Which Training Intensity Distribution Intervention will Produce the Greatest Improvements in Maximal Oxygen Uptake and Time-Trial Performance in Endurance Athletes? A Systematic Re…

Article *in* Sports Medicine · January 2025

DOI: 10.1007/s40279-024-02149-3

CITATIONS
0

READS
864

17 authors, including:

Michael A Rosenblat
Sylvan Adams Sports Institute
12 PUBLICATIONS **197** CITATIONS
SEE PROFILE

Jem Arnold
University of British Columbia
14 PUBLICATIONS **52** CITATIONS
SEE PROFILE

Gunnar Treff
Paracelsus Medical University
85 PUBLICATIONS **750** CITATIONS
SEE PROFILE

Jonathan Esteve
European University of Madrid
59 PUBLICATIONS **1,792** CITATIONS
SEE PROFILE

Sports Medicine
https://doi.org/10.1007/s40279-024-02149-3

SYSTEMATIC REVIEW

Which Training Intensity Distribution Intervention will Produce the Greatest Improvements in Maximal Oxygen Uptake and Time-Trial Performance in Endurance Athletes? A Systematic Review and Network Meta-analysis of Individual Participant Data

Michael A. Rosenblat[1,2] · Jennifer A. Watt[3,4,5] · Jem I. Arnold[6] · Gunnar Treff[7] · Øyvind B. Sandbakk[8] · Jonathan Esteve-Lanao[9] · Luca Festa[10] · Luca Filipas[11,12] · Stuart D. Galloway[13] · Iker Muñoz[14] · Domingo J. Ramos-Campo[15] · Patrick Schneeweiss[16,17] · Sergio Sellés-Pérez[18] · Thomas Stöggl[19] · Rune K. Talsnes[8] · Christoph Zinner[20] · Stephen Seiler[21]

Accepted: 13 November 2024

Abstract

Background Endurance athletes tend to accumulate large training volumes, the majority of which are performed at a low intensity and a smaller portion at moderate and high intensity. However, different training intensity distributions (TID) are employed to maximize physiological and performance adaptations.

Objective The objective of this study was to conduct a systematic review and network meta-analysis of individual participant data to compare the effect of different TID models on maximal oxygen uptake (VO_{2max}) and time-trial (TT) performance in endurance-trained athletes.

Methods Studies were included if: (1) they were published in peer reviewed academic journals, (2) they were in English, (3) they were experimental or quasi-experimental studies, (4) they included trained endurance athletes, (5) they compared a polarized (POL) TID intervention to a comparator group that utilized a different TID model, (6) the duration in each intensity domain could be quantified, and (7) they reported VO_{2max} or TT performance. Medline and SPORTDiscus were searched from inception until 11 February 2024.

Results We included 13 studies with 348 ($n = 296$ male, $n = 52$ female) recreational ($n = 150$) and competitive ($n = 198$) endurance athletes. Mean age ranged from 17.6 to 41.5 years and VO_{2max} ranged from 46.6 to 68.3 $mL \cdot kg^{-1} \cdot min^{-1}$, across studies respectively. Based on the time in heart rate zone approach, there was no difference in VO_{2max} (SMD = − 0.06, $p = 0.68$) or TT performance (SMD = − 0.05, $p = 0.34$) between POL and pyramidal (PYR) interventions. There were no statistically significant differences between POL and any of the other TID interventions. Subgroup analysis showed a statistically significant difference in the response of VO_{2max} between recreational and competitive athletes for POL and PYR (SMD = − 0.63, $p < 0.05$). Competitive athletes may have greater improvements to VO_{2max} with POL, while recreational athletes may improve more with a PYR TID.

Conclusions Our results indicate that the adaptations to VO_{2max} following different TID interventions are dependent on performance level. Athletes at a more competitive level may benefit from a POL TID intervention and recreational athletes from a PYR TID intervention.

Published online: 31 January 2025

Key Points

When training load was quantified by time in heart rate zone, our results indicate that the adaptations to maximal oxygen uptake following different training intensity distribution (TID) interventions is dependent on performance level. Athletes at a more competitive level may benefit from a polarized (POL) TID intervention and recreational athletes from a pyramidal (PYR) TID intervention.

A pooled analysis using different methods to estimate borders between training zones among the included studies did not add to the statistical heterogeneity. This suggests that the precise method of determining training zones may be less important for predicting performance outcomes.

Small sample size studies continue to be a major issue in sport science research. Even with pooling data, we were not able to overcome this limitation for several intervention groups and therefore were unable to provide conclusions regarding the effects of threshold, low, and high TID models. The direction and magnitude of the effect for these interventions may be interpreted as a result of sampling error.

A high degree of collaboration, communication, and transparency between laboratories made this study achievable, and we strongly encourage multicenter collaboration among sport science researchers to improve statistical power to detect small but important effects of training interventions on performance outcomes.

1 Introduction

Observational studies have shown that endurance athletes following a structured training program tend to accumulate large training volumes, the majority of which is accumulated at a low intensity, with a smaller portion at higher intensity [1, 2]. There are several methods used to describe the different intensity zones/domains [2–5]. The three-intensity zone model is commonly used in scientific literature to describe training intensity distribution [TID: with Z1 demarcated below the first lactate threshold (LT_1) or ventilatory threshold (VT_1); Z2 between LT_1/VT_1 and the second lactate (LT_2) or ventilatory (VT_2) threshold; and Z3 above LT_2/VT_2] [2, 6]. The three-intensity zone model will be used to describe intensity moving forward in this study.

The term "polarized" (POL) training was introduced to describe a commonly observed TID in elite endurance athletes by using a session-goal approach to determine training load. POL consists of approximately 75–80% of training sessions performed in Z1, < 10% in Z2, and 15% to 20% in Z3 [2]. Alternatively, the term "threshold" (THR) TID is used to describe training programs which incorporated a greater portion of training sessions in Z2 (e.g., 40–50–10%) [7–10]. This TID may be more common in untrained and/or recreational athletes [11].

Several early experimental studies were conducted to determine the effectiveness of POL compared with THR on endurance performance [7–10]. These studies found mixed results, partially explained by small sample sizes (6–15 participants per group). A meta-analysis of these early experimental studies was conducted to increase the ability to detect a significant effect [12] and showed that a POL model was superior to a THR model for improving time-trial (TT) performance [13].

Further examination into the approaches used to determine TID led to the addition of a pyramidal (PYR) model [11]. PYR consists of the same relative emphasis on Z1 but with the next largest intensity component in Z2 and the smallest component in Z3 (e.g., 75–15–10%). A PYR model has been observed as the primary TID model in several programs of endurance athletes [2, 14–17]. An important note, however, is that TID has been shown to vary depending on training phase [6, 11, 14–17] and across sports [6].

It is necessary to consider the method for which TID is determined when comparing TID across studies. For example, an intervention executed as a POL TID (75–8–17%) using a session-goal approach, can be quantified as a PYR TID (91–6–3%) using heart rate (HR) based time-in-zone (TIZ) [2]. In addition, there are several methods to determine TIZ including internal load measurements such as HR [2, 14, 17], blood lactate concentration [2], and training impulse (TRIMP) [15, 16], external load such as running pace [18, 19] and mechanical power output in cycling and rowing [20], and qualitative metrics such as rate of perceived exertion (RPE) [21]. Internal and external load measurements may not entirely align with each other and with the prescribed or intended TID target [22].

Since the earlier meta-analysis comparing POL with THR [13], there have been several additional experimental and quasiexperimental studies published that compared POL with other TID models. Therefore, a more up-to-date synthesis of the evidence is possible. A recent meta-analysis that compared POL with other TID interventions indicated that POL was superior [23]. However, the authors conducted

a pairwise analysis by pooling results from interventions that differed in TID as well as exercise type (i.e., endurance versus strength), which can be misleading. A network meta-analysis (NMA) allows for a comparison of multiple interventions simultaneously to determine the effect of TID on improvement in endurance sport performance in greater precision and provides the ability to rank all the interventions in a coherent way [24].

One limitation to synthesizing aggregate data from exercise intervention studies is the inability to account for individual alterations in the training program, as results are based on the original group allocation, not necessarily the completed program [8, 25]. Furthermore, differences in individual participant characteristics such as age and baseline fitness are subject to regression to the mean [26, 27]. Therefore, conducting an NMA of individual participant data (IPD) would allow for improved accuracy, as covariate analyses would not be subject to these limitations [28]. Accordingly, the objective of this study is to conduct a systematic review and NMA of IPD to compare the effect of different TID models on maximal oxygen uptake ($\dot{V}O_{2max}$) and TT performance in endurance-trained athletes.

2 Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension statement for reporting systematic reviews incorporating NMAs was used as a guideline to structure this study and increase transparency [29].

2.1 Eligibility Criteria

Study eligibility criteria were based on the PICOS framework. Studies were included if they met the following criteria: (1) published in peer reviewed academic journals, (2) in the English language, (3) were experimental or quasi-experimental studies, (4) included recreational or competitive endurance athletes, (5) compared a POL TID intervention with a comparator group that utilized a different TID model, (6) the duration completed in each intensity domain could be quantified, and (7) the study reported results for $\dot{V}O_{2max}$ or TT performance, pre- and postintervention. Participants were considered endurance athletes if they participated in a structured training program that was specific to a mode of exercise, as previously described [30].

Studies were excluded for the following reasons: (1) the study was a companion report for an article that was included in the review, (2) the IPD was unavailable, (3) the study contained nutritional interventions (supplements, hydration, fed state, etc.), (4) participants were subject to changes in environmental conditions (heat, cold, altitude, hypoxia, hyperoxia, etc.), (5) the study included potential ergogenic devices/modalities (cooling vests, compression garments, etc.), or (6) the study included pharmacological agents.

2.2 Information Sources

Two electronic databases, Medline (Ovid) and SPORTDiscus (EBSCOhost), were used to conduct the literature search. The original search was conducted on 28 March 2023. An updated search was performed on 11 February 2024.

2.3 Search

The full line-by-line search strategy for each database is available in Supplementary Tables 1 and 2 of the Electronic Supplementary Material Appendix 1. The search strategy included commonly used terms to describe the different TID models and endurance sports. In addition, several different tenses and synonyms were used to further broaden the search. Search limits included the following: titles, abstracts, academic journal articles, and the English language.

2.4 Selection Process

Two reviewers (M.R. and G.T.) conducted two levels of screening [(1) titles and abstracts and (2) full-text articles]. Disagreements were resolved by a third reviewer (J.A.).

2.5 Data Collection Process

The corresponding authors for each study were contacted by one of the reviewers (M.R.) to obtain IPD for participant characteristics, intervention results, and the results for the outcome variables. Two reviewers (M.R. and J.A.) conducted a separate data collection from each report to obtain relevant information regarding the study design and outcome characteristics. The corresponding authors for the respective studies were contacted via email to review the extracted IPD and correct potential errors.

When multiple studies meeting the eligibility criteria reported data from the same participant sample, the publication with the most complete outcome data was considered the primary publication; otherwise, the publication with the largest sample size was considered the primary publication. If a companion publication reported data for an additional outcome of interest, the data were extracted, but publication details (e.g., sample size) were abstracted from the primary publication only.

2.6 Data Items

All data are presented as a mean [and standard deviations (SD)]. Study characteristics were extracted for the following components: participant allocation method (i.e., randomization and stratification), wash-in duration (weeks), and intervention duration (weeks).

