oxygen uptake

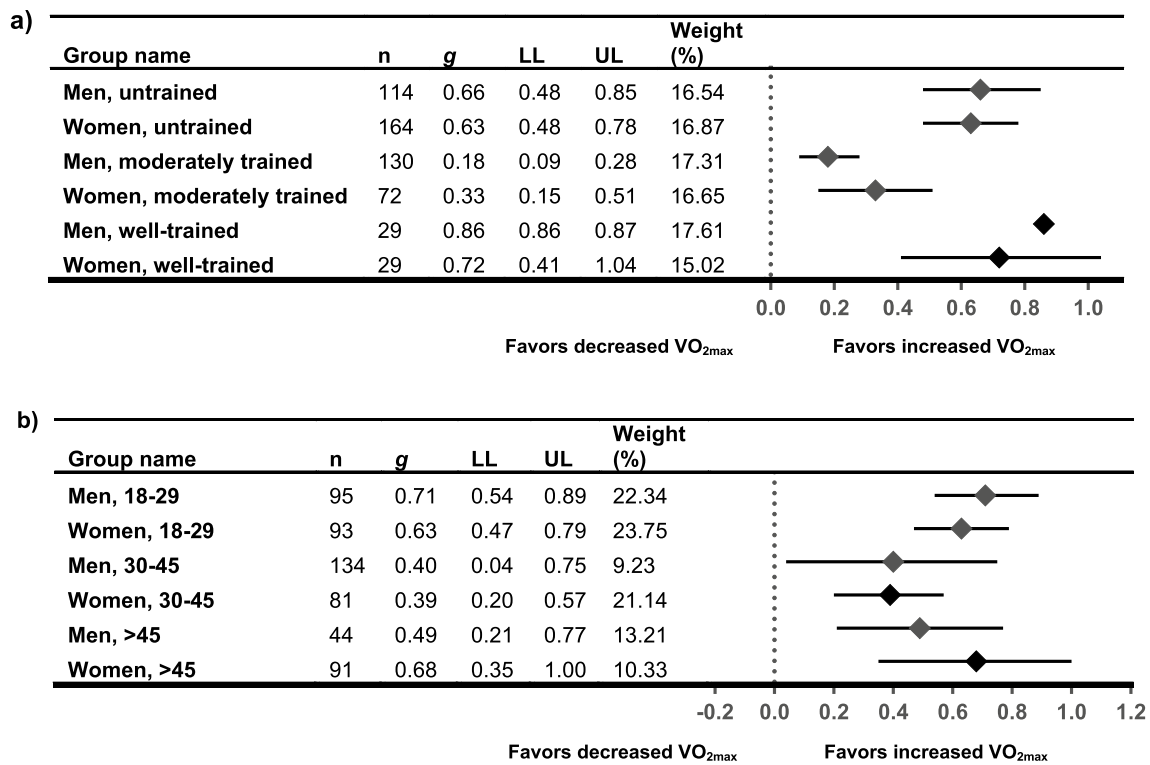
**Fig. 2** Meta-analysis of  $\dot{V}O_{2max}$ , pre- and post-HIIT or SIT intervention. Standardized mean differences and 95% confidence intervals. *HIIT* high-intensity interval training, *SIT* sprint interval training,  $\dot{V}O_{2max}$  maximal oxygen uptake, *LL* confidence interval lower limit, *UL* confidence interval upper limit



Although the observed asymmetry in these funnel plots may indicate some risk of publication bias, most included studies were laboratory studies where results were based on adherence to strict protocols (see also the fidelity check in Table 2). Due to the dose–response effects of exercise load and cardiorespiratory outcomes, it is unlikely that many studies with high adherence would have produced negative overall effect sizes. Similarly, the objectives of the review necessitated this level of adherence since the aim of the review was to explore and quantify the impact of biological sex on observed physiological adaptations. As such, the presence of some asymmetry within the funnel plots may

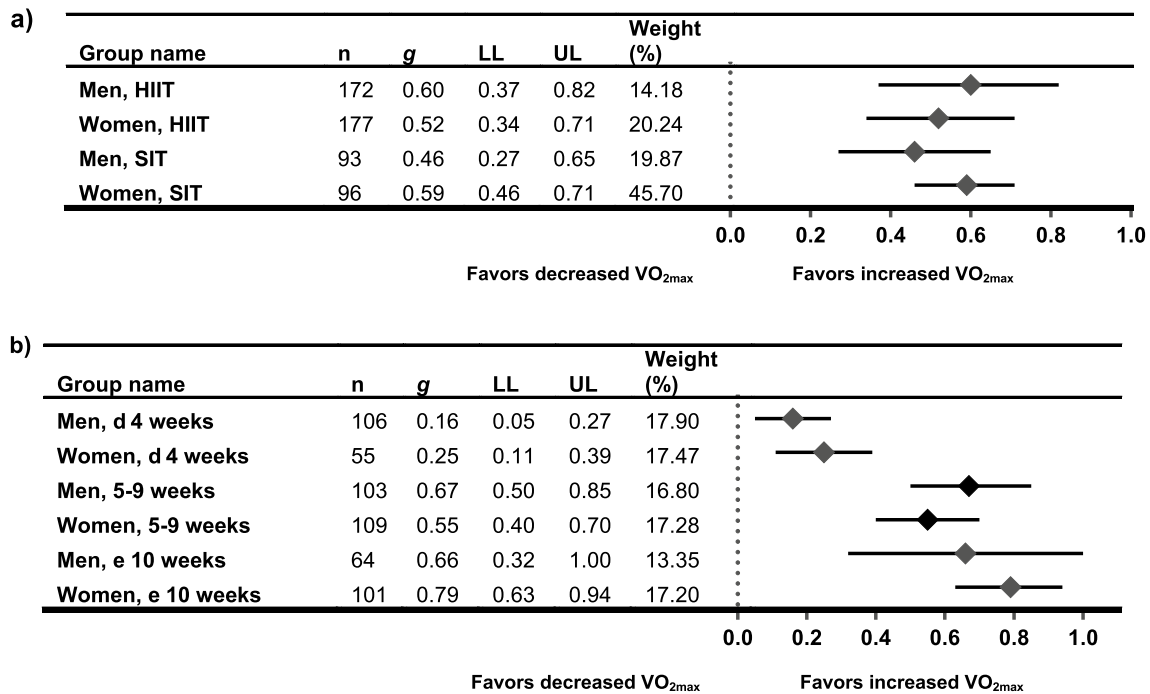
not indicate excessive publication bias within the current objectives of this review.

A visual inspection of the symmetry of the datapoint distribution for  $power_{AT}$  shows no asymmetry as an indication of notable publication bias in the pre-post effect meta-analysis. Observed and adjusted effect sizes were equivalent at 0.37 (95% CI 0.22–0.53); however, it should be noted that this meta-analysis only included data from five studies. Additionally, some reporting bias for all meta-analyzed outcomes was noted with some studies reporting a lack of sex differences in outcomes and thereby pooling male and female data. Although efforts were taken to obtain data from



**Fig. 3** Sub-grouping of meta-analysis of  $\dot{V}O_{2max}$ , pre- and post-HIIT or SIT intervention, **a** by baseline training status, and **b** by mean group age. Standardized mean differences and 95% confidence inter-

