Additional file 1: Supplementary Information

Anaerobic Speed and Power Reserve for Individualizing High Intensity Training in Trained Athletes: A Systematic Review

Journal: Sports Medicine Open

Suyan Xu, Shiwu Li, Jianhao Lu, Yan Chen,Zhenxiang Guo, Guoyan Zhang, Dapeng Bao, Øyvind Sandbakk and Grégoire P. Millet

Corresponding author: Dapeng Bao

Dapeng Bao: China Institute of Sport and Health Science, Beijing Sport University, Beijing 100084, China; Key Laboratory for Performance Training and Recovery of General Administration of Sport, Beijing Sport University, Beijing 100084, China;

Email: baodp@outlook.com; Tel: +86-10-62989304

S1 Full electronic search strategies

Searches were conducted from database inception to the date shown below. No limits were applied for publication year, language, sex, age, or publication type. Eligibility restrictions were applied during study screening. Record counts refer to results retrieved before deduplication.

Table S1 Information sources and search yields

Database	Platform or interface	Search date	Records retrieved
PubMed	PubMed	6 August 2026	66
Embase	Embase.com	6 August 2026	87
Web of Science Core Collection	Web of Science	6 August 2026	132
Scopus	Publish or Perish using the Elsevier Scopus API	6 August 2026	94
SPORTDiscus with Full Text	EBSCOhost	7 August 2026	61
CENTRAL	Cochrane Library	7 August 2026	14

S1.1 PubMed

Database: PubMed

Platform: PubMed

Search date: 6 August 2026

Fields and limits: Title and abstract terms [tiab] plus the MeSH term High-Intensity Interval Training. No filters applied.

Records retrieved: 66

Search strategy

```
(
"anaerobic speed reserve"[tiab]
 OR "anaerobic speed reserves"[tiab]
 OR "relative anaerobic speed reserve"[tiab]
 OR "anaerobic power reserve"[tiab]
 OR "anaerobic power reserves"[tiab]
 OR "relative anaerobic power reserve"[tiab]
 OR "speed reserve ratio"[tiab]
 OR "speed reserve ratios"[tiab]
 OR "power reserve ratio"[tiab]
 OR "power reserve ratios"[tiab]
 OR "locomotor profile"[tiab]
 OR "locomotor profiles"[tiab]
 OR "locomotor profiling"[tiab]
 OR "reserve based prescription"[tiab]
 OR "reserve-based prescription"[tiab]
 OR (
  (
   "speed reserve"[tiab]
   OR "speed reserves"[tiab]
   OR "power reserve"[tiab]
   OR "power reserves"[tiab]
  )
  AND
  (
   anaerob*[tiab]
   OR aerobic[tiab]
   OR sprint*[tiab]
   OR supramaximal[tiab]
   OR "critical speed"[tiab]
   OR "critical power"[tiab]
  )
 )
)
AND
(
 "High-Intensity Interval Training"[Mesh]
 OR "high intensity interval training"[tiab]
 OR "high-intensity interval training"[tiab]
 OR "high intensity interval exercise"[tiab]
 OR "high-intensity interval exercise"[tiab]
 OR "high intensity intermittent training"[tiab]
 OR "high-intensity intermittent training"[tiab]
 OR "high intensity intermittent exercise"[tiab]
 OR "high-intensity intermittent exercise"[tiab]
 OR "sprint interval training"[tiab]
 OR "sprint interval exercise"[tiab]
 OR "repeated sprint training"[tiab]
 OR "repeated-sprint training"[tiab]
 OR "repeated sprint exercise"[tiab]
 OR "repeated-sprint exercise"[tiab]
 OR "supramaximal interval training"[tiab]
 OR "supramaximal interval exercise"[tiab]
 OR "interval training"[tiab]
 OR "interval exercise"[tiab]
 OR "intermittent training"[tiab]
 OR "intermittent exercise"[tiab]
 OR repeated sprint*[tiab]
 OR exercise toleran*[tiab]
 OR time to exhaustion[tiab]
 OR training adapt*[tiab]
 OR physiologic* respons*[tiab]
 OR perceptual respons*[tiab]
 OR interindividual variab*[tiab]
 OR inter-individual variab*[tiab]
 OR response variab*[tiab]
 OR response consisten*[tiab]
 OR homogeneous adapt*[tiab]
 OR homeostatic stress[tiab]
 OR predict*[tiab]
 OR moderat*[tiab]
)
```

S1.2 Embase

Database: Embase

Platform: Embase.com

Search date: 6 August 2026

Fields and limits: Title, abstract, and keyword fields [ti,ab,kw] plus Emtree explosion terms. No filters applied.