IPD were requested and obtained for the following participant characteristics: age (years), sex, sport, performance level, baseline and follow up body mass (kg), and standing height (cm). Performance level was classified as competitive if participants competed at a high-performance level at tier 3 or above (i.e., university/college, provincial/state, national, international, or professional); otherwise they were classified as recreational [31, 32].

Intervention characteristics were extracted for the following components: the method by which exercise intensity zones were determined, the intended TID model, the intended relative TID distribution, periodization model, mesocycle structure, inclusion of resistance training, and the duration of a taper period (weeks) if included. IPD were requested for the total time completed in the different exercise intensity zones. The duration (minutes) within each intensity zone was determined from the HR values corresponding to the respective zone borders.

Exercise intensity zones were determined using the three intensity-zone model described by Seiler et al. [2], as it is the most common method used in the TID literature and is associated with a physiological framework. To facilitate data aggregation across multiple studies, the first lactate turn-point/threshold (LT_1), the gas exchange threshold (GET), and the first ventilatory threshold (VT_1) were all accepted to define the transition between zone 1 (Z1) and zone 2 (Z2). Similarly, the second ventilatory threshold (VT_2), respiratory compensation point (RCP), critical power (CP), second lactate turn-point/threshold (LT_2), and maximal lactate steady state (MLSS), were all accepted as methods to define the transition between Z2 and zone 3 (Z3).

TID models were classified by the relative proportion of total time spent in each zone: POL, Z1 > Z3 > Z2 and the polarization index > 2.00; PYR, Z1 > Z2 > Z3; THR, Z2 > Z1 > Z3; HIGH, Z3 > Z1 and Z3 > Z2; LOW, Z1 = 100%. The polarization index was used to distinguish between POL and non-POL TID using Eq. S1 when $Z2 \neq 0$ and Eq. S2 when $Z2 = 0$ [33]. The equations are provided in Supplementary Material Appendix 1.

The following outcome characteristics were extracted for $\dot{V}O_{2max}$ or peak oxygen uptake ($\dot{V}O_{2peak}$) (mL/kg/min): metabolic gas analysis system, exercise equipment and mode, initial stage intensity, stage duration, increment intensity, and measurement criteria. The terms $\dot{V}O_{2max}$ and $\dot{V}O_{2peak}$ will be referred to exclusively as $\dot{V}O_{2peak}$ moving forward, unless otherwise specified. For TT performance (seconds), the following variables were extracted: exercise mode, TT distance, TT location, inclusion of familiarization, and inclusion of competition during the TT test. IPD were requested for the timepoints including prewash in, baseline, midintervention, and follow up.

2.7 Geometry of Network

Network connectivity was visually assessed using network graphs for interventions that included $\dot{V}O_{2peak}$ and TT performance.

2.8 Risk of Bias of Individual Studies

The Cochrane Collaboration Risk of Bias 2.0 Tool [34] was used to assess the degree of bias across studies included in the review. Two reviewers (M.R. and J.A.) independently assessed the individual articles with a third reviewer (G.T.) to resolve discrepancies.

2.9 Summary Measures

$\dot{V}O_{2peak}$ was evaluated as a mean difference (MD) with associated 95% confidence intervals (CI) and prediction intervals (PI) between groups at follow up in its original units (mL/kg/min) as well as by using the standardized mean difference (SMD). TT performance was evaluated by using the SMD between intervention groups at follow up. The SMD with associated 95% CIs and PIs were used in place of expressing TT performance as MD in seconds as each study used different distance events (e.g., 2 versus 40 km). Hedges' g was used to account for small sample size bias [35].

2.10 Planned Methods of Analysis

Data were aggregated in two different ways to conduct separate analyses. First, participants were placed in the groups for which they were originally allocated (intention-to-treat analysis). The second method involved placing participants into groups based on the completed TID model described in Sect. 2.6 determined by HR TIZ (per-protocol analysis).

All statistical analyses were performed using R (version 4.3.3) [36]. Study groups were compared at baseline using a one-way ANOVA to confirm that there were no significant differences in the primary outcomes ($\dot{V}O_{2peak}$ and/or TT performance) within studies. If there were significant differences, then the comparisons were excluded from the analyses. The NMA for the MD for $\dot{V}O_{2peak}$ and the SMD for $\dot{V}O_{2peak}$ and TT performance were conducted using the NETMETA package [37], which uses a frequentist approach. A two-stage approach was used to combine the results of the

individual studies [28, 38]. The analyses were completed using a random effects model and the DerSimonian–Laird estimator [39].

Tables were used to describe the study characteristics, risk of bias, and results for $\dot{V}O_{2peak}$ and TT performance from the individual studies. A flow diagram was used to describe the article screening and selection process. A network graph was used to demonstrate the number of pairwise connections between interventions and the number of observations per treatment arm. A figure that contained forest plots was used to describe the pooled analysis of the results for $\dot{V}O_{2peak}$ and TT performance.

2.11 Assessment of Heterogeneity, Inconsistency, and Transitivity

The I^2 statistic was used to describe the degree of statistical heterogeneity [40]. Global inconsistency was assessed using a design-by-treatment interaction model [41]. Local inconsistency was assessed using the back-calculation method [separate indirect from direct evidence (SIDE)] [42]. Network transitivity was assessed by visually inspecting the study design tables in Supplementary Material Appendix 1 for characteristics including performance level and mean age.

2.12 Risk of Bias Across Studies

A comparison adjusted funnel plot was used to visually assess small study effects for the NMA [43].

2.13 Additional Analysis

Participant level covariates were examined using the two-stage approach as described by Riley et al. [28, 38] and included age and baseline $\dot{V}O_{2peak}$. The METAFOR package [44], which conducts pairwise analyses, was used to determine the effect of covariates on $\dot{V}O_{2peak}$ and TT performance for comparisons with sufficient sample size. Study level covariates included sport, performance level, weeks, weekly training duration, polarization index, and TT distance.

3 Results

3.1 Study Selection

The electronic databases Medline (Ovid) and SPORTDiscus (EBSCOhost) produced a total of 559 results. Following the removal of 94 duplicates, 465 titles and abstracts were screened. A total of 29 full-text articles were identified, retrieved, and screened for eligibility. Thirteen studies were included in the study (Fig. 1). The final literature search, which was performed on 11 February 2024, found 45 additional studies from the initial search; however, none met the eligibility criteria. The results of the line-by-line literature search and the explanations for study selection can be found in Supplementary Tables 1–3 of Supplementary Material Appendix 1.

3.2 Summary of Network Geometry

There were 16 pairwise comparisons for $\dot{V}O_{2peak}$ and 12 for TT performance included in the intention-to-treat analysis. There were two open connections for $\dot{V}O_{2peak}$ (PYR versus LOW and THR versus LOW) and five open connections for TT performance (PYR versus THR, PYR versus HIGH, PYR versus LOW, THR versus LOW, and THR versus HIGH). The graphs for both $\dot{V}O_{2peak}$ and TT performance demonstrated a nonstar-shaped network without open connections (Fig. 2a, c).

The number of pairwise comparisons increased to 26 for the per-protocol analysis for $\dot{V}O_{2peak}$ and 27 for TT performance. The graphs for the per-protocol analysis for both $\dot{V}O_{2peak}$ and TT performance revealed a star-shaped network (Fig. 2b, d).

3.3 Study Characteristics

A description of the study design characteristics for all included studies can be found in Table 1. Full details of the study characteristics, participant characteristics, intervention characteristics, and outcome characteristics are located in Supplementary Tables 4–8 of Supplementary Material Appendix 1.

One of the studies [9] included a crossover design. Therefore, only the results following the initial intervention for each group in this study were included. The studies ranged from 4 to 21 weeks in duration, including the wash-in and intervention periods.

There was a total of 348 ($n = 296$ male, $n = 52$ female) recreational ($n = 150$) and competitive ($n = 198$) endurance athletes across five sports (cross-country skiing, cycling, rowing, running, and triathlon) included in the quantitative analysis. Additional participant data that were not available in the original reports were provided for three studies [7, 8, 45]. The range for mean age, body mass index, and $\dot{V}O_{2peak}$ was 17.6–41.5 years, 19.8–24.9 kg/m^2, and 46.6–68.3 mL/kg/min respectively, across the studies respectively.

One study [7] prescribed a relative distribution of 80–10–10%, which did not fit any of the TID models described in Sect. 2.6. This study was included in the review as analysis of the IPD using HR TIZ revealed a completed distribution of 78(5)%, 7(3)%, and 14(4)%, consistent with a POL TID. The comparator groups included PYR ($k = 9$), THR ($k = 2$), HIGH ($k = 2$), and LOW ($k = 3$) TID models.

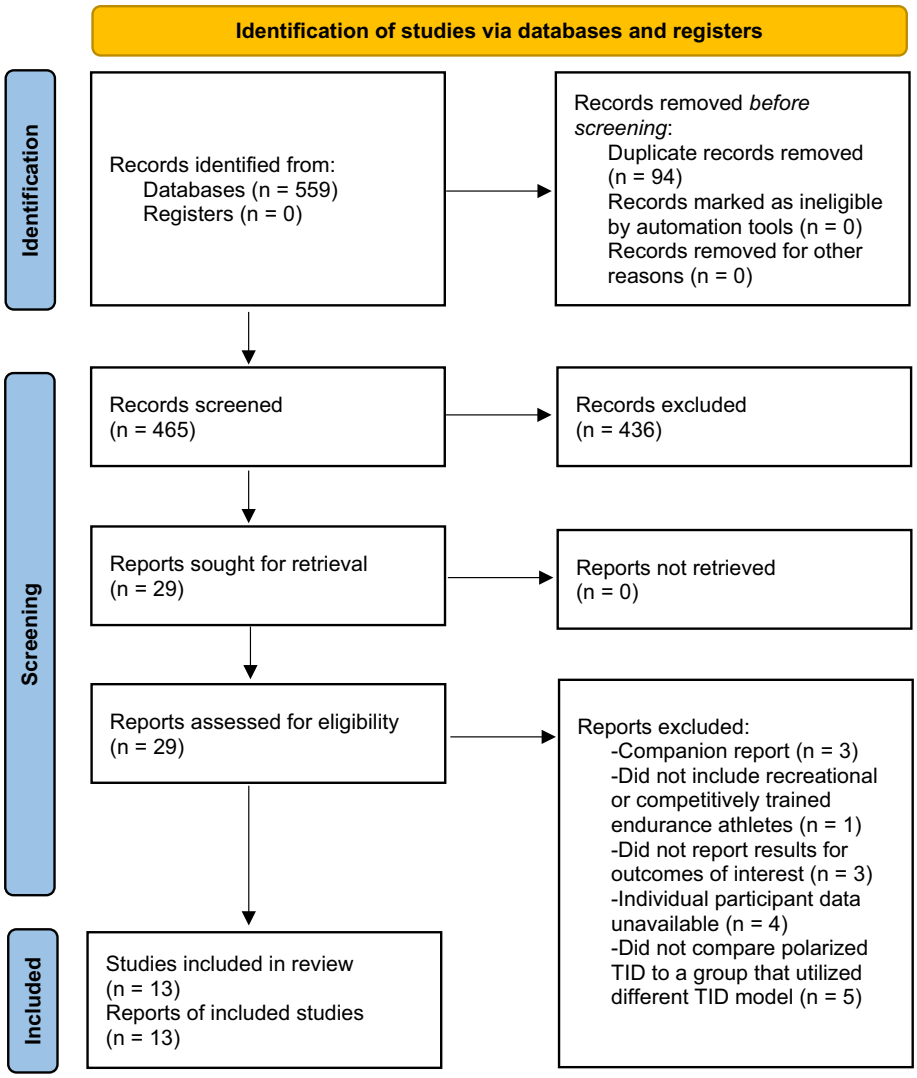

Fig. 1 PRISMA flow diagram. *TID* training intensity distribution

Eleven of the 13 studies included $\dot{V}O_{2peak}$ ($n = 264$) and 10 included TT performance ($n = 260$). Regarding the TT performance tests, two studies [8, 46] included competitors, one study [8] used a different course at follow-up, and in one study [46] the environmental conditions at follow up resulted in slower TT times.