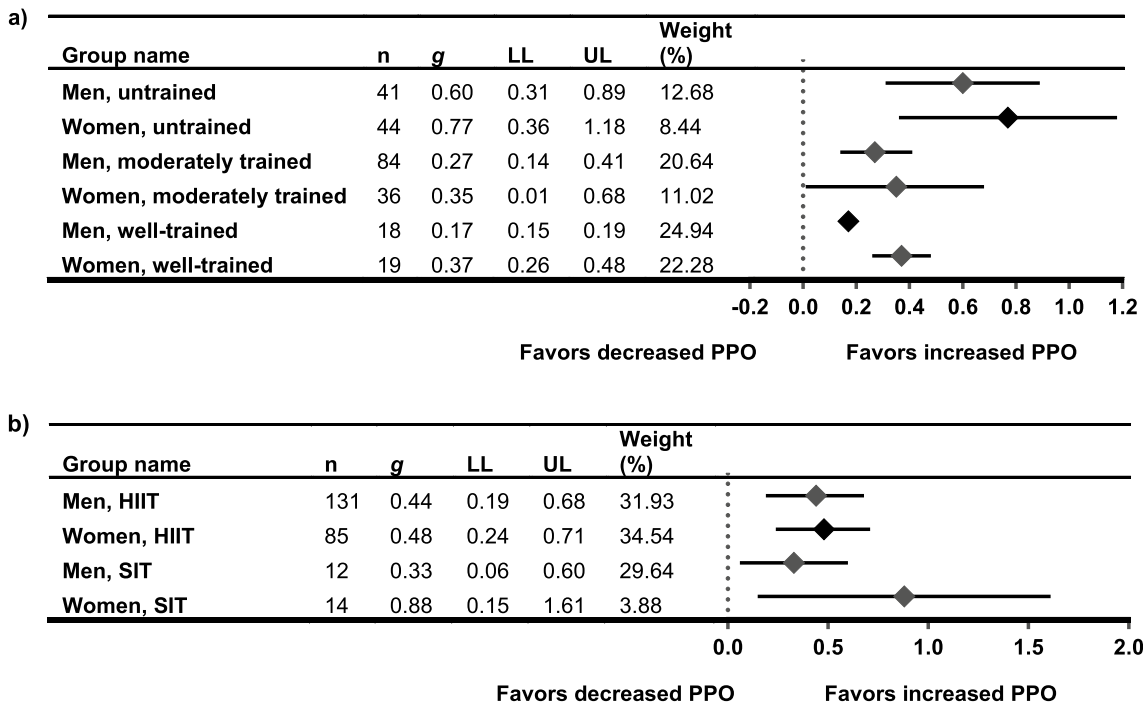
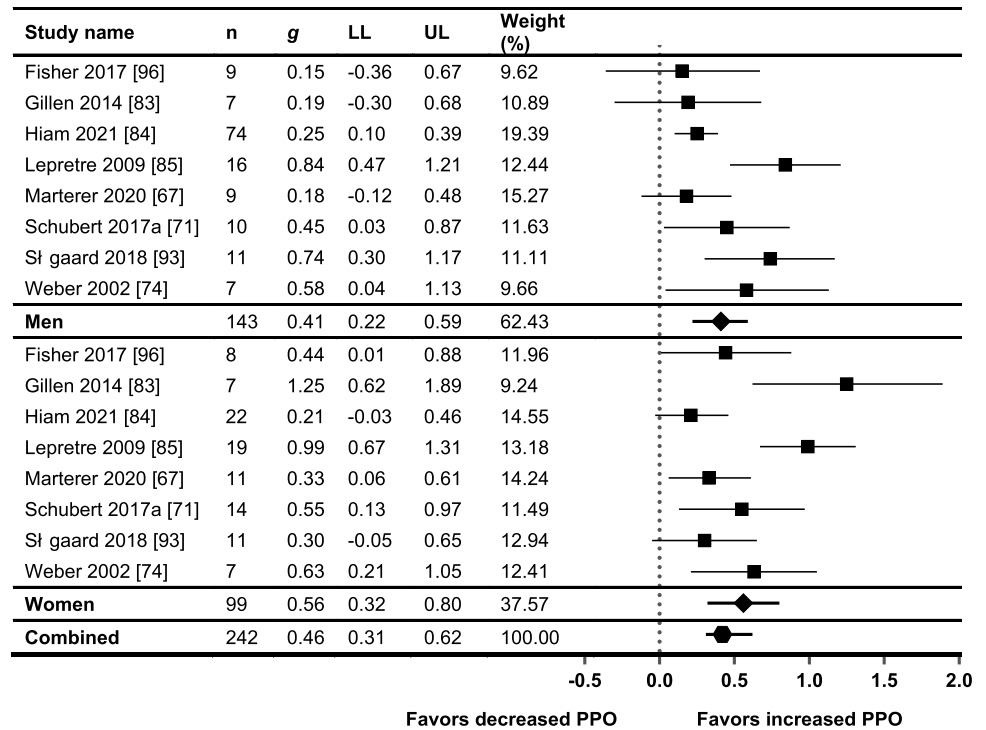
vals. *HIIT* high-intensity interval training, *SIT* sprint interval training,  $\dot{V}O_{2max}$  maximal oxygen uptake, *LL* confidence interval lower limit, *UL* confidence interval upper limit



**Fig. 4** Sub-grouping of meta-analysis of  $\dot{V}O_{2max}$ , pre- and post-HIIT or SIT intervention, **a** by intervention type, and **b** by intervention length. Standardized mean differences and 95% confidence intervals.

*HIIT* high-intensity interval training, *SIT* sprint interval training,  $\dot{V}O_{2max}$  maximal oxygen uptake, *LL* confidence interval lower limit, *UL* confidence interval upper limit

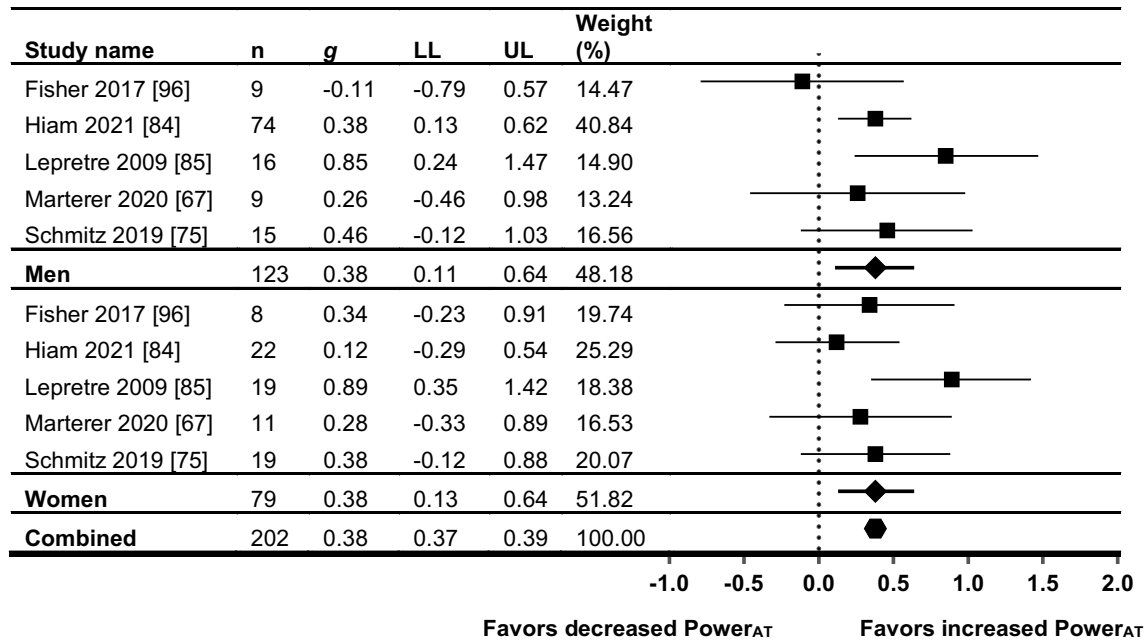
**Fig. 5** Meta-analysis of peak power out from incremental exercise testing (PPO), pre- and post-HIIT or SIT intervention. Standardized mean differences and 95% confidence intervals. *HIIT* high-intensity interval training, *SIT* sprint interval training, *LL* confidence interval lower limit, *UL* confidence interval upper limit



**Fig. 6** Sub-grouping of meta-analysis of peak power out from incremental exercise testing (PPO), pre- and post-HIIT or SIT intervention, **a** by baseline training status, and **b** by intervention type.

Standardized mean differences and 95% confidence intervals. *HIIT* high-intensity interval training, *SIT* sprint interval training, *LL* confidence interval lower limit, *UL* confidence interval upper limit





**Fig. 7** Meta-analysis of threshold power ( $\text{Power}_{\text{AT}}$ ), pre- and post-HIIT or SIT intervention. Standardized mean differences and 95% confidence intervals. *HIIT* high-intensity interval training, *SIT* sprint interval training, *LL* confidence interval lower limit, *UL* confidence interval upper limit

authors where possible, this type of reporting has resulted in some missing data from the meta-analyses. It is, however, unlikely that this missing data would have significantly affected the results for  $\dot{V}\text{O}_{2\text{max}}$  and  $\text{power}_{\text{AT}}$ , since these strongly indicated no between-group differences in the magnitude of adaptation. The meta-analysis of PPO could have benefitted from additional data which may have influenced the final results.