Records retrieved: 87

Search strategy

```
(
 'anaerobic speed reserve':ti,ab,kw
 OR 'anaerobic speed reserves':ti,ab,kw
 OR 'relative anaerobic speed reserve':ti,ab,kw
 OR 'anaerobic power reserve':ti,ab,kw
 OR 'anaerobic power reserves':ti,ab,kw
 OR 'relative anaerobic power reserve':ti,ab,kw
 OR 'speed reserve ratio':ti,ab,kw
 OR 'speed reserve ratios':ti,ab,kw
 OR 'power reserve ratio':ti,ab,kw
 OR 'power reserve ratios':ti,ab,kw
 OR 'locomotor profile':ti,ab,kw
 OR 'locomotor profiles':ti,ab,kw
 OR 'locomotor profiling':ti,ab,kw
 OR 'reserve based prescription':ti,ab,kw
 OR
 (
  (
   'speed reserve':ti,ab,kw
   OR 'speed reserves':ti,ab,kw
   OR 'power reserve':ti,ab,kw
   OR 'power reserves':ti,ab,kw
  )
  AND
  (
   anaerob*:ti,ab,kw
   OR aerobic:ti,ab,kw
   OR sprint*:ti,ab,kw
   OR supramaximal:ti,ab,kw
   OR 'critical speed':ti,ab,kw
   OR 'critical power':ti,ab,kw
  )
 )
)
AND
(
 'high intensity interval training'/exp
 OR 'anaerobic exercise'/exp
 OR 'exercise tolerance'/exp
 OR 'exercise training'/exp
 OR 'high intensity interval training':ti,ab,kw
 OR 'high intensity interval exercise':ti,ab,kw
 OR 'high intensity intermittent training':ti,ab,kw
 OR 'high intensity intermittent exercise':ti,ab,kw
 OR 'sprint interval training':ti,ab,kw
 OR 'sprint interval exercise':ti,ab,kw
 OR 'repeated sprint training':ti,ab,kw
 OR 'repeated sprint exercise':ti,ab,kw
 OR 'supramaximal interval training':ti,ab,kw
 OR 'supramaximal interval exercise':ti,ab,kw
 OR 'interval training':ti,ab,kw
 OR 'interval exercise':ti,ab,kw
 OR 'intermittent training':ti,ab,kw
 OR 'intermittent exercise':ti,ab,kw
 OR 'repeated sprint':ti,ab,kw
 OR 'repeated sprints':ti,ab,kw
 OR 'exercise tolerance':ti,ab,kw
 OR 'time to exhaustion':ti,ab,kw
 OR 'training adaptation':ti,ab,kw
```

```
(
 'anaerobic speed reserve':ti,ab,kw
 OR 'anaerobic speed reserves':ti,ab,kw
 OR 'relative anaerobic speed reserve':ti,ab,kw
 OR 'anaerobic power reserve':ti,ab,kw
 OR 'anaerobic power reserves':ti,ab,kw
 OR 'relative anaerobic power reserve':ti,ab,kw
 OR 'speed reserve ratio':ti,ab,kw
 OR 'speed reserve ratios':ti,ab,kw
 OR 'power reserve ratio':ti,ab,kw
 OR 'power reserve ratios':ti,ab,kw
 OR 'locomotor profile':ti,ab,kw
 OR 'locomotor profiles':ti,ab,kw
 OR 'locomotor profiling':ti,ab,kw
 OR 'reserve based prescription':ti,ab,kw
 OR
 (
  (
   'speed reserve':ti,ab,kw
   OR 'speed reserves':ti,ab,kw
   OR 'power reserve':ti,ab,kw
   OR 'power reserves':ti,ab,kw
  )
  AND
  (
   anaerob*:ti,ab,kw
 OR 'training adaptations':ti,ab,kw
 OR 'physiological response':ti,ab,kw
 OR 'physiological responses':ti,ab,kw
 OR 'perceptual response':ti,ab,kw
 OR 'perceptual responses':ti,ab,kw
 OR 'interindividual variability':ti,ab,kw
 OR 'inter individual variability':ti,ab,kw
 OR 'response variability':ti,ab,kw
 OR 'response consistency':ti,ab,kw
 OR 'homogeneous adaptation':ti,ab,kw
 OR 'homogeneous adaptations':ti,ab,kw
 OR 'homeostatic stress':ti,ab,kw
 OR predict*:ti,ab,kw
 OR moderat*:ti,ab,kw
)
```