3.4 Risk of Bias within Studies

A total of 4 of the 11 studies that included $\dot{V}O_{2peak}$, and 3 of the 10 studies that included TT performance were considered to have a high risk of bias. The full risk of bias results can be found in Supplementary Tables 9 and 10 of Supplementary Material Appendix 1.

3.5 Results of Individual Studies

The results for the weekly TID as determined by HR for the groups as initially allocated (intention-to-treat) and the retrospective group reallocation (per-protocol) are in Supplementary Tables 11 and 12 of Supplementary Material Appendix 1. Group adherence was maintained in five of the 13 studies [46, 48–51] with a total of 67 (19%) participants altering their TID model during their training programs.

There were no differences between the intervention and comparator groups at baseline for $\dot{V}O_{2peak}$ and TT performance across studies. Following group reallocation, there was a significant difference between groups for $\dot{V}O_{2peak}$ at baseline for one study [8], which was subsequently excluded from the per-protocol analysis.

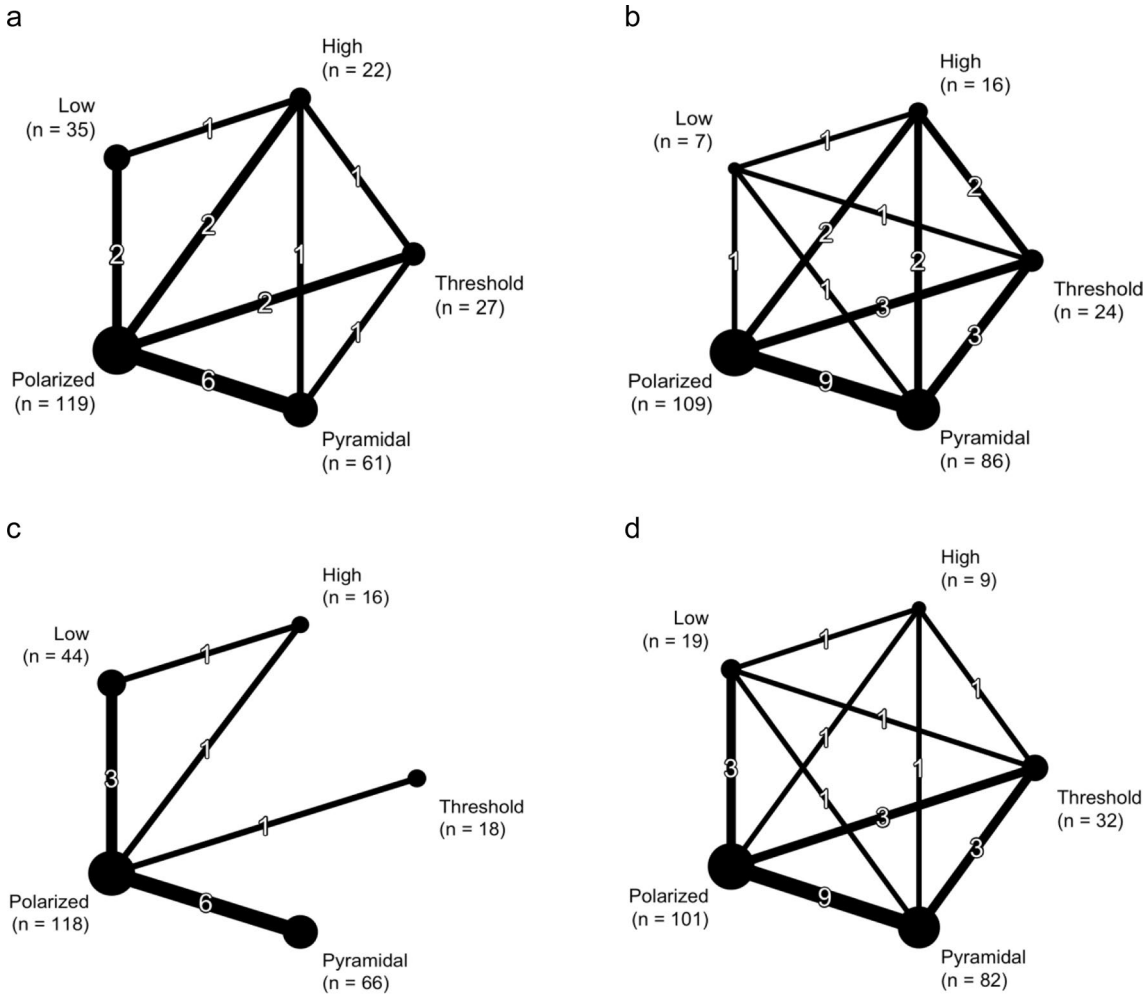

Fig. 2 Network graphs of training intensity distribution interventions. **a** Peak oxygen uptake ($\dot{V}O_{2peak}$) (mL/kg/min) for the intention-to-treat analysis. (**b**) $\dot{V}O_{2peak}$ (mL/kg/min) for the per-protocol analysis. **c** Time-trial (TT) performance for the intention-to-treat analysis. **d** TT performance for the per-protocol analysis. Line widths are proportional to the number of studies directly comparing treatments. The numbers on the connection lines indicates the number of direct pairwise comparisons between interventions. Point sizes are proportional to the number of observations in treatment arms

The results for $\dot{V}O_{2peak}$ and TT performance with the participants in their original groups can be found in Supplementary Tables 13 and 14 of Supplementary Material Appendix 1. The results for $\dot{V}O_{2peak}$ and TT performance following a retrospective group reallocation based on the completed TID using the IPD are presented in Tables 2 and 3. Two studies did not conduct follow up testing for $\dot{V}O_{2peak}$ [7, 8]. However, $\dot{V}O_{2peak}$ was not used as an outcome measure to determine the effectiveness of the interventions in the respective studies. Two studies combined participants that completed TT tests that differed in distance into a single analysis [46, 53]. In the per-protocol analysis, the participants in both studies, respectively, were placed into groups based on TT distance (i.e., the same outcome measure).

3.6 Synthesis of Results

3.6.1 $\dot{V}O_{2peak}$

The synthesized results for the MD and SMD for $\dot{V}O_{2peak}$ for the intention-to-treat analysis and for the per-protocol analysis can be found in Fig. 3. There were nine studies ($n = 264$) that were included in the intention-to-treat analysis for $\dot{V}O_{2peak}$ [10, 25, 45, 47–51, 53]. The NMA consisted of POL ($n = 119$), PYR ($n = 61$), THR ($n = 27$), HIGH ($n = 22$), and LOW ($n = 35$) TID models. When compared with a POL TID model, there were no significant differences in the SMD for PYR ($p = 0.81$), THR ($p = 0.27$), HIGH ($p = 0.25$), or LOW ($p = 0.15$).

Following group reallocation (per-protocol analysis) there was a change in sample size for each intervention

Table 1 Study characteristics

Study	Study characteristics		Participant characteristics					Intervention characteristics	Outcome characteristics
	Design	Duration (weeks)	Sample size	Sport	Performance level	Age (years)	$\dot{V}O_{2peak}$(mL/kg/min)	Group (TID)	
Esteve-Lano et al. [7]	RCT	21	15	Running	Competitive	26.4 (5.4)	67.0 (6.1)	Z1 (POL) Z2 (PYR)	$\dot{V}O_{2max}$, 10.4 km TT
Festa et al. [47]	RCT	8	35	Running	Recreational	41.5 (8.6)	53.3 (6.5)	PET (POL) FOC (THR)	$\dot{V}O_{2max}$, 2.0 km TT
Filipas et al. [48]	RCT	22	30	Running	Competitive	36.6 (5.6)	68.3 (3.8)	POL PYR	$\dot{V}O_{2peak}$, 5.0 km TT
Filipas et al. [49]	RCT	12	30	Running	Competitive	34.4 (6.3)	68.1 (2.8)	POL PYR	$\dot{V}O_{2peak}$, 5.0 km TT
Muñoz et al. [8]	RCT	18	35	Running	Recreational	34.3 (7.4)	61.6 (7.9)	PET (POL) BThET (PYR)	$\dot{V}O_{2max}$, 10.0 km TT
Neal et al. [9]	RCT-X	10	11	Cycling	Competitive	37.5 (6.2)	–	POL THR (PYR)	40.0 km TT
Pérez et al. [50]	RCT	12	20	Running	Recreational	37.9 (8.9)	56.8 (5.2)	POL THR (PYR)	$\dot{V}O_{2max}$
Schneeweiss et al. [46]	RCT	3	18	Cycling	Competitive	17.9 (3.6)	–	POL LIT (LOW)	8.4/12.6 km TT
Sellés-Pérez et al. [51]	NRCT	23	15	Triathlon	Recreational	29.1 (6.7)	55.7 (4.5)	POL PYR	$\dot{V}O_{2max}$
Stöggl et al. [52]	RCT	9	41	Combined	Competitive	31.4 (6.7)	62.0 (7.5)	POL HVT (PYR) THR HIIT (HIGH)	$\dot{V}O_{2peak}$
Talsnes et al. [53]	RCT	13	39	XC-skiing	Competitive	17.6 (0.7)	66.7 (7.1)	HITG (POL) LOWG (LOW)	$\dot{V}O_{2peak}$, 4.5/6.4 km TT
Treff et al. [25]	NRCT	11	14	Rowing	Competitive	20.1 (2.0)	65.6 (5.2)	POL PYR	$\dot{V}O_{2max}$, 2.0 km TT
Zinner et al. [45]	NRCT	7	45	Running	Recreational	27.6 (6.9)	46.6 (6.6)	POL HIIT (HIGH) HVT (LOW)	$\dot{V}O_{2peak}$, 5.0 km TT

All data are expressed as a mean (SD)

BM body mass index, combined (cross-country skiing cycling running triathlon), *COMP* competitive, *XC-skiing* cross-country skiing, $\dot{V}O_{2max}$ maximal oxygen uptake, *NRCT* nonrandomized controlled trial, $\dot{V}O_{2peak}$ peak oxygen uptake, *RCT* randomized controlled trial, *RCT-X* randomized controlled trial with crossover, *REC* recreational. *BThET* between thresholds endurance training, *FOC* focused endurance training, *HIIT* high-intensity interval training, *HITG* high-intensity training group, *HVT* high volume training, *LIT, LOW* low intensity training, *LITG* low intensity training group, *PET, POL* polarized, *PYR* pyramidal, *THR* threshold hold training, *TID* training intensity distribution, *Z1* zone 1, *Z2* zone 2

group (POL: $n = 109$, PYR: $n = 86$, THR: $n = 24$, HIGH: $n = 16$, LOW: $n = 7$). As in the intention-to-treat analysis, there were no significant differences between POL and PYR ($p = 0.68$), THR (0.17), HIGH ($p = 0.33$), LOW ($p = 0.13$).