## 4 Discussion

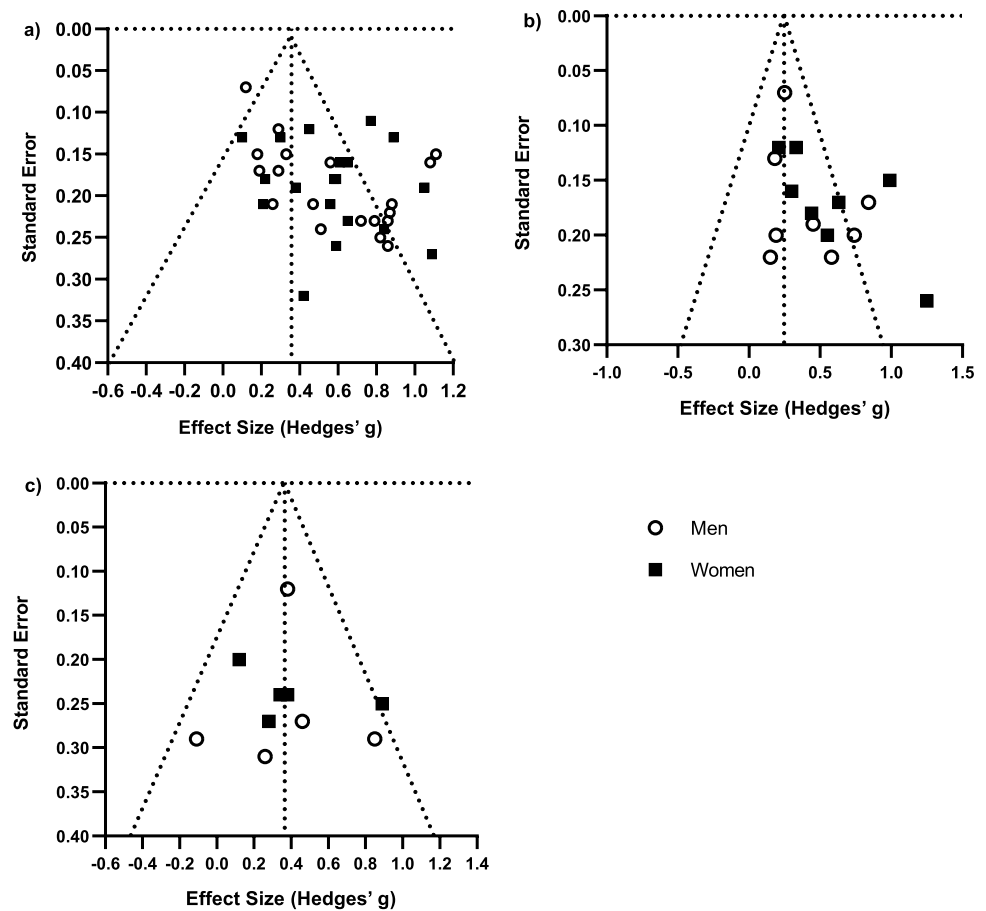
The main finding of the current review with meta-analyses is that men and women improve fitness and performance outcomes to a similar extent following equivalent HIIT and SIT interventions. In particular, meta-analyzed outcomes for  $\dot{V}\text{O}_{2\text{max}}$  and  $\text{power}_{\text{AT}}$  revealed strikingly similar small to moderate increases for men and women. These findings are consistent with those of Weston and colleagues [50], who found moderate improvements in  $\dot{V}\text{O}_{2\text{max}}$  in active and sedentary adults in response to HIIT. Until recently, there has been insufficient research to conclude whether or not sex differences exist in fitness adaptations to HIIT interventions. Our findings expand on the work of Weston and colleagues [50], who could not come to a conclusion regarding a sex-specific response to HIIT, by demonstrating near-identical effect sizes in women and men and a

lack of significant between-group differences for the primary meta-analysis of any outcome.

The sensitivity analysis that was undertaken to estimate raw mean differences found that baseline  $\dot{V}\text{O}_{2\text{max}}$  in men was significantly higher, the equivalent of  $1.06 \text{ L}\cdot\text{min}^{-1}$  or  $5.88 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ , compared with women, as expected. The pre-post analysis using a raw mean difference indicated significant overall improvements of approximately  $0.23 \text{ L}\cdot\text{min}^{-1}$  and  $3.40 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ , without the presence of significant between-group differences for men and women. This analysis indicates that the general results were not altered by the use of a standardized mean difference designed for dependent data and provides more practical estimates of effect. The results of this analysis should be taken with some caution, however, since the meta-analytical approach assumes independent data. In particular, the low levels of heterogeneity in the pre-post analyses appear to be underestimated.

While larger effect sizes for PPO were demonstrated for women compared with men across all sub-groupings, this difference only reached the threshold of significance in participants who were well trained at baseline. Despite the presence of this apparent difference, it must be noted that the well-trained male and female sub-groups only consisted of two studies and had a small sample size of only 18 men and 19 women. Baseline training status accounted for all significant heterogeneity in  $\dot{V}\text{O}_{2\text{max}}$  and PPO outcomes in

**Fig. 8** Funnel plots of the pre-post effect size (standardized mean differences; Hedges'  $g$ ) versus standard error for **a** maximal oxygen uptake ( $\dot{V}O_{2\max}$ ), **b** peak power output from incremental exercise testing (PPO), and **c** threshold power ( $Power_{AT}$ ) in response to HIIT or SIT interventions. Funnels represent the 95% prediction interval for pooled, observed effect size. *HIIT* high-intensity interval training, *SIT* sprint interval training



moderately trained and well-trained men and women. While decreases in the variability of outcomes for participants who were trained at baseline could make sense from a physiological standpoint as fitness and performance outcomes move closer to a theoretical physiological ceiling, it is unclear from the current analysis what factors contributed to the variability of outcomes for untrained participants.

A significant between-group difference was present overall for  $\dot{V}O_{2\max}$  when sub-grouped by baseline training status, with smaller effect sizes present for the moderately trained groups. This likely reflected some confounding from shorter intervention lengths in the moderately trained groups. All except one study within the moderately trained groups had an intervention length ranging between 2 and 4 weeks, while studies in the well-trained category had a range of 6–8 weeks. Consistent with this, the sub-grouping by intervention length revealed significant between-group differences with smaller effect sizes seen for interventions of  $\leq 4$  weeks in duration for both sexes. While the female data demonstrated a gradual increase in effect size with longer interventions, the male data demonstrated similar effect sizes but greater variability with all sub-groups longer than 4 weeks in duration. Interestingly, these findings are consistent with the study by Hirsch et al. [72],

who highlighted potential differences in the rate of adaptation between men and women, with significant changes in  $\dot{V}O_{2\max}$  occurring during the first 4 weeks for the men in their study, while the significant changes in women occurred during the second 4 weeks.

Conversely, the sub-group analyses for mean group age and intervention type (HIIT versus SIT) did not significantly influence the heterogeneity in  $\dot{V}O_{2\max}$  or PPO outcomes and revealed no significant between-group differences. Furthermore, the observed sensitivity of the mean age sub-group analysis indicated fundamental issues with the robustness of this grouping, and as such these results should be interpreted with caution. Overall, it appears that baseline training status and intervention length may be important factors influencing the variability of outcomes in response to HIIT and SIT interventions for both sexes, rather than age or intervention type.

Except for conflicting evidence regarding fatigue and speed/power decrement, the qualitative analysis of results of the additional performance outcomes (outlined in Table 7) demonstrated a similar lack of differences between men and women regarding the magnitude of change for most performance outcomes. Generally, the included studies supported current knowledge that women are less fatigable than men

[66, 73, 76, 91, 95, 96] prior to HIIT. While some of the studies included in this review indicated that men may be able to adapt this to a greater extent than women through HIIT [91, 93], some studies found significant improvements in fatigability or power output through repeated efforts in women only [95, 97], while still other studies found no sex differences at all regarding the change in fatigability [66, 73, 76, 96].