S1.3 Web of Science Core Collection

Database: Web of Science Core Collection

Platform: Web of Science

Search date: 6 August 2026

Fields and limits: Topic field, all Core Collection editions available through the interface. No filters applied.

Records retrieved: 132

Note: The two Topic blocks were entered as separate rows in Fielded Search and combined with AND.

Topic block 1

```
(
 "anaerobic speed reserve"
 OR "anaerobic speed reserves"
 OR "relative anaerobic speed reserve"
 OR "anaerobic power reserve"
 OR "anaerobic power reserves"
 OR "relative anaerobic power reserve"
 OR "speed reserve ratio"
 OR "speed reserve ratios"
 OR "power reserve ratio"
 OR "power reserve ratios"
 OR "locomotor profile"
 OR "locomotor profiles"
 OR "locomotor profiling"
 OR "reserve based prescription"
 OR
 (
  (
   "speed reserve"
   OR "speed reserves"
   OR "power reserve"
   OR "power reserves"
  )
  AND
  (
   anaerob*
   OR aerobic
   OR sprint*
   OR supramaximal
   OR "critical speed"
   OR "critical power"
  )
 )
)
```

AND Topic block 2

```
(
 "high intensity interval training"
 OR "high intensity interval exercise"
 OR "high intensity intermittent training"
 OR "high intensity intermittent exercise"
 OR "sprint interval training"
 OR "sprint interval exercise"
 OR "repeated sprint training"
 OR "repeated sprint exercise"
 OR "supramaximal interval training"
 OR "supramaximal interval exercise"
 OR "interval training"
 OR "interval exercise"
 OR "intermittent training"
 OR "intermittent exercise"
 OR "repeated sprint"
 OR "repeated sprints"
 OR "exercise tolerance"
 OR "time to exhaustion"
 OR "training adaptation"
 OR "training adaptations"
 OR "physiological response"
 OR "physiological responses"
 OR "perceptual response"
 OR "perceptual responses"
 OR "interindividual variability"
 OR "response variability"
 OR "response consistency"
 OR "homogeneous adaptation"
 OR "homogeneous adaptations"
 OR "homeostatic stress"
 OR predict*
 OR moderat*
)
```

S1.4 Scopus

Database: Scopus

Platform: Publish or Perish using the Elsevier Scopus API

Search date: 6 August 2026

Fields and limits: Publish or Perish Scopus Query interface, Keywords field only; all other search fields left blank; no year restriction.

Records retrieved: 94

Note: A broad reserve construct query was used for Scopus to maximize sensitivity in the Publish or Perish interface. Eligibility criteria were applied during screening.

Keywords query

```
"anaerobic speed reserve" | "anaerobic speed reserves" | "relative anaerobic speed reserve" | "anaerobic power reserve" | "anaerobic power reserves" | "relative anaerobic power reserve" | "speed reserve ratio" | "speed reserve ratios" | "power reserve ratio" | "power reserve ratios"
```

S1.5 SPORTDiscus with Full Text

Database: SPORTDiscus with Full Text

Platform: EBSCOhost

Search date: 7 August 2026

Fields and limits: XB Title and Abstract field for both search rows. No filters applied.

Records retrieved: 61

Note: Only records returned from SPORTDiscus with Full Text were retained for the formal SPORTDiscus search set.