3.6.2 TT Performance

The synthesized results for the SMD for TT performance for the intention-to-treat analysis, and the per-protocol analysis can be found in Fig. 3. There were ten studies ($n = 262$) that were included in the intention-to-treat analysis for TT performance [7–9, 25, 45–49, 53]. The NMA consisted of POL ($n = 118$), PYR ($n = 66$), THR ($n = 18$), HIGH ($n = 16$), and LOW ($n = 44$) TID models. When compared with a POL TID model, there were no significant differences in the SMD for PYR ($p = 0.34$), THR ($p = 0.91$), HIGH ($p = 0.52$), or LOW ($p = 0.75$).

Following group reallocation (per-protocol analysis) there was a change in sample size for each intervention group (POL: $n = 101$, PYR: $n = 82$, THR: $n = 32$, HIGH: $n = 9$, LOW: $n = 19$). As in the intention-to-treat analysis, there were no significant differences between POL and PYR ($p = 0.75$), THR (0.19), HIGH ($p = 0.70$), LOW ($p = 0.30$).

3.7 Exploration of Heterogeneity, Inconsistency, and Transitivity

There was no statistical heterogeneity for the intention-to-treat analyses for the MD in $\dot{V}O_{2peak}$ ($I^2 = 5.1\%$, 95% CI 0.0–66.6%), the SMD for $\dot{V}O_{2peak}$ ($I^2 = 3.5\%$, 95% CI 0.0–66.6%), or for the SMD for TT performance ($I^2 = 0.0\%$, 95% CI 0.0–67.6%). There was also no statistical heterogeneity for the per-protocol analyses for the MD in $\dot{V}O_{2peak}$ ($I^2 = 0.0\%$, 95% CI 0.0–58.3%), the SMD for $\dot{V}O_{2peak}$ ($I^2 = 0.0\%$, 95% CI 0.0–58.3%), or for the SMD for TT performance ($I^2 = 0.0\%$, 95% CI 0.0–56.6%).

Table 2 Results for peak oxygen uptake of individual studies (per-protocol analysis)

Study	Group	Group size	Completed TID model	Baseline (mL/kg/min)	Follow-up (mL/kg/min)	Delta (mL/kg/min)	Delta (%)
Festa et al. [47]	1	19	Polarized	53.0 (5.9)	53.6 (4.8)	0.6 (2.9)	1.5 (5.7)
	2	4	Pyramidal	56.7 (6.5)	56.6 (6.7)	−0.0 (1.2)	−0.1 (2.0)
	3	12	Threshold	52.7 (7.7)	52.1 (7.9)	−0.6 (3.2)	−1.1 (7.2)
Filipas et al. [48]	1	15	Polarized	68.5 (3.3)	69.9 (3.6)	1.4 (2.0)	2.1 (3.0)
	2	15	Pyramidal	68.1 (4.3)	68.9 (3.5)	0.8 (1.5)	1.3 (2.2)
Filipas et al. [49]	1	15	Polarized	68.1 (2.9)	68.1 (3.9)	0.1 (1.4)	0.1 (2.1)
	2	15	Pyramidal	68.2 (2.8)	68.8 (2.6)	0.6 (1.3)	0.9 (1.9)
Pérez et al. [50]	1	11	Polarized	56.7 (5.5)	56.1 (6.3)	−0.6 (3.0)	−1.1 (5.3)
	2	9	Pyramidal	56.8 (5.0)	58.4 (4.7)	1.7 (2.9)	3.1 (5.2)
Sellés-Pérez et al. [51]	1	7	Polarized	54.7 (4.3)	55.1 (5.5)	0.4 (6.0)	1.2 (11.4)
	2	8	Pyramidal	56.5 (4.8)	57.6 (4.9)	1.1 (2.9)	2.1 (5.6)
Stöggl et al. [52]	1	9	Polarized	61.7 (8.1)	67.0 (8.5)	5.3 (1.8)	8.8 (3.1)
	2	10	Pyramidal	60.9 (8.4)	61.9 (9.1)	0.9 (3.2)	1.6 (5.1)
	3	8	Threshold	63.3 (5.2)	60.5 (7.9)	−2.8 (4.4)	−4.7 (7.5)
	4	9	High	63.7 (7.1)	66.6 (5.8)	2.8 (3.5)	4.8 (5.9)
Talsnes et al. [53]	1a	9	Polarized	57.1 (3.9)	59.1 (4.0)	2.1 (2.7)	3.7 (4.9)
	2a	1	Pyramidal	60.1 (0.0)	58.6 (0.0)	−1.5 (0.0)	−2.5 (0.0)
	1b	19	Polarized	70.7 (4.9)	70.3 (3.6)	−0.4 (3.7)	−0.4 (5.5)
	2b	4	Pyramidal	68.7 (4.5)	65.8 (5.6)	−2.9 (1.6)	−4.3 (2.5)
Treff et al. [25]	1	5	Polarized	67.7 (7.9)	67.4 (7.1)	−0.3 (1.4)	−0.3 (1.8)
	2	5	Pyramidal	63.5 (3.0)	65.1 (0.8)	1.6 (3.5)	2.7 (5.6)
Zinner et al. [45]	1	9	Polarized	45.7 (5.8)	46.0 (5.2)	0.3 (2.8)	1.0 (6.1)
	2	16	Pyramidal	44.5 (7.0)	46.3 (7.5)	1.8 (2.5)	4.1 (5.7)
	3	4	Threshold	43.2 (2.3)	45.7 (2.9)	2.5 (1.3)	5.7 (2.8)
	4	7	High	47.5 (3.2)	48.7 (5.0)	1.2 (4.6)	2.7 (9.4)
	5	7	Low	50.7 (7.7)	51.2 (9.9)	0.5 (3.1)	0.6 (5.0)

All data are expressed as a mean (SD). Talsnes et al. [53] groups 1a and 2a include female athletes who completed a 4.6 km TT, and groups 1b and 2b include male athletes who completed a 6.4 km TT

Table 3 Results for time-trial performance of individual studies (per-protocol analysis)

Study	Group	Group size	Completed TID model	Baseline (s)	Follow-up (s)	Delta (s)	Delta (%)
Esteve-Lano et al. [7]	1	3	Polarized	2203.7 (131.5)	2082.0 (96.4)	− 121.7 (35.2)	− 5.5 (1.3)
	2	11	Pyramidal	2257.9 (138.0)	2116.3 (128.8)	− 141.6 (58.4)	− 6.2 (2.4)
Festa et al. [47]	1	18	Polarized	532.9 (79.7)	513.2 (70.9)	− 19.8 (34.7)	− 3.4 (5.5)
	2	4	Pyramidal	543.2 (69.8)	512.5 (70.1)	− 30.8 (13.0)	− 5.7 (2.4)
	3	12	Threshold	520.7 (65.9)	510.1 (60.3)	− 10.6 (13.4)	− 1.9 (2.4)
Filipas et al. [48]	1	15	Polarized	997.7 (47.9)	986.3 (47.0)	− 11.5 (11.0)	− 1.1 (1.1)
	2	15	Pyramidal	992.8 (56.6)	986.7 (55.7)	− 6.1 (5.8)	− 0.6 (0.6)
Filipas et al. [49]	1	15	Polarized	997.1 (34.5)	989.7 (37.5)	− 7.5 (5.9)	− 0.8 (0.6)
	2	15	Pyramidal	992.9 (31.8)	983.5 (32.6)	− 9.3 (7.0)	− 0.9 (0.7)
Muñoz et al. [8]	1	5	Polarized	2346.4 (312.8)	2235.0 (350.3)	− 111.4 (51.4)	− 5.0 (2.6)
	2	20	Pyramidal	2373.4 (309.6)	2242.4 (274.0)	− 131.1 (76.4)	− 5.4 (2.6)
	3	7	Threshold	2600.4 (247.0)	2536.4 (228.8)	− 64.0 (109.8)	− 2.4 (4.1)
Neal et al. [9]	1	6	Polarized	3891.8 (151.3)	3764.7 (146.2)	− 127.2 (78.6)	− 3.3 (2.0)
	2	4	Pyramidal	3830.0 (325.0)	3896.5 (394.7)	66.5 (248.8)	1.8 (6.8)
	3	1	Threshold	3761.0 (0.0)	3780.0 (0.0)	19.0 (0.0)	0.5 (0.0)
Schneeweiss et al. [46]	1a	5	Polarized	2481.6 (213.1)	2571.6 (343.9)	90.0 (141.1)	3.3 (5.1)
	2a	3	Low	2718.7 (352.8)	2855.7 (464.6)	137.0 (167.6)	4.8 (6.2)
	1b	5	Polarized	3239.0 (85.3)	3334.0 (222.1)	95.0 (139.9)	2.8 (4.3)
	2b	5	Low	3262.0 (145.4)	3217.2 (109.2)	− 44.8 (108.8)	− 1.3 (3.3)
Talsnes et al. [53]	1a	9	Polarized	1450.8 (59.5)	1429.6 (81.8)	− 21.2 (43.0)	− 1.5 (3.0)
	2a	1	Pyramidal	1410.0 (0.0)	1353.0 (0.0)	− 57.0 (0.0)	− 4.0 (0.0)
	1b	15	Polarized	1756.5 (139.6)	1719.9 (145.8)	− 36.6 (50.7)	− 2.1 (2.8)
	2b	2	Pyramidal	1814.5 (149.2)	1827.5 (201.5)	13.0 (52.3)	0.6 (2.8)
Treff et al. [25]	1	5	Polarized	371.9 (11.5)	370.1 (9.4)	− 1.8 (3.5)	− 0.5 (1.0)
	2	6	Pyramidal	368.6 (8.3)	367.2 (7.0)	− 1.4 (2.0)	− 0.4 (0.5)
Zinner et al. [45]	1	9	Polarized	1576.6 (196.5)	1514.4 (188.6)	− 62.1 (31.0)	− 3.9 (1.9)
	2	16	Pyramidal	1681.4 (339.4)	1616.5 (329.7)	− 64.9 (199.0)	− 3.3 (10.2)
	3	4	Threshold	1717.2 (195.5)	1573.5 (231.6)	− 143.8 (213.9)	− 8.1 (12.2)
	4	7	High	1538.4 (153.3)	1502.4 (154.4)	− 36.0 (66.6)	− 2.3 (4.5)
	5	5	Low	1511.2 (394.2)	1381.6 (311.1)	− 129.6 (85.1)	− 8.0 (2.9)

All data are expressed as a mean (SD). Schneeweiss et al. [46] groups 1a and 2a include junior and female athletes who completed an 8.4 km time-trial (TT), and groups 1b and 2b include senior athletes who completed a 12.6 km TT. Talsnes et al. [53] groups 1a and 2a include female athletes who completed a 4.6 km TT, and groups 1b and 2b include male athletes who completed a 6.4 km TT

There was no indication of global inconsistency for the intention-to-treat analyses for the MD in $\dot{V}O_{2peak}$ ($Q = 3.51$, $p = 0.48$), the SMD for $\dot{V}O_{2peak}$ ($Q = 3.57$, $p = 0.47$), or for the SMD for TT performance ($Q = 0.12$, $p = 0.73$). There was also no indication of global inconsistency for the per-protocol analyses for the MD in $\dot{V}O_{2peak}$ ($Q = 3.16$, $p = 0.79$), the SMD for $\dot{V}O_{2peak}$ ($Q = 2.20$, $p = 0.90$), or for the SMD for TT performance ($Q = 2.47$, $p = 0.65$). Furthermore, there was also no indication of local inconsistency for any of the analyses; the results can be found in Supplementary Tables 27–32 in Supplementary Material Appendix 1.