Despite the notable lack of sex differences in the magnitude of adaptation of fitness and performance outcomes in response to HIIT, some of the findings outlined in Tables 7 and 8 indicated potential differences regarding the underlying mechanisms contributing to these improvements. Most notably, all significant sex  $\times$  HIIT interactions reported in the primary studies that related to central cardiorespiratory adaptations favored men, with women generally demonstrating smaller, and often non-significant changes. Examples of this included the observed increases in peak cardiac output, peak stroke volume, and peak cardiac index in men only after SIT, as demonstrated by Bostad et al. [80]; the significant sex  $\times$  HIIT interaction for oxygen pulse (the amount of oxygen ejected from the ventricles with each cardiac contraction) with a greater increase in men compared with women as demonstrated by Marterer and colleagues [67]; the increases in accumulated oxygen uptake and decreases in accumulated oxygen deficit after HIIT seen only in men, as demonstrated by Weber and Schneider [74]; and finally, the increases in minute ventilation and decreases in maximum heart rate and the ventilatory equivalents for oxygen and carbon dioxide in men only, as demonstrated by Mucci and colleagues [70]. These findings, together with the increase in coupled and uncoupled mitochondrial respiratory capacity that was demonstrated in men only, as outlined by Chrøis et al. [81], indicate that improvements in oxygen delivery and uptake in men may play a greater role in achieving fitness and performance adaptations after HIIT compared with women. While adaptations to high-intensity exercise via improvements in oxygen delivery and uptake may seem counter-intuitive on initial contemplation, previous studies have indicated that the greater oxygen availability during exercise in women provides an advantage with regard to the fatigability of muscular contractions, even at maximal intensities [43, 99]. Consistent with this, it appears that the sex differences in fatigability can be eliminated under ischemic conditions [100]. As such, it is conceivable that such adaptations may contribute to improvements in maximal fitness and performance outcomes. Similar findings have also been reported in a previous review of responses to endurance training, where increases in left ventricular end-diastolic volume and stroke volume increased to a greater extent in men compared with women [101]. Although this meta-analysis reported outcomes using mean differences, thereby reporting absolute changes in these outcomes, the

results reported in the primary studies included in the current review indicate that these differences may also exist in these changes when considered relative to baseline.

In contrast, there was a relative lack of evidence surrounding the mechanisms of adaptation that account for the equivalent improvements in fitness and performance outcomes in women. Interestingly, the study by Hoffmann and colleagues [73] reported a significant correlation where a change in threshold power ( $LT_2$ ) accounted for 77% of the performance improvement in the 40 km cycling time trial for women. This same result was not observed for men, nor were there any significant correlations between change in time trial performance and measures of heart rate or blood lactate at  $LT_2$ , absolute or relative peak power output, or incremental time to fatigue. Another study in the current review [97] found a significant increase in mean power during repeated Wingate tests and greater increases in the cross-sectional area of type IIB muscle fibers in women only in response to four weeks of SIT. Although these are the findings of only two studies, which could have been influenced by exercise mode since both used cycle-based testing and training protocols, some additional information can be gathered from outcomes that have been presented in other contexts. Spina and colleagues [102] compared mechanisms of adaptation in older men and women who participated in 9–12 months of moderate to vigorous uphill walking and running-based endurance training and found that 66% of the  $\dot{V}O_{2max}$  adaptation in older men was accounted for by a 15% increase in stroke volume in combination with a 7% increase in arteriovenous oxygen difference at maximal exercise. In contrast, this study found no change in stroke volume in older women and the whole change in  $\dot{V}O_{2max}$  could only be contributed to an increase in peripheral oxygen extraction. In another example, similar to the relationship noted by Hoffmann and colleagues [73], one cross-sectional study [103] found that 60% of the variability in 10-km performance for highly trained female runners aged 23–47 years was explained by running velocity at the lactate threshold. The authors noted, however, that this relationship was age-dependent, with  $\dot{V}O_{2max}$  explaining 74% of the variability in performance in women aged 37–56 years. The potential differences highlighted here alongside the strikingly similar effect sizes seen for  $\dot{V}O_{2max}$  in the current meta-analysis suggest some potential sex differences in the adaptive responses to HIIT/SIT that may warrant further investigation.

Despite highlighting potential differences in the adaptive responses to HIIT and SIT in men and women, these appear to be only different means to the same end. Improvements in maximal cardiorespiratory fitness along with many other performance measures outlined in the results of the current review were found to occur to a similar magnitude in men and women. In contrast to this, a previous meta-analysis by Diaz-Canestro and Montero [52] reported greater

increases in absolute and relative  $\dot{V}O_{2\max}$  in men compared with women in response to moderate-intensity endurance training. This discrepancy between the results of these two meta-analyses may be influenced by the differing interventions (endurance training vs HIIT/SIT) or the methodological differences between the two meta-analyses. The earlier review by Diaz-Canestro and Montero [52] presented absolute change in  $\dot{V}O_{2\max}$  ( $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$  and  $\text{mL}\cdot\text{min}^{-1}$ ) using raw mean differences, whereas the current review primarily focused on changes relative to baseline reflected as standardized effect sizes, percentage change, and sex  $\times$  HIIT interactions (see also the comment by Senefeld and colleagues [104]). Despite these differences in overall approach, the sensitivity analysis using a raw mean difference in the current review persisted to indicate that there were no significant differences between men and women for either absolute or relative  $\dot{V}O_{2\max}$  in response to HIIT/SIT interventions. Overall, these findings seem to indicate that sex differences in  $\dot{V}O_{2\max}$  response may be protocol dependent, and could warrant further investigation.

## 5 Limitations

The current review has a number of limitations. Firstly, in order to minimize confounding from different exercise protocols (exercise dose) across studies, only studies that presented both male and female data were included in the current review. While this ensures that exercise dose is normalized between male and female sub-groups and provides relatively similar numbers of men and women within each analysis, the majority of the included studies were small, which limits the number of participants within the meta-analyses overall. Since studies with small sample sizes tend to be associated with larger effect sizes and greater error, some of the effect sizes seen here have the potential to be overestimated or unduly influenced by a small number of participants. Despite this, the focus of the current review was to identify differences between the relative change in outcomes for men and women rather than quantifying effect sizes. The general effects of HIIT, particularly with relation to  $\dot{V}O_{2\max}$ , have been demonstrated for mixed male and female groups in previous meta-analyses [4, 9, 50], and the current analysis strongly indicates that sex differences in the relative change in these outcomes are minimal.

Another potential limitation of the current meta-analysis could be that while prescribed training doses were matched between men and women in all studies, there was an inability to properly assess the actual dose received in many studies and whether this was consistent with the prescribed dose (dose delivered). Many studies provided only basic details regarding compliance to prescribed exercise, with only a few reporting that this remained consistent between men and

women. Despite this, many of the included studies appeared to have been tightly controlled interventions, in which case the differences between the prescribed and actual training doses are unlikely to have been substantially different.

In addition to the small pooled sample size and challenges associated with assessing the fidelity of the interventions, the use of pre-post meta-analysis techniques and multiple effect sizes from individual studies in the same analysis (such as the matched male and female sub-group data) have been widely used, but also widely debated in the literature [105–107]. While the use of dependent pre-post data has been largely accounted for with the use of correlation coefficients in the meta-analyses, sufficient data was not available to calculate correlation coefficients for all studies, therefore many of these were imputed. Furthermore, while the design of the current review ensures that prescribed exercise dose is matched between male and female groups, the wide range of intervention lengths and exercise protocols included in the literature makes it difficult to precisely examine the influence of different protocols on these outcomes. Overall, the approaches used within the current review with meta-analyses aimed to minimize statistical and methodological errors as much as possible in the face of the unique set of challenges associated with the research question; however, the results should be considered within the constraints of the limitations that are outlined here.

Finally, while the current review provides some insight into generalized similarities and differences between men and women regarding physiological adaptations to HIIT interventions, evaluation of the influence of hormonal status on outcomes was outside of the scope of this review. While the influence of hormonal fluctuations in women may be somewhat offset by the current focus on adaptation over several sessions, generally spanning weeks or months, these concepts appeared to be largely overlooked in the included studies and should be a focus of future primary and secondary research. Despite the limitations outlined here, the findings from the current review will be critical in order to fill the research gaps and to promote better optimization of exercise prescription and health for both women and men.