S1

```
"anaerobic speed reserve"
OR "anaerobic speed reserves"
OR "relative anaerobic speed reserve"
OR "anaerobic power reserve"
OR "anaerobic power reserves"
OR "relative anaerobic power reserve"
OR "speed reserve ratio"
OR "speed reserve ratios"
OR "power reserve ratio"
OR "power reserve ratios"
OR "locomotor profile"
OR "locomotor profiles"
OR "locomotor profiling"
OR "reserve based prescription"
OR
(
 (
  speed N1 reserve*
  OR power N1 reserve*
 )
 AND
 (
  anaerob*
  OR aerobic
  OR sprint*
  OR supramaximal
  OR "critical speed"
  OR "critical power"
 )
)
```

S2

```
"high intensity interval training"
OR "high intensity interval exercise"
OR "high intensity intermittent training"
OR "high intensity intermittent exercise"
OR "sprint interval training"
OR "sprint interval exercise"
OR "repeated sprint training"
OR "repeated sprint exercise"
OR "supramaximal interval training"
OR "supramaximal interval exercise"
OR "interval training"
OR "interval exercise"
OR "intermittent training"
OR "intermittent exercise"
OR repeated N1 sprint*
OR "exercise tolerance"
OR "time to exhaustion"
OR training N1 adapt*
OR physiologic* N1 respons*
OR perceptual N1 respons*
OR interindividual N1 variab*
OR response N1 variab*
OR response N1 consisten*
OR homogeneous N1 adapt*
OR "homeostatic stress"
OR predict*
OR moderat*
```

Combination

```
S1 AND S2
```

S1.6 Cochrane Central Register of Controlled Trials

Database: Cochrane Central Register of Controlled Trials (CENTRAL)

Platform: Cochrane Library Search Manager

Search date: 7 August 2026

Fields and limits: Title, abstract, and keyword fields [ti,ab,kw]. No additional limits applied.

Records retrieved: 14

#1

```
("anaerobic speed reserve" OR "anaerobic speed reserves" OR "relative anaerobic speed reserve" OR "anaerobic power reserve" OR
"anaerobic power reserves" OR "relative anaerobic power reserve" OR "speed reserve ratio" OR "speed reserve ratios" OR "power reserve
ratio" OR "power reserve ratios" OR "locomotor profile" OR "locomotor profiles" OR "locomotor profiling" OR "reserve based
prescription" OR (((speed near/1 reserve*) OR (power near/1 reserve*)) AND (anaerob* OR aerobic OR sprint* OR supramaximal OR
"critical speed" OR "critical power"))):ti,ab,kw
```

#2

```
("high intensity interval training" OR "high intensity interval exercise" OR "high intensity intermittent training" OR "high intensity
intermittent exercise" OR "sprint interval training" OR "sprint interval exercise" OR "repeated sprint training" OR "repeated sprint
exercise" OR "supramaximal interval training" OR "supramaximal interval exercise" OR "interval training" OR "interval exercise" OR
"intermittent training" OR "intermittent exercise" OR (repeated near/1 sprint*) OR "exercise tolerance" OR "time to exhaustion" OR
(training near/1 adapt*) OR (physiologic* near/1 respons*) OR (perceptual near/1 respons*) OR (interindividual near/1 variab*) OR (response
near/1 variab*) OR (response near/1 consisten*) OR (homogeneous near/1 adapt*) OR "homeostatic stress" OR predict* OR
moderat*):ti,ab,kw
```

#3

```
#1 AND #2
```

Abbreviations. CENTRAL, Cochrane Central Register of Controlled Trials; MeSH, Medical Subject Headings; N1, terms within one word of each other in EBSCOhost; near/1, terms within one word of each other in Cochrane Search Manager. Supplementary completeness searching was completed on 12 August 2026 by screening the reference list of Sandford et al. (2021) [1], a key ASR/APR review and framework paper. Fourteen unique records that were not present in the original database set were identified through this reference-list screening. Because several early reserve studies used terminology that was not consistently represented in titles or abstracts, targeted legacy terms including AnSR, anaerobic scope, peak anaerobic scope, mean anaerobic scope, and power reserve were used to verify candidate terminology and eligibility. These terminology checks did not identify any additional unique records beyond the 14 obtained from the Sandford et al. (2021) reference list [1]. Eleven of the 14 records were excluded during citation screening and three reports were assessed in full text. Buchheit et al. (2012) [2] was included; Blondel et al. (2001) [3] and Barnett et al. (1996) [4] were excluded because the trained athlete population criterion could not be established. Barnett et al [4]. reported male students of mixed athletic ability without a separable trained subgroup. The supplementary reference-list screening was introduced after initial screening and did not change the prespecified eligibility criteria or outcome hierarchy.