With respect to network transitivity, there was no imbalance across studies for study characteristics including performance level and mean age.

3.8 Risk of Bias Across Studies

Visual inspection of funnel plot asymmetry did not indicate the presence of small sample size bias in the results for any of the pooled analyses. Funnel plots for each of the analyses are in Supplementary Fig. 1 of Supplementary Material Appendix 1.

3.9 Results of Additional Analyses

3.9.1 $\dot{V}O_{2peak}$

Stratified meta-analysis showed a statistically significant difference between recreational and competitive athletes (SMD = − 0.63, 95% CI = − 1.24 to − 0.02, $p < 0.05$;

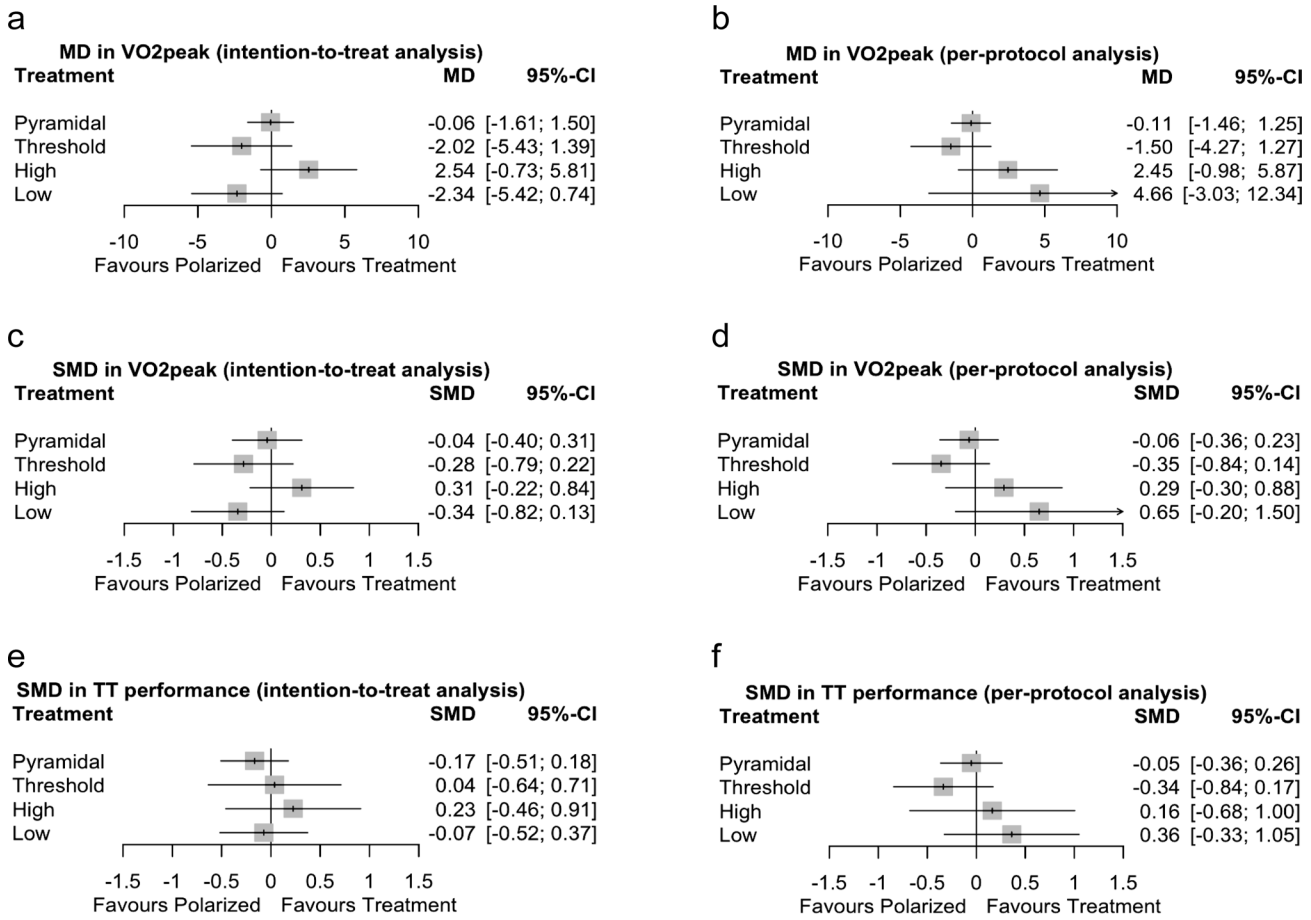

Fig. 3 Forest plots of **a** mean difference (MD) in peak oxygen uptake ($\dot{V}O_{2peak}$) (mL/kg/min) for the intention-to-treat analysis; **b** MD in $\dot{V}O_{2peak}$ (mL/kg/min) for the per-protocol analysis; **c** standardized mean difference (SMD) in $\dot{V}O_{2peak}$ for the intention-to-treat analysis; **d** SMD in $\dot{V}O_{2peak}$ for the per-protocol analysis; **e** SMD in time-trial (TT) performance for the intention-to-treat analysis; and **f** SMD in TT performance for the per-protocol analysis

$I^2 = 0\%$, 95% CI = 0.0% to 63.5%; MD of approximately 2.7 mL·kg^{-1}·min^{-1}). Although not statistically significant but clinically meaningful, recreational athletes favored PYR (SMD = 0.32, 95% CI = − 0.14 to 0.79, $p = 0.18$; $I^2 = 0.0$, 95% CI = 0.0–68.2%) and competitive athletes favored POL (SMD = − 0.31, 95% CI = − 0.70 to 0.09, $p = 0.13$; $I^2 = 2.2\%$, 95% CI = 0.0–88.2%) (Fig. 4). There was a significant difference ($p < 0.001$) in baseline $\dot{V}O_{2peak}$ (mL·kg^{-1}·min^{-1}) between recreational (53.0, 95% CI = 49.7–56.3) and competitive athletes (67.0, 95% CI = 64.0–69.8). An additional subgroup analysis was conducted to determine if the difference in $\dot{V}O_{2peak}$ was because of sex-based differences between the recreational and competitive subgroups; however, this was not found to be the case.

There was no significant difference across studies for the other covariates including age ($p = 0.86$), sport ($p = 0.16$), weekly training duration (minutes) ($p = 0.22$), weeks ($p = 0.83$), or polarization index ($p = 0.98$). In addition to the covariate analysis, we conducted a bivariate meta-analysis to determine if the total weekly training duration differed between POL and PYR interventions. The results of the analysis showed there to be no statistically significant difference in weekly training duration between interventions.

3.9.2 TT Performance

There was no significant effect on TT performance from covariates including age ($p = 0.18$), baseline $\dot{V}O_{2peak}$ ($p = 0.10$), sport ($p = 0.71$), performance level ($p = 0.72$), weekly training duration (minutes) ($p = 0.95$), weeks ($p = 0.64$), polarization index ($p = 0.65$), or TT distance ($p = 0.48$).

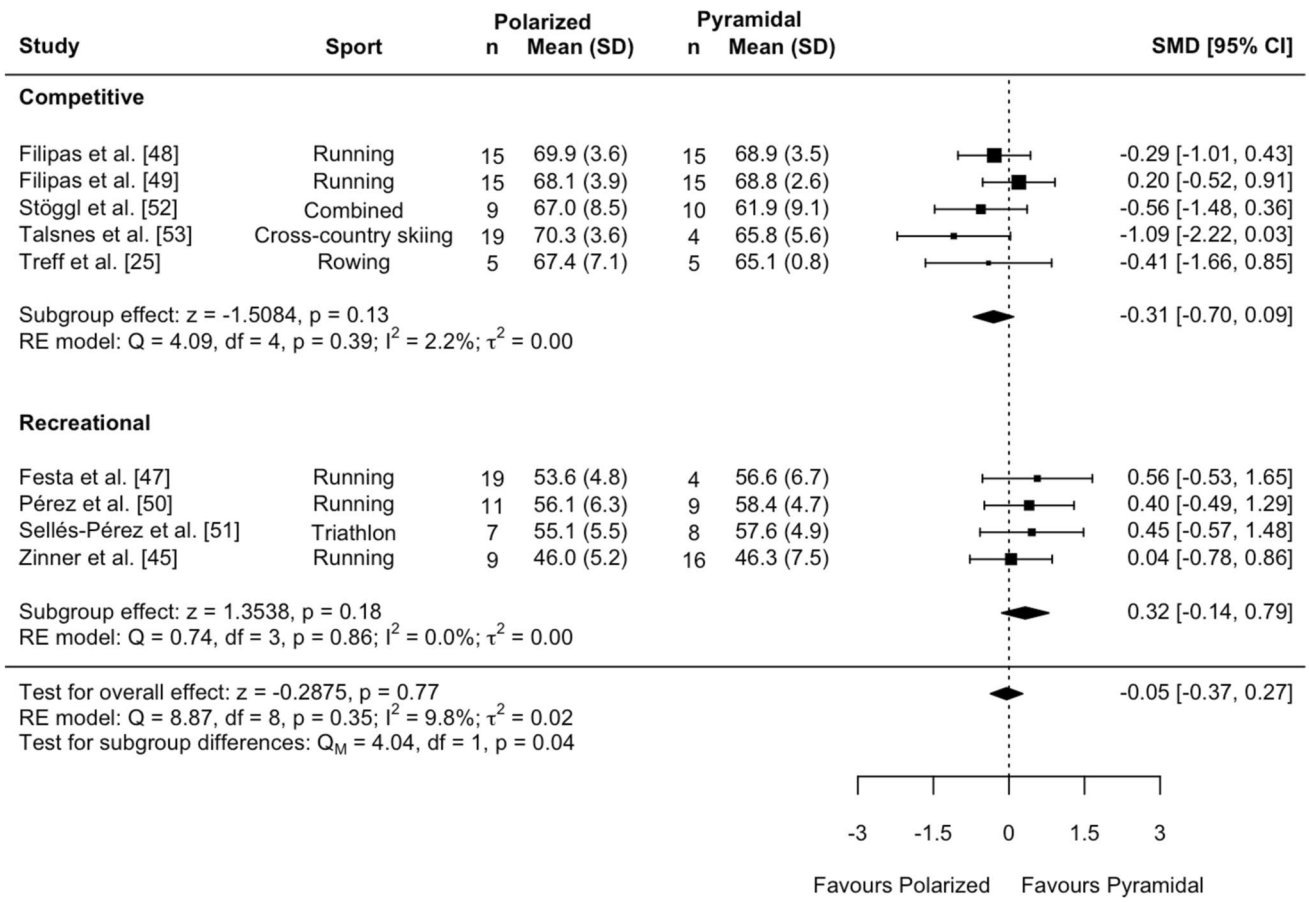

Fig. 4 Standardized mean difference (SMD) in peak oxygen uptake ($\dot{V}O_{2peak}$) (mL/kg/min) between polarized and pyramidal training intensity distribution (TID) models (per-protocol analysis). Mean (SD) values are results for $\dot{V}O_{2peak}$ at follow up [combined (cross-country skiing, cycling, running, and triathlon)]

4 Discussion

4.1 General Interpretation of the Results

We performed the first NMA of IPD on the effects of different TID interventions on $\dot{V}O_{2peak}$ and endurance performance. The primary analysis did not show a statistically significant difference in $\dot{V}O_{2peak}$ or TT performance between POL and any of the other TID models (PYR, THR, HIGH, and LOW). However, we found a significant difference in the response of $\dot{V}O_{2peak}$ between recreational and competitive athletes when quantifying training load by HR TIZ (SMD − 0.63, $p < 0.05$). Competitive athletes may have greater improvements in $\dot{V}O_{2peak}$ with POL, whereas recreational athletes may benefit more from PYR.