## 6 Conclusions

The current review with meta-analyses aimed to clarify sex differences in the adaptations of fitness and performance outcomes in response to HIIT and SIT interventions. The main findings of this review indicated that the magnitude of change in  $\dot{V}O_{2\max}$  and  $\text{power}_{\text{AT}}$  in response to HIIT interventions is similar for men and women. While a borderline significant sex difference was found for PPO in well-trained men and women, the sub-groups consisted of small sample sizes and therefore should be interpreted with caution.

Additionally, qualitative analysis of performance outcomes and concurrent measures of physiological adaptation indicated potential differences in the underlying mechanisms of adaptation for men and women. Lastly, it appears that baseline training status and intervention length may play a role in influencing the variability of  $\dot{V}O_{2\max}$  and PPO outcomes in both sexes, including significantly smaller effect sizes for interventions with a duration of 4 weeks or less.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s40279-023-01914-0>.

**Acknowledgements** The authors would like to acknowledge the Hamad Bin Khalifa University (HBKU) librarians, Samy and Mohamed for their assistance with the planning of the search strategy and obtaining articles via inter-library loans. Thank you to the authors who responded and supplied data to contribute to the review and meta-analysis, and all of the authors of the included articles for their valuable work that enabled us to undertake this project. The authors would like to acknowledge HBKU, the College of Health and Life Sciences (CHLS), and the Division of Exercise Science, Health and Epidemiology (ESHE) for their logistical support through the post-doctoral and post-graduate scholarship funding for the first and second authors, and the additional funding for the Covidence license to assist with the management of this review. Lastly, we would like to acknowledge the Qatar National Library for funding the open access publication of this manuscript.

**Funding** Open Access funding provided by the Qatar National Library.

## Declarations

**Conflict of interest** The authors declare that they have no financial or non-financial conflicts of interest. Initial results from this study have contributed to the thesis of the second author for the degree of Master of Science in Exercise Science at Hamad Bin Khalifa University (HBKU). Initial findings have also been presented as a conference poster for Research to Practice 2022 (online May 2022, Exercise and Sports Science Australia).

**Funding** Open access funding was provided by the Qatar National Library. Funding was received from HBKU, College of Health and Life Sciences for the Covidence license that was used to manage this review. A post-graduate research scholarship was granted to the second author from HBKU during the course of the Master's degree.

**Data availability** The extracted data that was used for the current manuscript is available in Supplementary Online Resource 1 in the ESM.

**Author contributions** ML and IY undertook the processes of searching, screening, and risk of bias appraisal. Initial analysis of  $\dot{V}O_{2\max}$  outcomes and write-up were undertaken by IY as a part of her Master's thesis. Subsequent analyses and results synthesis were undertaken by ML, and subsequent write-ups were undertaken collaboratively by IY and ML. Physiological content expertise and guidance was provided by NT and BM, and review/meta-analytical content expertise and guidance was provided by HM and ML. All authors significantly contributed to the design and development of the project at all stages. All authors contributed to all versions of the manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source,

provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. American College of Sports Medicine, Riebe D, Ehrman JK, Liguori G, Magal M. ACSM's guidelines for exercise testing and prescription. 2018.
2. Gibala MJ, Gillen JB, Percival ME. Physiological and health-related adaptations to low-volume interval training: influences of nutrition and sex. *Sports Med.* 2014;44:127–37. <https://doi.org/10.1007/s40279-014-0259-6>.
3. Kubukeli ZN, Noakes TD, Dennis SC. Training techniques to improve endurance exercise performances. *Sports Med Auckl NZ.* 2002;32:489–509. <https://doi.org/10.2165/00007256-200232080-00002>.
4. Milanović Z, Sporiš G, Weston M. Effectiveness of high-intensity interval training (HIT) and continuous endurance training for  $\dot{V}O_{2\max}$  improvements: a systematic review and meta-analysis of controlled trials. *Sports Med.* 2015;45:1469–81. <https://doi.org/10.1007/s40279-015-0365-0>.
5. Wen D, Utesch T, Wu J, Robertson S, Liu J, Hu G, et al. Effects of different protocols of high intensity interval training for  $\dot{V}O_{2\max}$  improvements in adults: a meta-analysis of randomised controlled trials. *J Sci Med Sport.* 2019;22:941–7. <https://doi.org/10.1016/j.jsams.2019.01.013>.
6. Weston KS, Wisløff U, Coombes JS. High-intensity interval training in patients with lifestyle-induced cardiometabolic disease: a systematic review and meta-analysis. *Br J Sports Med.* 2014;48:1227–34. <https://doi.org/10.1136/bjsports-2013-092576>.
7. Ramos JS, Dalleck LC, Tjonna AE, Beetham KS, Coombes JS. The impact of high-intensity interval training versus moderate-intensity continuous training on vascular function: a systematic review and meta-analysis. *Sports Med.* 2015;45:679–92. <https://doi.org/10.1007/s40279-015-0321-z>.
8. Astorino T, Schubert M, Palumbo E, Stirling D, McMillan D, Cooper C, et al. Magnitude and time course of changes in maximal oxygen uptake in response to distinct regimens of chronic interval training in sedentary women. *Eur J Appl Physiol.* 2013. <https://doi.org/10.1007/s00421-013-2672-1>.
9. Bacon AP, Carter RE, Ogle EA, Joyner MJ.  $\dot{V}O_{2\max}$  trainability and high intensity interval training in humans: a meta-analysis. *PLoS ONE.* 2013;8: e73182. <https://doi.org/10.1371/journal.pone.0073182>.
10. Dun Y, Smith JR, Liu S, Olson TP. High-intensity interval training in cardiac rehabilitation. *Clin Geriatr Med.* 2019;35:469–87. <https://doi.org/10.1016/j.cger.2019.07.011>.
11. Maturana FM, Martus P, Zipfel S, NIEB AM. Effectiveness of HIIE versus MICT in improving cardiometabolic risk factors in health and disease: a meta-analysis. *Med Sci Sports Exerc.* 2021;53:559–73. <https://doi.org/10.1249/MSS.0000000000002506>.
12. Palma S, Hasenoehrl T, Jordakieva G, Ramzanova D, Crevenna R. High-intensity interval training in the prehabilitation of cancer patients—a systematic review and meta-analysis. *Support Care Cancer.* 2021;29:1781–94. <https://doi.org/10.1007/s00520-020-05834-x>.