Table S1A Post-extraction eligibility adjudication for Heaney et al. (2014) [5]

Issue	Source-report finding and review adjudication
Assessment timing	The Methods state that YOYO IR1 testing was completed preintervention and mid-intervention and that a postintervention YOYO IR1 test could not be completed because of competition and scheduling constraints. Table 4 is titled “Pre & Post HIIT intervention” and the accompanying Results text describes the later test as postintervention. The review-relevant endpoint could therefore not be assigned reliably.
Intervention exposure	The Methods describe eight HIIT sessions over four weeks. The Results state that the intervention consisted of seven HIIT sessions when interpreting improvement per session. The completed intervention dose could not be reconciled.
Analytic denominator	Sixteen athletes are described in the participant section. Table 4 displays 15 baseline YOYO IR1 distances because participant 4 has no baseline value, but 16 values at the later assessment. A paired analytic denominator is not reported.
Percentage change	The displayed group means of 1525 m and 1720 m imply a 12.8% increase, while the mean of the 15 displayed participant-level percentage changes is 12.0%. Table 4 reports a mean percentage difference of 14.6%. The reported summary could not be reproduced from the displayed data.
Effect size	Table 4 reports a mean effect size of 0.8, but the calculation method is not defined and the 15 displayed participant-level effect sizes average approximately 0.65. A reproducible review-level effect-size input could not be identified.
Adjudication	The report was excluded after data extraction because intervention timing, completed exposure, analytic denominators, and quantitative summaries could not be reconciled with sufficient reliability for synthesis. This was recorded as a review-process deviation because report-level internal

Issue	Source-report finding and review adjudication
	consistency had not been prespecified as a standalone eligibility criterion.

Note. Values were transcribed from the published report. No missing participant value was imputed and no preferred interpretation was selected when report sections were internally inconsistent.

S2 Risk of bias assessment

Risk of bias was assessed at the level of the review-relevant result or evidence contribution. Randomized parallel trials were assessed with RoB 2 [6], randomized crossover trials with the RoB 2 crossover extension , the nonrandomized intervention study with ROBINS-I 2016[7], and predictor, profile, and longitudinal exposure studies with QUIPS [8]. One reviewer completed the initial domain-level assessments and supporting rationales, and a second reviewer checked the judgements against the source reports. Disagreements were resolved by discussion to consensus. QUIPS domain judgements were retained without calculating a formal overall score. The label Higher concern was used only as an operational synthesis flag for QUIPS-assessed evidence and was not treated as an official QUIPS overall category. Companion reports arising from the same cohort were linked for overlapping evidence contributions so the same participants were not treated as independent risk-of-bias units. Girard and Racinais (2023) [9] was therefore linked to Soo et al. (2021) [10] for overlapping review-relevant contributions rather than assigned a separate overall study-level judgement.