The per-protocol analysis of SMD for $\dot{V}O_{2peak}$ and TT performance (Fig. 3d, f) showed greater similarities than for the intention-to-treat results (Fig. 3c, e), suggesting that these results may have been related. $\dot{V}O_{2peak}$ is one of the key determinants of endurance performance [54], although it is often observed that change in $\dot{V}O_{2peak}$ alone is poorly correlated with change in TT performance during short-term training intervention studies or when monitored across a competitive season [55–58]. Using the per-protocol dataset, a linear mixed effect model analysis with random effects for study found no significant effect of the percent change in $\dot{V}O_{2peak}$ on percent change in TT performance ($\beta = 0.004$, 95% CI − 0.19 to 0.20, $p = 0.96$). Thus, while the per-protocol analysis of $\dot{V}O_{2peak}$ and TT performance did show similar results at the group level (Fig. 3d, f), those changes were not well correlated at the individual level. Longitudinal change in endurance performance can be driven by a number of interrelated factors, including fractional utilization of $\dot{V}O_{2peak}$ at the maximal metabolic steady state (MMSS), gross mechanical efficiency (economy), and peripheral muscle oxidative capacity [59, 60]. Therefore, to obtain a valid indicator that considers these

factors, TT performance should be included as an outcome measure to evaluate the effectiveness of endurance training interventions.

When considering performance level, there was a significant difference in how recreational and competitive athletes responded to POL versus PYR TID models. The data indicate greater improvements in $\dot{V}O_{2peak}$ with POL for competitive athletes, while for recreational athletes, improvements with PYR were greater (Fig. 4). While the differences between POL and PYR within each subgroup alone were not significant, the difference between subgroups was significant. This suggests that athletes at a higher competitive level may obtain greater improvements in $\dot{V}O_{2peak}$ with a POL training structure. Even though all participants were endurance-trained, only those who were reported as competing at a university/college, provincial/state, national/international, or professional level were considered competitive. Classification was not made on the basis of $\dot{V}O_{2peak}$ but on the information reported by the study authors. However, there was a significant difference in baseline $\dot{V}O_{2peak}$ between recreational and competitive subgroups (53 versus 67 mL/kg/min, respectively). Besides competitive level, there were no differences observed with any other covariates. This is consistent with previous meta-analyses [30, 31, 61] and suggests that athlete classification may be an important indicator of responsiveness to TID models [32] and should be taken into account when personalizing training interventions.

The MD for $\dot{V}O_{2peak}$ between POL and PYR was − 0.11 mL/kg/min (95% CI − 1.46 to 1.25). This is well within the typical error often reported for $\dot{V}O_{2peak}$ of approximately 1–5% (~ 50–250 mL/min or ~ 1–5 mL/kg/min) [62–65]. The inability to detect significant differences between TID models may partially be explained by small sample sizes, particularly in the THR, HIGH, and LOW groups. We previously found that a minimum of 81 participants per group would be required to show a significant difference of SMD of 0.44 for change in $\dot{V}O_{2peak}$ between different TID models [31]. In the current review, only POL and PYR exceeded this sample size after reallocation. The lower sample sizes for the remaining TID groups and for the competitive and recreational subgroups led to wider confidence intervals for the comparisons between those models (Figs. 3 and 4). The direction and magnitude of the effect for these intervention groups (THR, LOW, and HIGH) may be a result of sampling error [66]. Future research will need to ensure sufficient statistical power to detect whether these TID models have small but potentially important effects on performance outcomes among trained athletes.

Reallocation of individuals based on completed TID (i.e., per-protocol analysis) tended to reduce the uncertainty around the estimates for $\dot{V}O_{2peak}$ and TT performance. The 95% CIs for SMD between interventions were generally smaller for the per-protocol analyses than for the intention-to-treat analysis. The exception was the LOW group, which had only 7 participants for $\dot{V}O_{2peak}$ and 13 for TT performance after reallocation (reduced from $n = 25$ and 35, respectively (Fig. 3)). Athletes prescribed to LOW training groups tended to accumulate a larger than intended volume of HR in Z2, pushing some of them into a PYR TID. This may be partly explained by a divergence between HR as internal training load and external load measured with cycling power or running pace [22]. For example, if athletes were prescribed training workloads near their Z1 ceiling, their HR may have drifted into Z2 despite them adhering to the prescribed training. On the other hand, individuals prescribed to HIGH groups also tended to accumulate more HR volume in Z2 than Z3, resulting reallocation into a PYR group (from $n = 25$ to 16 for $\dot{V}O_{2peak}$ and $n = 13$ to 7 for TT performance). Given that the 95% CIs of the between-group differences were reduced after reallocation based on HR, this suggests that monitoring the completed internal training load had less variability when predicting performance outcomes compared with the originally prescribed TID model. Monitoring external training load where available may further improve these results [6], particularly for LOW and HIGH TID models where HR tended to converge toward a more PYR completed distribution.

Additionally, the 95% PIs, which represent the range of individual observations, also tended to be smaller after reallocation. The 95% PI for the difference between POL and PYR was SMD of − 0.06, 95% PI − 0.40 to 0.28 for $\dot{V}O_{2peak}$, and SMD − 0.05, 95% PI − 0.40 to 0.30 for TT performance. This is consistent with previous observations that when training is prescribed by individualized intensity zones, the variability in responsiveness to the intervention is reduced [3, 67, 68]. PIs show the real variability in responses among the population of endurance-trained athletes to different TID models. While there were no detectable differences between TID models at the group level in this review, individual athletes may respond better or worse to a particular intervention within a wide range around the group mean.

The meta-analysis by Rosenblat et al. [13] suggested that POL was associated with improved TT performance in endurance-trained athletes. However, the review did not collect individual participant data or reallocate based on intended versus completed TID. In addition, three of the studies pooled to compare TT performance included groups described as THR, when in fact they followed a PYR TID [7–9]. Therefore, a reconsideration of these studies would have suggested that POL showed enhanced TT performance compared with PYR not THR. Furthermore, the analysis was performed by comparing percentage change from baseline.

Since one study performed baseline and follow up testing on different TT courses [8], this approach would not be valid. A reanalysis of the differences between groups at follow up no longer resulted in a significant difference between groups (SMD 0.20, 95% CI − 0.46 to 0.87, $p = 0.55$).

A more recent meta-analysis by Silva Oliveira et al. [23] indicated that POL was superior to other exercise interventions. However, the study has several limitations including double counting of participants, comparing the change from baseline instead of comparing intervention groups at follow up, pooling results that include both sprint and long distance TTs, and combining all interventions (including strength training) into a single pairwise analysis. Therefore, the results may not accurately represent the effect of POL when compared with other exercise interventions.

4.2 Limitations of the Evidence Included in the Review

There were various limitations with the available studies and IPD included in this review. For example, three studies did not include a randomization process [25, 45, 51]. Fewer than 85% of initially allocated participants completed the intervention and follow-up testing in five of the studies [7, 46, 47, 50, 51], meaning the intention-to-treat analysis could not include these participants. For this reason, we considered the intention-to-treat analysis to be the original group allocation, and the per-protocol analysis for the reallocated groups by completed TID. Several studies did not report the intended TID [10, 25, 45, 50, 51, 53]. Females comprised only 15% of participants and were not represented at all in seven studies [7, 9, 25, 48–51]), which may limit generalizability of these results to female athletes.

TT test procedures used across studies were particularly heterogeneous. Only two studies included a familiarization bout [9, 45], likely resulting in wider variability and less ability to detect a significant difference between interventions. Two studies included racing with competitors that may have influenced pacing [8, 46]. One study used different courses for baseline and follow-up testing [8], one study had markedly different environmental conditions at follow-up [46], and two studies included TTs of different distances and times for male and female participants and pooled the results when analyzing between-group differences [46, 53].

4.3 Limitations of the Review Process

A limitation of the available data is the heterogeneity with which different studies classified exercise "thresholds" and zones, and the methods used to prescribe training intensity [2–5]. As mentioned, internal and external training loads may lead to different estimates for TID [22]. HR is a commonly used metric for internal training load allowing for quantification across multiple sports [6]. The discrepancy between prescribed (both external and internal across the included studies) and completed (internal) TID tended to result in more athletes converging into PYR from both LOW and HIGH TID models when reallocated by HR, as more time was accumulated in Z2. However, the statistical heterogeneity in the current review was low, reallocation of groups resulted in smaller confidence intervals around the MD and SMD for $\dot{V}O_{2peak}$ and TT performance, and there were no differences in outcomes between the intention-to-treat and per-protocol analyses. Therefore, monitoring TID by HR improved prediction of performance outcomes; however, including external training load may further improve these results.

The studies included in the meta-analysis differed to some extent in terms of intensity zone definitions based on, for example, lactate or ventilatory "thresholds" or critical power concepts. Since previous findings indicate that the methods used to determine the 'borders' between domains can occur at different percentages of $\dot{V}O_{2peak}$ [69] and of maximum mechanical power output [70], it is probable that the precise boundaries between intensity zones varied across the studies applying these different approaches. Subsequently, combining studies that use different methods to program exercise might add inaccuracy to the results [71]. Therefore, we used a random effects model as opposed to a fixed effect model to adjust the weight of the included studies. A random effects method considers both the within-study (standard error which is linked to sample size) and between-study variance (the variation in the effect measures across studies) [72]. The benefits of considering both forms of variance are that studies with a large standard error (i.e., small sample sizes) as well as those where differences in methodology (i.e., clinical heterogeneity) influence the effect measure, will have less weight on the synthesized effect estimate.

One method to detect if the between study variance influenced the pooled effect is by using Higgin's I^2 to calculate the percentage of the total variation in the estimated effect across studies (i.e., the statistical heterogeneity) [40]. If the difference in the measurements (e.g., LT_1 vs VT_1) had systematically influenced the results, there would have been a high degree of statistical heterogeneity. There was virtually no evidence of statistical heterogeneity for any of the results of the present NMA (I^2 values ranged from 0% to 5%), as described in Sect. 3.7. Therefore, it is unlikely that the different methods used to define intensity zones in the studies had a meaningful statistical impact on our findings. This suggests that although there are a variety of methods

used to estimate boundaries between training zones, with potentially large differences between estimates [73, 74], the particularly selected method may not be as important for predicting performance outcomes.