13. Serrablo-Torres I, Lopez-Valenciano A, Ayuso M, Horton E, Mayo X, Medina-Gomez G, et al. High intensity interval training exercise-induced physiological changes and their potential influence on metabolic syndrome clinical biomarkers: a meta-analysis. *BMC Endocr Disord.* 2020;20:167. <https://doi.org/10.1186/s12902-020-00640-2>.
14. Leahy AA, Mavilidi MF, Smith JJ, Hillman CH, Eather N, Barker D, et al. Review of high-intensity interval training for cognitive and mental health in youth. *Med Sci Sports Exerc.* 2020;52:2224–34. <https://doi.org/10.1249/MSS.0000000000002359>.
15. Martland R, Mondelli V, Gaughran F, Stubbs B. Can high-intensity interval training improve physical and mental health outcomes? A meta-review of 33 systematic reviews across the lifespan. *J Sports Sci.* 2020;38:430–69. <https://doi.org/10.1080/02640414.2019.1706829>.
16. Martland R, Mondelli V, Gaughran F, Stubbs B. Can high intensity interval training improve health outcomes among people with mental illness? A systematic review and preliminary meta-analysis of intervention studies across a range of mental illnesses. *J Affect Disord.* 2020;263:629–60. <https://doi.org/10.1016/j.jad.2019.11.039>.
17. Moreau D, Chou E. The acute effect of high-intensity exercise on executive function: a meta-analysis. *Perspect Psychol Sci J Assoc Psychol Sci.* 2019;14:734–64. <https://doi.org/10.1177/1745691619850568>.
18. Reljic D, Lampe D, Wolf F, Zopf Y, Herrmann HJ, Fischer J. Prevalence and predictors of dropout from high-intensity interval training in sedentary individuals: a meta-analysis. *Scand J Med Sci Sports.* 2019;29:1288–304. <https://doi.org/10.1111/sms.13452>.
19. Stubbs B, Vancampfort D, Hallgren M, Firth J, Veronese N, Solmi M, et al. EPA guidance on physical activity as a treatment for severe mental illness: a meta-review of the evidence and Position Statement from the European Psychiatric Association (EPA), supported by the International Organization of Physical Therapists in Mental Health (IOPTMH). *Eur Psychiatry J Assoc Eur Psychiatr.* 2018;54:124–44. <https://doi.org/10.1016/j.eurpsy.2018.07.004>.
20. Gibala MJ, Little JP, Macdonald MJ, Hawley JA. Physiological adaptations to low-volume, high-intensity interval training in health and disease. *J Physiol.* 2012;590:1077–84. <https://doi.org/10.1113/jphysiol.2011.224725>.
21. Gibala MJ, Little JP, van Essen M, Wilkin GP, Burgomaster KA, Safdar A, et al. Short-term sprint interval versus traditional endurance training: similar initial adaptations in human skeletal muscle and exercise performance. *J Physiol.* 2006;575:901–11. <https://doi.org/10.1113/jphysiol.2006.112094>.
22. Santisteban KJ, Lovering AT, Halliwill JR, Minson CT. Sex differences in  $VO_{2max}$  and the impact on endurance-exercise performance. *Int J Environ Res Public Health.* 2022;19:4946. <https://doi.org/10.3390/ijerph19094946>.
23. Ansdell P, Thomas K, Hicks KM, Hunter SK, Howatson G, Goodall S. Physiological sex differences affect the integrative response to exercise: acute and chronic implications. *Exp Physiol.* 2020;105:2007–21. <https://doi.org/10.1113/EP088548>.
24. Altavilla G, Tore PAD, Riela L, D'Isanto T. Anthropometric, physiological and performance aspects that differentiate male athletes from females and practical consequences. *J Phys Educ Sport.* 2017. <https://doi.org/10.7752/jpes.2017.s5226>.
25. Sparling PB. A meta-analysis of studies comparing maximal oxygen uptake in men and women. *Res Q Exerc Sport.* 1980;51:542–52. <https://doi.org/10.1080/02701367.1980.10608077>.
26. Issurin VB, Kaufman LE, Tenenbaum G. Modeling of velocity regimens for anaerobic and aerobic power exercises in high-performance swimmers. *J Sports Med Phys Fit.* 2001;41:433–40.
27. Chevront SN, Carter R, Deruisseau KC, Moffatt RJ. Running performance differences between men and women: an update. *Sports Med Auckland NZ.* 2005;35:1017–24. <https://doi.org/10.2165/00007256-200535120-00002>.
28. Drinkwater BL. Physiological responses of women to exercise. *Exerc Sport Sci Rev.* 1973;1:125–53.
29. Weber CL, Chia M, Inbar O. Gender differences in anaerobic power of the arms and legs—a scaling issue. *Med Sci Sports Exerc.* 2006;38:129–37. <https://doi.org/10.1249/01.mss.0000179902.31527.2c>.
30. Huxley VH. Sex and the cardiovascular system: the intriguing tale of how women and men regulate cardiovascular function differently. *Adv Physiol Educ.* 2007;31:17–22. <https://doi.org/10.1152/advan.00099.2006>.
31. Pelliccia A, Maron BJ, Spataro A, Proschan MA, Spirito P. The upper limit of physiologic cardiac hypertrophy in highly trained elite athletes. *N Engl J Med.* 1991;324:295–301. <https://doi.org/10.1056/NEJM199101313240504>.
32. Vasan RS, Larson MG, Levy D, Evans JC, Benjamin EJ. Distribution and categorization of echocardiographic measurements in relation to reference limits. *Circulation.* 1997;96:1863–73. <https://doi.org/10.1161/01.CIR.96.6.1863>.
33. Whyte GP, George K, Sharma S, Firooz S, Stephens N, Senior R, et al. The upper limit of physiological cardiac hypertrophy in elite male and female athletes: the British experience. *Eur J Appl Physiol.* 2004;92:592–7. <https://doi.org/10.1007/s00421-004-1052-2>.
34. Ceridon ML, Beck KC, Olson TP, Bilezikian JA, Johnson BD. Calculating alveolar capillary conductance and pulmonary capillary blood volume: comparing the multiple- and single-inspired oxygen tension methods. *J Appl Physiol.* 2010;109:643–53. <https://doi.org/10.1152/jappphysiol.01411.2009>.
35. Dominelli PB, Ripoll JG, Cross TJ, Baker SE, Wiggins CC, Welch BT, et al. Sex differences in large conducting airway anatomy. *J Appl Physiol.* 2018;125:960–5. <https://doi.org/10.1152/jappphysiol.00440.2018>.
36. McClaran SR, Harms CA, Pegelow DF, Dempsey JA. Smaller lungs in women affect exercise hyperpnea. *J Appl Physiol.* 1998;84:1872–81. <https://doi.org/10.1152/jappphysiol.1998.84.6.1872>.
37. Martin TR, Castile RG, Fredberg JJ, Wohl ME, Mead J. Airway size is related to sex but not lung size in normal adults. *J Appl Physiol Bethesda Md* 1985. 1987;63:2042–7. <https://doi.org/10.1152/jappphysiol.1987.63.5.2042>.
38. Sheel AW, Guenette JA. Mechanics of breathing during exercise in men and women: sex versus body size differences? *Exerc Sport Sci Rev.* 2008;36:7. <https://doi.org/10.1097/JES.0b013e31817be7f0>.
39. Thurlbeck WM. Postnatal human lung growth. *Thorax.* 1982;37:564–71. <https://doi.org/10.1136/thx.37.8.564>.
40. Guenette JA, Witt JD, McKenzie DC, Road JD, Sheel AW. Respiratory mechanics during exercise in endurance-trained men and women. *J Physiol.* 2007;581:1309–22. <https://doi.org/10.1113/jphysiol.2006.126466>.
41. Wilkie SS, Dominelli PB, Sporer BC, Koehle MS, Sheel AW. Heliox breathing equally influences respiratory mechanics and cycling performance in trained males and females. *J Appl Physiol.* 2015;118:255–64. <https://doi.org/10.1152/jappphysiol.00400.2014>.
42. Staron RS, Hagerman FC, Hikida RS, Murray TF, Hostler DP, Crill MT, et al. Fiber type composition of the vastus lateralis muscle of young men and women. *J Histochem Cytochem.* 2000;48:623–9. <https://doi.org/10.1177/002215540004800506>.