Table S2 Consensus risk-of-bias judgements and synthesis implications

Study	Tool	Formal overall judgement	Operational concern class	Key domain concern	Synthesis implication
Thron et al., 2025 [11]	QUIPS	No formal overall score	Higher concern	Confounding High; Statistical analysis/reporting High	Retain structured evidence; downgrade certainty
Buchheit et al., 2012 [2]	QUIPS	No formal overall score	Higher concern	Confounding High; Statistical analysis/reporting High	Retain structured evidence; downgrade certainty
Di Gennaro et al., 2025 JSM [12]	RoB 2 crossover	Some concerns	Intermediate	D1 Some concerns; D5 Some concerns	Retain
Peng et al., 2025 [13]	RoB 2	Some concerns	Intermediate	D5 Some concerns	Retain in main model
Wang and Ye, 2024 [14]	RoB 2	Some concerns	Intermediate	D1 Some concerns; D5 Some concerns	Retain in main model
Zhu et al., 2025 [15]	RoB 2	Some concerns	Intermediate	D1 Some concerns; D5 Some concerns	Retain narrative decision evidence
Luo et al., 2024 [16]	RoB 2	Some concerns	Intermediate	D1 Some concerns; D2 Some concerns; D5 Some concerns	Retain; formula-definition sensitivity
Dai and Xie, 2023 [17]	RoB 2	Some concerns	Intermediate	D1 Some concerns; D5 Some concerns	Retain in main model
Yu Tongwu et al., 2025 [18]	RoB 2	High	Higher concern	D5 High; D1/D2 Some concerns	Exclude from primary model; sensitivity only
Du and Tao, 2023 [19]	RoB 2	Some concerns	Intermediate	D1 Some concerns; D5 Some concerns	Supportive only
Aspin et al., 2024 [20]	QUIPS	No formal overall score	Higher concern	Attrition High; Statistical analysis/reporting High	Retain within-study only; emphasize attrition
Julio et al., 2020 [21]	RoB 2 crossover	Some concerns	Intermediate	D3 Some concerns; D5 Some concerns	Retain paired author data
Wang and Zhao, 2023 [22]	RoB 2	Some concerns	Intermediate	D3 Some concerns; D5 Some concerns	Retain n=10 per arm; reporting discrepancy note
Soo et al., 2021 [10]	RoB 2 crossover	Some concerns	Intermediate	D5 Some concerns	Retain environmental context; Girard and Racinais (2023) is a companion report from the same study and shares this assessment
Collison et al., 2022 [23]	RoB 2 crossover	Some concerns	Intermediate	D1 Some concerns; D5 Some concerns	Retain
Di Gennaro et al., 2025 APNM [24]	QUIPS	No formal overall score	Higher concern	Statistical analysis/reporting High; Confounding Moderate	Retain recalculated partial associations with caution
Salter et al., 2025 [25]	QUIPS	No formal overall score	Higher concern	Confounding High; Statistical analysis/reporting High	Narrative only
Thron et al., 2026	ROBINS-I	Serious	Higher concern	Confounding Serious; Missing data	De-emphasize; no primary pooled inference

Study	Tool	Formal overall judgement	Operational concern class	Key domain concern	Synthesis implication
[26]	2016			Serious	
Belechris et al., 2025 [27]	QUIPS	No formal overall score	Higher concern	Confounding High; Statistical analysis/reporting High	Supportive only
Girard et al., 2021 [28]	RoB 2 crossover	Some concerns	Intermediate	D5 Some concerns	Retain environmental context as a separate cohort
Fitzpatrick et al., 2018 [29]	QUIPS	No formal overall score	Higher concern	Confounding High; Statistical analysis/reporting High	Within-study only
Franchini et al., 2022 [30]	QUIPS	No formal overall score	Higher concern	Confounding High; Statistical analysis/reporting High	Narrative only
Kavanagh et al., 2025 [31]	QUIPS	No formal overall score	Higher concern	Participation High; Attrition High; Confounding High; Statistical analysis/reporting High	Exploratory evidence only

Note. Higher concern is an operational synthesis label used to guide structured synthesis of QUIPS-assessed evidence. It is not an official QUIPS overall risk-of-bias category and is not equivalent to High risk in RoB 2 or Serious risk in ROBINS-I.