This study is also limited by the data available regarding the duration and periodization of the training interventions. Training interventions ranged between 3 and 18 weeks. Individuals completed an average of 385 (198) min of training per week. We also did not consider the influence of mesocycle and microcycle training volume distribution within each training intervention. Four studies did not progress training load within the intervention period [9, 25, 46, 53], while four used 2 weeks of build to 1 week taper [7, 8, 10, 47] and five used 3 weeks of build to 1 week taper [45, 48–51]. Filipas et al. 2022 [37] showed that changing the TID model during a program may influence results. Therefore, we cannot say how differences in mesocycle periodization may have influenced results within 18 weeks nor can we speculate how results may change over longer seasonal or quadrennial training programs. Furthermore, the available data did not allow definitive conclusions to be drawn on the relationship between the TID effect and sport specificity, primarily due to small sample sizes. Such associations are not unlikely as the disciplines differ significantly in terms of mechanical load, competition duration, intensity, and strength/endurance ratio.

4.4 Implications of the Results for Practice and Future Research

Our results indicate that athletes of varying performance levels may benefit from distinct TID models. Specifically, recreational athletes appear to gain greater benefit to $\dot{V}O_{2peak}$ from a PYR model, whereas competitive athletes may benefit from a POL model. Recreational athletes with less training experience likely have more potential to improve performance with any periodization model, while athletes with more training experience may need more specific stimuli to continue to improve [30]. This discrepancy could suggest that competitive athletes require sufficient amounts of high intensity training combined with large volumes of low intensity training to stimulate further cardiovascular (and central) adaptations, especially over shorter time periods which are used in most of the included studies.

The mean change in TT performance was around 1.8% across all participants, with wide individual variability. While a 1.8% improvement might be considered a meaningful improvement and could be decisive of victory in competition [75], the variable nature of TT testing makes it difficult to detect small differences between interventions [76]. Likely for this reason, we did not detect any differences in TT performance between POL and PYR or the other TID models (Fig. 3e,f). Future studies need to increase sample sizes to achieve more robust conclusions, which we also discussed in our previous review [31]. With the availability and advancements of wearable technology, large-crowd data-sourced studies and retrospective analysis of available datasets may be a viable approach to achieve this objective [6, 31]. In addition, all but three of the included studies employed a wash-in period to control for the participants' previous training [10, 46, 50]. Standardising wash-in protocols or at the very least recording the previous training leading into the study period may help to better account for effects of the intervention, per se, in participants already engaged in regular training.

Finally, the majority of included studies employed HR to monitor exercise intensity and subsequently determine the TID. There are several issues with this method, as previously mentioned [22]. Therefore, to improve accuracy, studies should record external work in addition to internal work where possible, as this provides more precise information regarding exercise intensity. This approach would further enhance our understanding of the longitudinal relationship between internal and external training load measurements across a range of endurance sports and participant characteristics [6].

5 Conclusions

The objective of this study was to compare the effects of different TID models on the changes in $\dot{V}O_{2peak}$ and TT performance among endurance-trained athletes, reallocating individual participant data into their completed TID models based on HR as a measurement of internal training load. Among the entire dataset, we observed no statistically significant differences in $\dot{V}O_{2peak}$ or TT performance between POL and any of the other TID models (PYR, THR, HIGH, and LOW). However, we found a statistically significant and meaningful difference in the response of $\dot{V}O_{2peak}$ between recreational and competitive athletes when quantifying training load by HR TIZ. This suggests that individual athletes can gain similar endurance performance benefits with different TID models but that athletes at higher competitive levels with greater training experience might benefit more from a POL TID model.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40279-024-02149-3.

Acknowledgements We thank Andrew J. Carnes, Paulina Hebisz, Rafał Hebisz and Stephen A. Ingham for sharing and discussing their data with us. However, during the process it became apparent that the studies in question did not fulfil the eligibility criteria for the present study and were, therefore, not included.

Declarations

Registration and Protocol Not applicable.

Support No sources of funding were used to assist in the preparation of this article.

Competing Interests M.R., J.W., J.A., G.T., Ø.S., S.S., J.E.-L., L.F., L.F., I.M., S.G., D.J.R.-C., P.S., S.S.-P., T.S., R.T., and C.Z. declare that they have no conflicts of interest relevant to the content of this review.

Availability of Data, Code and Other Material All aggregate data generated or analyzed during this study are included in this published article (and its Supplementary Information files). The IPD for the studies included in the review may be available upon request to the corresponding authors of the respective articles.

Authors' Contributions Michael Rosenblat was the project lead, conceived and designed the study, conducted article screening, data extraction, risk of bias analysis, statistical analysis, drafted the manuscript, and incorporated revisions. Jennifer Watt participated in the statistical analysis and reviewed the study methodology critically for important intellectual content. Jem Arnold participated in data extraction, risk of bias analysis, drafted the manuscript, and made substantial contributions before and during the review process of the manuscript. Gunnar Treff participated in article screening, drafted the manuscript, and made substantial contributions before and during the review process of the manuscript. Øyvind Sandbakk drafted the manuscript and incorporated revisions. Stephen Seiler participated in study design and drafted the manuscript. Jonathan Esteve-Lanao, Luca Festa, Luca Filipas, Iker Muñoz, Stuart Galloway, Domingo J. Ramos-Campo, Patrick Schneeweiss, Sergio Sellés-Pérez, Thomas Stöggl, Rune Talsnes, and Christoph Zinner all participated in data extraction, data collection and the aggregation of individual participant intervention data of the respective studies, and reviewed the manuscript critically for important intellectual content. All authors read and approved the final version of the manuscript.

References

1. Fiskerstrand Å, Seiler KS. Training and performance characteristics among Norwegian International Rowers 1970–2001. Scand J Med Sci Sports. 2004;14(5):303–10.
2. Seiler KS, Kjerland GØ. Quantifying training intensity distribution in elite endurance athletes: is there evidence for an "optimal" distribution? Scand J Med Sci Sports. 2006;16(1):49–56.
3. Iannetta D, Inglis EC, Mattu AT, Fontana FY, Pogliaghi S, Keir DA, et al. A critical evaluation of current methods for exercise prescription in women and men. Med Sci Sports Exerc. 2020;52(2):466–73.
4. Jamnick NA, Pettitt RW, Granata C, Pyne DB, Bishop DJ. An examination and critique of current methods to determine exercise intensity. Sports Med. 2020;50(10):1729–56.
5. Poole DC, Jones AM. Oxygen uptake kinetics. Compr Physiol. 2012;2(2):933–96.
6. Sperlich B, Matzka M, Holmberg H-C. The proportional distribution of training by elite endurance athletes at different intensities during different phases of the season. Front Sports Act Living. 2023;5:1258585.
7. Esteve-Lanao J, Foster C, Seiler S, Lucia A. Impact of training intensity distribution on performance in endurance athletes. J Strength Cond Res. 2007;21(3):943–9.
8. Muñoz I, Seiler S, Bautista J, Espana J, Larumbe E, Esteve-Lanao J. Does polarized training improve performance in recreational runners? Int J Sports Physiol Perform. 2014;9(2):265–72.
9. Neal CM, Hunter AM, Brennan L, O'Sullivan A, Hamilton DL, De Vito G, et al. Six weeks of a polarized training-intensity distribution leads to greater physiological and performance adaptations than a threshold model in trained cyclists. J Appl Physiol. 2013;114(4):461–71.
10. Stöggl T, Sperlich B. Polarized training has greater impact on key endurance variables than threshold, high intensity, or high volume training. Front Physiol. 2014;5:33.
11. Stöggl TL, Sperlich B. The training intensity distribution among well-trained and elite endurance athletes. Front Physiol. 2015;6:295.
12. Cohn LD, Becker BJ. How meta-analysis increases statistical power. Psychol Methods. 2003;8(3):243–53.
13. Rosenblat MA, Perrotta AS, Vicenzino B. Polarized vs. threshold training intensity distribution on endurance sport performance: a systematic review and meta-analysis of randomized controlled trials. J Strength Cond Res. 2019;33(12):3491–500.
14. Enoksen E, Tjelta AR, Tjelta LI. Distribution of training volume and intensity of elite male and female track and marathon runners. Int J Sports Sci Coach. 2011;6(2):273–94.
15. Muñoz I, Cejuela R, Seiler S, Larumbe E, Esteve-Lanao J. Training-intensity distribution during an ironman season: relationship with competition performance. Int J Sports Physiol Perform. 2014;9(2):332–9.
16. Neal CM, Hunter AM, Galloway SD. A 6-month analysis of training-intensity distribution and physiological adaptation in Ironman triathletes. J Sports Sci. 2011;29(14):1515–23.
17. Tjelta LI, Enoksen E. Training characteristics of male junior cross country and track runners on European top level. Int J Sports Sci Coach. 2010;5(2):193–203.
18. Billat VL, Demarle A, Slawinski J, Paiva M, Koralsztein JP. Physical and training characteristics of top-class marathon runners. Med Sci Sports Exerc. 2001;33(12):2089–97.
19. Kenneally M, Casado A, Gomez-Ezeiza J, Santos-Concejero J. Training intensity distribution analysis by race pace vs. physiological approach in world-class middle- and long-distance runners. Eur J Sport Sci. 2021;21(6):819–26.
20. van Erp T, Sanders D, de Koning JJ. Training characteristics of male and female professional road cyclists: a 4-year retrospective analysis. Int J Sports Physiol Perform. 2020;15(4):534–40.
21. Ieno C, Baldassarre R, Pennacchi M, La Torre A, Bonifazi M, Piacentini MF. Monitoring rating of perceived exertion time in zone: a novel method to quantify training load in elite open-water swimmers? Int J Sports Physiol Perform. 2021;16(10):1551–5.
22. Sylta Ø, Tønnessen E, Seiler S. From heart-rate data to training quantification: a comparison of 3 methods of training-intensity analysis. Int J Sports Physiol Perform. 2014;9(1):100–7.
23. Silva Oliveira P, Boppre G, Fonseca H. Comparison of polarized versus other types of endurance training intensity distribution on athlete' endurance performance: a systematic review with meta-analysis. Sports Med. 2024;54(8):2071–95.
24. Watt J, Tricco AC, Straus S, Veroniki AA, Naglie G, Drucker AM. Research techniques made simple: network meta-analysis. J Invest Dermatol. 2019;139(1):4–12.
25. Treff G, Winkert K, Sareban M, Steinacker JM, Becker M, Sperlich B. Eleven-week preparation involving polarized intensity