43. Ansdell P, Brownstein CG, Škarabot J, Hicks KM, Howatson G, Thomas K, et al. Sex differences in fatigability and recovery relative to the intensity–duration relationship. *J Physiol*. 2019;597:5577–95. <https://doi.org/10.1113/JP278699>.
44. Tarnopolsky MA. Sex differences in exercise metabolism and the role of 17-beta estradiol. *Med Sci Sports Exerc*. 2008;40:648–54. <https://doi.org/10.1249/mss.0b013e31816212ff>.
45. Esbjörnsson M, Sylvén C, Holm I, Jansson E. Fast twitch fibres may predict anaerobic performance in both females and males. *Int J Sports Med*. 1993;14:257–63. <https://doi.org/10.1055/s-2007-1021174>.
46. Russ DW, Lanza IR, Rothman D, Kent-Braun JA. Sex differences in glycolysis during brief, intense isometric contractions. *Muscle Nerve*. 2005;32:647–55. <https://doi.org/10.1002/mus.20396>.
47. Astorino TA, Sheard AC. Does sex mediate the affective response to high intensity interval exercise? *Physiol Behav*. 2019;204:27–32. <https://doi.org/10.1016/j.physbeh.2019.02.005>.
48. de Poli RAB, Gonzalez JAM, Fonsati N, Zagatto AM. Differences between genders in anaerobic capacity during a supramaximal effort. *Mot Rev Educ Física*. 2019;25: e101910. <https://doi.org/10.1590/s1980-6574201900030018>.
49. Shephard RJ. Exercise and training in women, part I: influence of gender on exercise and training responses. *Can J Appl Physiol*. 2000;25:19–34. <https://doi.org/10.1139/h00-002>.
50. Weston M, Taylor KL, Batterham AM, Hopkins WG. Effects of low-volume high-intensity interval training (HIIT) on fitness in adults: a meta-analysis of controlled and non-controlled trials. *Sports Med Auckl NZ*. 2014;44:1005–17. <https://doi.org/10.1007/s40279-014-0180-z>.
51. Maillard F, Pereira B, Boisseau N. Effect of high-intensity interval training on total, abdominal and visceral fat mass: a meta-analysis. *Sports Med Auckl NZ*. 2018;48:269–88. <https://doi.org/10.1007/s40279-017-0807-y>.
52. Diaz-Canestro C, Montero D. Sex dimorphism of  $VO_{2max}$  trainability: a systematic review and meta-analysis. *Sports Med*. 2019;49:1949–56. <https://doi.org/10.1007/s40279-019-01180-z>.
53. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021. <https://doi.org/10.1136/bmj.n71>.
54. Wells G, Shea B, O'Connell D, Robertson J, Peterson J, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analysis
55. Hartling L, Hamm M, Milne A, Vandermeer B, Santaguida PL, Ansari M, et al. Validity and inter-rater reliability testing of quality assessment instruments. 2012. <http://europepmc.org/abstract/MED/22536612>.
56. Hartling L, Milne A, Hamm MP, Vandermeer B, Ansari M, Tsertsvadze A, et al. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *J Clin Epidemiol*. 2013;66:982–93. <https://doi.org/10.1016/j.jclinepi.2013.03.003>.
57. AHRQ. Agency for Healthcare Research and Quality. Newcastle-Ottawa Quality Assessment Form for Cohort Studies. <https://www.ncbi.nlm.nih.gov/books/NBK115843/bin/appe-fm3.pdf>.
58. Higgins JPT, Green. *Cochrane handbook for systematic reviews of interventions*. <https://handbook-5-1.cochrane.org/>.
59. van Rhee H, Suurmond R, Hak T. User manual for Meta-Essentials: workbooks for meta-analysis. <https://doi.org/10.2139/ssrn.3241355>
60. Suurmond R, van Rhee H, Hak T. Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Res Synth Methods*. 2017;8:537–53. <https://doi.org/10.1002/jrsm.1260>.
61. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc*. 2009;41:3–12. <https://doi.org/10.1249/MSS.0b013e31818cb278>.
62. Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol*. 1992;45:769–73. [https://doi.org/10.1016/0895-4356\(92\)90054-Q](https://doi.org/10.1016/0895-4356(92)90054-Q).
63. Harbord RM, Egger M, Sterne JAC. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med*. 2006;25:3443–57. <https://doi.org/10.1002/sim.2380>.
64. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002. <https://doi.org/10.1136/bmj.d4002>.
65. Bratland-Sanda S, Pedersen FG, Haave MN, Helgerud J, Støren Ø. Large inter-individual differences in responses to a block of high intensity aerobic interval training: a case series in national-level cyclists and triathletes. *Int J Exerc Sci*. 2020;13:480–487. <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc7241617/>.
66. Astorino TA, Allen RP, Roberson DW, Jurancich M (2012) Effect of high-intensity interval training on cardiovascular function,  $\dot{V}O_{2max}$ , and muscular force. *J Strength Cond Res* 1:138–145. <https://doi.org/10.1519/JSC.0b013e318218dd77>.
67. Marterer N, Menz V, Amin S, Faulhaber M. 6-week high-intensity interval training (HIIT) of the lower extremities improves  $VO_{2max}$  of the upper extremities. *Int J Sports Med*. 2020;41:380–90. <https://doi.org/10.1055/a-1073-8016>.
68. Metcalfe RS, Tardif N, Thompson D, Vollaard NBJ. Changes in aerobic capacity and glycaemic control in response to reduced-exertion high-intensity interval training (REHIT) are not different between sedentary men and women. *Appl Physiol Nutr Metab*. 2016;41:1117–23. <https://doi.org/10.1139/apnm-2016-0253>.
69. Molina-Hidalgo C, De-la-O A, Dote-Montero M, Amaro-Gahete FJ, Castillo MJ. Influence of daily beer or ethanol consumption on physical fitness in response to a high-intensity interval training program. The BEER-HIIT study. *J Int Soc Sports Nutr*. 2020;17:1–13. <https://doi.org/10.1186/s12970-020-00356-7>.
70. Mucci P, Blondel N, Claudine F, Nourry C, Berthoin S. Evidence of exercise-induced  $O_2$  arterial desaturation in non-elite sportsmen and sportswomen following high-intensity interval-training. *Int J Sports Med*. 2004;25:6–13. <https://doi.org/10.1055/s-2003-45225>.
71. Schubert MM, Clarke HE, Seay RF, Spain KK. Impact of 4 weeks of interval training on resting metabolic rate, fitness, and health-related outcomes. *Appl Physiol Nutr Metab Physiol Appl Nutr Metab*. 2017;42:1073–81. <https://doi.org/10.1139/apnm-2017-0268>.
72. Hirsch KR, Greenwalt CE, Saylor HE, Gould LM, Harrison CH, Brewer GJ, et al. High-intensity interval training and essential amino acid supplementation: effects on muscle characteristics and whole-body protein turnover. *Physiol Rep*. 2021;9: e14655.
73. Hoffmann S, Skinner TL, van Rosendal SP, Emmerton LM, Jenkins DG. Comparison of training responses and performance adaptations in endurance-trained men and women performing high-intensity interval training. *J Sports Sci*. 2021;39:1010–20. <https://doi.org/10.1080/02640414.2020.1853960>.
74. Weber CL, Schneider DA. Increases in maximal accumulated oxygen deficit after high-intensity interval training are not gender dependent. *J Appl Physiol Bethesda Md*. 1985;2002(92):1795–801. <https://doi.org/10.1152/jappphysiol.00546.2001>.
75. Schmitz B, Thorwesten L, Klose A, Krüger M, Brand S-M. Progressive high-intensity interval training (HIIT) is not superior to unmodified non-progressive HIIT in an uncontrolled setting. *J Sports Med Phys Fit*. 2019;59:2022–9. <https://doi.org/10.23736/S0022-4707.19.09690-7>.