distribution is not superior to pyramidal distribution in national elite rowers. Front Physiol. 2017;8:515.
26. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. Int J Epidemiol. 2005;34(1):215–20.
27. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI, Anti-Lymphocyte Antibody Induction Therapy Study Group. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. Stat Med. 2002;21(3):371–87.
28. Riley RD, Debray TPA, Fisher D, Hattle M, Marlin N, Hoogland J, et al. Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning. Stat Med. 2020;39(15):2115–37.
29. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med. 2015;162(11):777–84.
30. Rosenblat MA, Lin E, da Costa BR, Thomas SG. Programming interval training to optimize time-trial performance: a systematic review and meta-analysis. Sports Med. 2021;51(8):1687–714.
31. Rosenblat MA, Arnold J, Nelson H, Watt J, Seiler S. The additional effect of training above the maximal metabolic steady state on VO2peak, Wpeak and time-trial performance in endurance-trained athletes: a systematic review, meta-analysis, and reality check. Sports Med. 2024;54(2):429–46.
32. McKay AK, Stellingwerff T, Smith ES, Martin DT, Mujika I, Goosey-Tolfrey VL, et al. Defining training and performance caliber: a participant classification framework. Int J Sports Physiol Perform. 2021;17(2):317–31.
33. Treff G, Winkert K, Sareban M, Steinacker JM, Sperlich B. The polarization-index: a simple calculation to distinguish polarized from non-polarized training intensity distributions. Front Physiol. 2019;10:707.
34. Sterne JA, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366: l4898.
35. Hedges L, Olkin I. Statistical methods for meta-analysis. New York: Academic Press; 1981.
36. R Core Team. R: A Language and Environment for Statistical Computing. 4.3.3 ed. Vienna, Austria: R Foundation for Statistical Computing; 2024.
37. Balduzzi S, Rücker G, Nikolakopoulou A, Papakonstantinou T, Salanti G, Efthimiou O, et al. netmeta: An R package for network meta-analysis using frequentist methods. J Stat Softw. 2023;106(2):1–40.
38. Riley RD, Tierney JF, Stewart LA. Individual participant data meta-analysis: a handbook for healthcare research. Chichester: Wiley; 2021.
39. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–88.
40. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539–58.
41. Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Methods. 2012;3(2):98–110.
42. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking Consistency in Mixed Treatment Comparison Meta-Analysis. Stat Med. 2010;29(7–8):932–44.
43. Chaimani A, Salanti G. Using network meta-analysis to evaluate the existence of small-study effects in a network of interventions. Res Synth Methods. 2012;3(2):161–76.
44. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010;36(3):1–48.
45. Zinner C, Schafer Olstad D, Sperlich B. Mesocycles with different training intensity distribution in recreational runners. Med Sci Sports Exerc. 2018;50(8):1641–8.
46. Schneeweiss P, Schellhorn P, Haigis D, Niess AM, Martus P, Krauss I. Effect of two different training interventions on cycling performance in mountain bike cross-country Olympic athletes. Sports. 2022;10(4):53.
47. Festa L, Tarperi C, Skroce K, La Torre A, Schena F. Effects of different training intensity distribution in recreational runners. Front Sports Act Living. 2019;1:70.
48. Filipas L, Bonato M, Gallo G, Codella R. Effects of 16 weeks of pyramidal and polarized training intensity distributions in well-trained endurance runners. Scand J Med Sci Sports. 2022;32(3):498–511.
49. Filipas L, Bonato M, Maggio A, Gallo G, Codella R. Effects of plyometric training on different 8-week training intensity distributions in well-trained endurance runners. Scand J Med Sci Sports. 2023;33(3):200–12.
50. Pérez A, Ramos-Campo DJ, Freitas TT, Rubio-Arias JÁ, Marín-Cascales E, Alcaraz PE. Effect of two different intensity distribution training programmes on aerobic and body composition variables in ultra-endurance runners. Eur J Sports Sci. 2019;19(5):636–44.
51. Sellés-Pérez S, Fernández-Sáez J, Cejuela R. Polarized and pyramidal training intensity distribution: relationship with a half-ironman distance triathlon competition. J Sports Sci Med. 2019;18(4):708–15.
52. Stöggl TL, Björklund G. High intensity interval training leads to greater improvements in acute heart rate recovery and anaerobic power as high volume low intensity training. Front Physiol. 2017;8:562.
53. Talsnes RK, Engdahl LJ, Sandbakk Ø. How do the effects of an 8-week intervention influence subsequent performance development in cross-country skiers? Int J Sports Physiol Perform. 2022;17(4):594–604.
54. Joyner MJ, Coyle EF. Endurance exercise performance: the physiology of champions. J Physiol. 2008;586(1):35–44.
55. Christensen PM, Andreasen JJ, Lyngholm J, Søgaard O, Lykkestrup J, Hostrup M, et al. Importance of training volume during intensified training in elite cyclists: maintained vs. reduced volume at moderate intensity. Scand J Med Sci Sports. 2024;34(1): e14362.
56. Legaz Arrese A, Serrano Ostáriz E, Jcasajús Mallén JA, Munguía ID. The changes in running performance and maximal oxygen uptake after long-term training in elite athletes. J Sports Med Phys Fitness. 2005;45(4):435–40.
57. Losnegard T, Myklebust H, Spencer M, Hallén J. Seasonal variations in VO2max, O2-cost, O2-deficit, and performance in elite cross-country skiers. J Strength Cond Res. 2013;27(7):1780–90.
58. Ramsbottom R, Williams C, Fleming N, Nute ML. Training induced physiological and metabolic changes associated with improvements in running performance. Br J Sports Med. 1989;23(3):171–6.
59. Batterson PM, Norton MR, Hetz SE, Rohilla S, Lindsay KG, Subudhi AW, et al. Improving biologic predictors of cycling endurance performance with near-infrared spectroscopy derived measures of skeletal muscle respiration: E pluribus unum. Physiol Rep. 2020;8(2): e14342.
60. Jacobs RA, Rasmussen P, Siebenmann C, Diaz V, Gassmann M, Pesta D, et al. Determinants of time trial performance and maximal incremental exercise in highly trained endurance athletes. J Appl Physiol. 2011;111(5):1422–30.
61. Rosenblat MA, Granata C, Thomas SG. Effect of interval training on the factors influencing maximal oxygen

consumption: a systematic review and meta-analysis. Sports Med. 2022;52(6):1329–52.
62. Amann M, Subudhi AW, Walker J, Eisenman P, Shultz B, Foster C. An evaluation of the predictive validity and reliability of ventilatory threshold. Med Sci Sports Exerc. 2004;36(10):1716–22.
63. Van Hooren B, Souren T, Bongers BC. Accuracy of respiratory gas variables, substrate, and energy use from 15 CPET systems during simulated and human exercise. Scand J Med Sci Sports. 2023;34(1): e14490.
64. Weston SB, Gabbett TJ. Reproducibility of ventilation of thresholds in trained cyclists during ramp cycle exercise. J Sci Med Sport. 2001;4(3):357–66.
65. Yogev A, Arnold J, Nelson H, Clarke DC, Guenette JA, Sporer BC, et al. Comparing the reliability of muscle oxygen saturation with common performance and physiological markers across cycling exercise intensity. Front Sports Act Living. 2023;5:1143393.
66. Lin L. Bias caused by sampling error in meta-analysis with small sample sizes. PLoS ONE. 2018;13(9): e0204056.
67. Bossi AH, Cole D, Passfield L, Hopker J. Conventional methods to prescribe exercise intensity are ineffective for exhaustive interval training. Eur J Appl Physiol. 2023;123(8):1655–70.
68. Meyler S, Bottoms L, Wellsted D, Muniz-Pumares D. Variability in exercise tolerance and physiological responses to exercise prescribed relative to physiological thresholds and to maximum oxygen uptake. Exp Physiol. 2023;108(4):581–94.
69. Greco CC, Carita RA, Dekerle J, Denadai BS. Effect of aerobic training status on both maximal lactate steady state and critical power. Appl Physiol Nutr Metab. 2012;37(4):736–43.
70. Galan-Rioja MA, Gonzalez-Mohino F, Poole DC, Gonzalez-Rave JM. Relative proximity of critical power and metabolic/ventilatory thresholds: systematic review and meta-analysis. Sports Med. 2020;50(10):1771–83.
71. Schmidt FL, Oh IS, Hayes TL. Fixed- versus random-effects models in meta-analysis: model properties and an empirical comparison of differences in results. Br J Math Stat Psychol. 2009;62(Pt 1):97–128.
72. da Costa BR, Juni P. Systematic reviews and meta-analyses of randomized trials: principles and pitfalls. Eur Heart J. 2014;35(47):3336–45.
73. Jamnick NA, Botella J, Pyne DB, Bishop DJ. Manipulating graded exercise test variables affects the validity of the lactate threshold and VO2peak. PLoS One. 2018;13(7):e0199794.
74. Pallarés JG, Morán-Navarro R, Ortega JF, Fernández-Elías VE, Mora-Rodriguez R. Validity and reliability of ventilatory and blood lactate thresholds in well-trained cyclists. PLoS One. 2016;11(9):e0163389.
75. Smith TB, Hopkins WG. Variability and predictability of finals times of elite rowers. Med Sci Sports Exerc. 2011;43(11):2155–60.
76. Laursen PB, Shing CM, Jenkins DG. Reproducibility of a laboratory-based 40-km cycle time-trial on a stationary wind-trainer in highly trained cyclists. Int J Sports Med. 2003;24(7):481–5.

Authors and Affiliations

Michael A. Rosenblat[1,2] · Jennifer A. Watt[3,4,5] · Jem I. Arnold[6] · Gunnar Treff[7] · Øyvind B. Sandbakk[8] · Jonathan Esteve-Lanao[9] · Luca Festa[10] · Luca Filipas[11,12] · Stuart D. Galloway[13] · Iker Muñoz[14] · Domingo J. Ramos-Campo[15] · Patrick Schneeweiss[16,17] · Sergio Sellés-Pérez[18] · Thomas Stöggl[19] · Rune K. Talsnes[8] · Christoph Zinner[20] · Stephen Seiler[21]

✉ Michael A. Rosenblat
michael@evidencebasedcoaching.ca
http://www.evidencebasedcoaching.ca

1 Sylvan Adams Sport Institute, Tel Aviv University, Tel Aviv, Israel

2 Evidence-Based Coaching, Toronto, ON, Canada

3 Knowledge Translation Program, Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, ON, Canada

4 Division of Geriatric Medicine, Department of Medicine, University of Toronto, Toronto, ON, Canada

5 Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

6 School of Kinesiology, University of British Columbia, Vancouver, BC, Canada

7 Institute for Molecular Sports and Rehabilitation Medicine, Paracelsus Medical University, Salzburg, Austria

8 Department of Neuromedicine and Movement Science, Centre for Elite Sports Research, Norwegian University of Science and Technology, 7491 Trondheim, Norway

9 All In Your Mind Training System, Mérida, Yucatán, Mexico

10 Italian Cycling Federation FCI, Rome, Italy

11 Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milan, Italy

12 TotalEnergies Pro Cycling Team, Essarts-en-Bocage, France

13 Physiology, Exercise and Nutrition Research Group, Faculty of Health Sciences and Sport, University of Stirling, Stirling, Scotland, UK

14 Department of Physical Activity and Sports, Faculty of Education and Sports, University of Deusto, Bilbao, Spain

15 LFE Research Group, Department of Health and Human Performance, Faculty of Physical Activity and Sport Science-INEF, Universidad Politécnica de Madrid, Madrid, Spain

16 Medical Clinic, Department of Sports Medicine, University of Tübingen, 72076 Tübingen, Germany

17 Interfaculty Research Institute for Sports and Physical Activity, University of Tübingen, 72074 Tübingen, Germany

18 General Didactic and Specific Didactics Department, Faculty of Education, University of Alicante, Alicante, Spain

19 Red Bull Athlete Performance Center, Salzburg, Austria

[20] Department of Sport, University of Applied Sciences for Police and Administration of Hesse, Wiesbaden, Germany

[21] Department of Sport Science and Physical Education, Faculty of Health and Sport Sciences, University of Agder, Kristiansand, Norway