76. Astorino TA, Allen R, Roberson D, Jurancich M, Lewis R, McCarthy K, et al. Adaptations to high-intensity training are independent of gender. *Eur J Appl Physiol.* 2011;111:1279–86. <https://doi.org/10.1007/s00421-010-1741-y>.
77. Bagley L, Slevin M, Bradburn S, Liu D, Murgatroyd C, Morrissey G, et al. Sex differences in the effects of 12 weeks sprint interval training on body fat mass and the rates of fatty acid oxidation and  $\text{VO}_{2\text{max}}$  during exercise. *BMJ Open Sport Exerc Med.* 2016. <https://doi.org/10.1136/bmjsem-2015-000056>.
78. Bagley L, Al-Shanti N, Bradburn S, Baig O, Slevin M, McPhee JS. Sex comparison of knee extensor size, strength and fatigue adaptation to sprint interval training. *J Strength Cond Res.* 2021;35:64–71. <https://doi.org/10.1519/jsc.0000000000002496>.
79. Bornath DPD, Kenno KA. Physiological responses to increasing battling rope weight during two 3-week high-intensity interval training programs. *J Strength Cond Res.* 2022;36:352–8. <https://doi.org/10.1519/JSC.0000000000003470>.
80. Bostad W, Valentino SE, McCarthy DG, Richards DL, MacInnis MJ, MacDonald MJ, et al. Twelve weeks of sprint interval training increases peak cardiac output in previously untrained individuals. *Eur J Appl Physiol.* 2021;121:2449–58. <https://doi.org/10.1007/s00421-021-04714-4>.
81. Chrøis KM, Dohlmann TL, Sogaard D, Hansen CV, Dela F, Helge JW, et al. Mitochondrial adaptations to high intensity interval training in older females and males. *Eur J Sport Sci.* 2020;20:135–45. <https://doi.org/10.1080/17461391.2019.1615556>.
82. Dalzell C, Nigam A, Juneau M, Guilbeault V, Latour E, Mauriège P, et al. Intensive lifestyle intervention improves cardio-metabolic and exercise parameters in metabolically healthy obese and metabolically unhealthy obese individuals. *Can J Cardiol.* 2014;30:434–40. <https://doi.org/10.1016/j.cjca.2013.11.033>.
83. Gillen JB, Percival ME, Skelly LE, Martin BJ, Tan RB, Tarnopolsky MA, et al. Three minutes of all-out intermittent exercise per week increases skeletal muscle oxidative capacity and improves cardiometabolic health. *PLoS ONE.* 2014;9: e111489. <https://doi.org/10.1371/journal.pone.0111489>.
84. Hiam D, Landen S, Jacques M, Voisin S, Alvarez-Romero J, Byrnes E, et al. Osteocalcin and its forms respond similarly to exercise in males and females. *Bone.* 2021;144: 115818. <https://doi.org/10.1016/j.bone.2020.115818>.
85. Lepretre P-M, Vogel T, Brechat P-H, Dufour S, Richard R, Kaltenbach G, et al. Impact of short-term aerobic interval training on maximal exercise in sedentary aged subjects. *Int J Clin Pract.* 2009;63:1472–8. <https://doi.org/10.1111/j.1742-1241.2009.02120.x>.
86. Liu H, Leng B, Li Q, Liu Y, Bao D, Cui Y. The effect of eight-week sprint interval training on aerobic performance of elite badminton players. *Int J Environ Res Public Health.* 2021;18:638. <https://doi.org/10.3390/ijerph18020638>.
87. Menz V, Strobl J, Faulhaber M, Gatterer H, Burtcher M. Effect of 3-week high-intensity interval training on  $\text{VO}_{2\text{max}}$ , total haemoglobin mass, plasma and blood volume in well-trained athletes. *Eur J Appl Physiol.* 2015;115:2349–56. <https://doi.org/10.1007/s00421-015-3211-z>.
88. Metcalfe RS, Babraj JA, Fawcner SG, Vollaard NBJ. Towards the minimal amount of exercise for improving metabolic health: beneficial effects of reduced-exertion high-intensity interval training. *Eur J Appl Physiol.* 2012;112:2767–75. <https://doi.org/10.1007/s00421-011-2254-z>.
89. Phillips BE, Kelly BM, Lilja M, Ponce-González JG, Brogan RJ, Morris DL, et al. A Practical and time-efficient high-intensity interval training program modifies cardio-metabolic risk factors in adults with risk factors for type II Diabetes. *Front Endocrinol.* 2017;8:229. <https://doi.org/10.3389/fendo.2017.00229>.
90. Sawashita J, Onitsuka S, Gen-no H, Ishikawa S, Iino F, Tateishi N, et al. Effects of mild calorie restriction and high-intensity interval walking in middle-aged and older overweight Japanese. *Exp Gerontol.* 2009;44:666–75. <https://doi.org/10.1016/j.exger.2009.07.007>.
91. Scalzo RL, Peltonen GL, Binns SE, Shankaran M, Giordano GR, Hartley DA, et al. Greater muscle protein synthesis and mitochondrial biogenesis in males compared with females during sprint interval training. *FASEB J.* 2014;28:2705–14. <https://doi.org/10.1096/fj.13-246595>.
92. Schubert MM, Palumbo E, Seay RF, Spain KK, Clarke HE. Energy compensation after sprint- and high-intensity interval training. *PLoS ONE.* 2017;12: e0189590. <https://doi.org/10.1371/journal.pone.0189590>.
93. Sogaard D, Lund MT, Scheuer CM, Dehlbæk MS, Dideriksen SG, Abildskov CV, et al. High-intensity interval training improves insulin sensitivity in older individuals. *Acta Physiol.* 2018;222:1–1. <https://doi.org/10.1111/apha.13009>.
94. Støren Ø, Sæbø M, Støa EM, Bratland-Sanda S, Helgerud J, Unhjem RJ, et al. The effect of age on the  $\dot{V}\text{O}_{2\text{max}}$  response to high-intensity interval training. *Med Sci Sports Exerc.* 2017;49:78–85. <https://doi.org/10.1249/MSS.0000000000001070>.
95. Schmitz B, Niehues H, Thorwesten L, Klose A, Krüger M, Brand S-M. Sex differences in high-intensity interval training—Are HIIT protocols interchangeable between females and males? *Front Physiol.* 2020;11:38. <https://doi.org/10.3389/fphys.2020.00038>.
96. Fisher S. Sex differences and training adaptations in relation to the lactate threshold and endurance exercise performance. Brisbane: The University of Queensland; 2017. <https://doi.org/10.14264/uq.2018.105>.
97. Esbjörnsson Liljedahl M, Holm I, Sylvén C, Jansson E. Different responses of skeletal muscle following sprint training in men and women. *Eur J Appl Physiol.* 1996;74:375–83. <https://doi.org/10.1007/BF02226935>.
98. Cicioni-Kolsky D, Lorenzen C, Williams M, Kemp J. Endurance and sprint benefits of high-intensity and supramaximal interval training. *Eur J Sport Sci.* 2013;13:304–11. <https://doi.org/10.1080/17461391.2011.606844>.
99. Ansdell P, Škarabot J, Atkinson E, Corden S, Tygart A, Hicks KM, et al. Sex differences in fatigability following exercise normalised to the power-duration relationship. *J Physiol.* 2020;598:5717–37. <https://doi.org/10.1113/JP280031>.
100. Russ DW, Kent-Braun JA. Sex differences in human skeletal muscle fatigue are eliminated under ischemic conditions. *J Appl Physiol Bethesda Md* 1985. 2003;94:2414–22. <https://doi.org/10.1152/jappphysiol.01145.2002>.
101. Diaz-Canestro C, Montero D. The impact of sex on left ventricular cardiac adaptations to endurance training: a systematic review and meta-analysis. *Sports Med Auckl NZ.* 2020;50:1501–13. <https://doi.org/10.1007/s40279-020-01294-9>.
102. Spina RJ, Ogawa T, Kohrt WM, Martin WH 3rd, Holloszy JO, Ehsani AA. Differences in cardiovascular adaptations to endurance exercise training between older men and women. *J Appl Physiol.* 1993. <https://doi.org/10.1152/jappphysiol.1993.75.2.849>.
103. Evans SL, Davy KP, Stevenson ET, Seals DR. Physiological determinants of 10-km performance in highly trained female runners of different ages. *J Appl Physiol Bethesda Md* 1985. 1995;78:1931–41. <https://doi.org/10.1152/jappphysiol.1995.78.5.1931>.
104. Senefeld JW, Smith JR, Wiggins CC, Joyner MJ. Comment on: “Sex dimorphism of  $\text{vo}_{2\text{max}}$  trainability: a systematic review and meta-analysis.” *Sports Med Auckl NZ.* 2020;50:1047–8. <https://doi.org/10.1007/s40279-020-01270-3>.
105. Cuijpers P, Weitz E, Cristea IA, Twisk J. Pre-post effect sizes should be avoided in meta-analyses. *Epidemiol Psychiatr Sci.* 2017;26:364–8. <https://doi.org/10.1017/S2045796016000809>.



106. Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol Methods*. 1996;1:170–7. <https://doi.org/10.1037/1082-989X.1.2.170>.
107. Hansen C, Steinmetz H, Block J. How to conduct a meta-analysis in eight steps: a practical guide. *Manag Rev Q*. 2021. <https://doi.org/10.1007/s11301-021-00247-4>.

## Authors and Affiliations

Merilyn Lock<sup>1</sup>  · Ibtisam Yousef<sup>1,2</sup>  · Bridget McFadden<sup>3,4</sup>  · Hend Mansoor<sup>1,5</sup>  · Nathan Townsend<sup>1</sup> 

✉ Merilyn Lock  
mlock@hbku.edu.qa

<sup>1</sup> Division of Exercise Science, Health and Epidemiology,  
College of Health and Life Sciences, Hamad Bin Khalifa  
University, Doha, Qatar

<sup>2</sup> Patient and Family Education Unit, Nursing Department,  
Hamad Medical Corporation, Doha, Qatar

<sup>3</sup> Department of Exercise Science, Arnold School of Public  
Health, University of South Carolina, Columbia, SC, USA

<sup>4</sup> Department of Family, Nutrition, and Exercise Sciences,  
Queens College, City University of New York, Queens, NY,  
USA

<sup>5</sup> Department of Pharmacy Practice and Science, College  
of Pharmacy, University of Kentucky, Lexington, KY, USA