A. Randomized studies assessed with RoB 2 and RoB 2 crossover

Study	D1 Randomization	D3 Period/ carryover	D2 Deviations	D3 Missing outcome data	D4 Outcome measurement	D5 Selection of reported result	Overall
Di Gennaro et al. [12]	SC	L	L	L	L	SC	SC
Peng et al. [13]	L	NA	L	L	L	SC	SC
Wang and Ye [14]	SC	NA	L	L	L	SC	SC
Zhu et al. [15]	SC	NA	L	L	L	SC	SC
Luo et al. [16]	SC	NA	SC	L	L	SC	SC
Dai and Xie [17]	SC	NA	L	L	L	SC	SC
Yu et al. [18]	SC	NA	SC	L	L	H	H
Du and Tao [19]	SC	NA	L	L	L	SC	SC
Julio et al. [21]	SC	L	L	SC	L	SC	SC
Wang and Zhao [22]	SC	NA	L	SC	L	SC	SC
Soo et al. [10]	L	L	L	L	L	SC	SC
Collison et al. [23]	SC	L	L	L	L	SC	SC
Girard et al. [28]	L	L	L	L	L	SC	SC

B. Nonrandomized intervention study assessed with ROBINS-I 2016

Study	D1 Confounding	D2 Selection	D3 Classification	D4 Deviations	D5 Missing data	D6 Outcome measurement	D7 Selection of reported result	Overall
Thron et al. [26]	S	M	L	M	S	L	M	S

C. Predictor, profile, and longitudinal exposure studies assessed with QUIPS

Study	Participation	Attrition	Predictor measurement	Outcome measurement	Confounding	Statistical analysis and reporting
Thron et al. [11]	L	L	L	M	H	H
Aspin et al. [20]	M	H	L	M	M	H
Di Gennaro et al. [24]	L	L	L	L	M	H
Salter et al. [25]	M	L	M	L	H	H
Belechris et al. [27]	M	L	M	M	H	H
Buchheit et al. [2]	M	L	L	L	H	H
Fitzpatrick et al. [29]	M	L	L	M	H	H
Franchini et al. [30]	M	L	L	L	H	H
Kavanagh et al. [31]	H	H	L	M	H	H

L Low SC/M Some concerns or Moderate H/S High or Serious NA Not applicable

Supplementary Fig. S1 Domain-level risk-of-bias assessment.

(a) Domain-level consensus judgements for randomized studies assessed with RoB 2 or the RoB 2 crossover extension. (b) Domain-level consensus judgements for the nonrandomized intervention study assessed with ROBINS-I 2016. (c) Domain-level consensus judgements for predictor, profile, and longitudinal exposure studies assessed with QUIPS. QUIPS domain judgements are presented without a formal overall score. Abbreviations: H high; L low; M moderate; NA not applicable; S serious; SC some concerns.

S3 Supplementary statistical methods

S3.1 Chronic mean adaptation

Postintervention raw mean differences in relative VO_2max were used for the primary chronic mean model because all contributing studies reported the same scale and within-group preintervention-to-postintervention correlations were unavailable. This choice avoided imputing unreported correlations but does not remove the possibility of chance baseline imbalance in the small groups; change-score sensitivity analyses were therefore used to assess robustness. Positive mean differences favoured reserve-based prescription. Random-effects models used REML estimation of τ^2 and Wald 95% confidence intervals. Cochran Q, I^2, and τ^2 described heterogeneity [32]. Sensitivity analyses changed eligible comparators, added or excluded prespecified studies, repeated leave-one-out models, examined change scores across assumed pre-post correlations from 0.0 to 0.9, and applied alternative τ^2 estimators and interval methods.

S3.2 Chronic observed response dispersion

CVR was the primary dispersion effect [33, 34]. For reserve and comparator coefficients of variation CVR and CVC with sample sizes nR and nC, the bias-corrected effect was:

$$\ln\left(\frac{CV_R}{CV_C}\right)+\frac{1}{2(n_R-1)}-\frac{1}{2(n_C-1)}+0.5\left(\frac{CV_C^2}{n_C}-\frac{CV_R^2}{n_R}\right)$$

The primary large-sample sampling variance was:

$$\frac{1}{2(n_R-1)}+\frac{CV_R^2}{n_R}+\frac{1}{2(n_C-1)}+\frac{CV_C^2}{n_C}$$

Effects were synthesized on the log scale with REML and back-transformed to the ratio scale. CVR values below 1 indicate lower observed relative dispersion under reserve-based prescription. Sensitivity analyses used alternative comparators, study additions and exclusions, leave-one-out models, alternative interval and variance estimators, the uncorrected log CVR, and exhaustive combinations of plausible source-data rounding.

For the VR sensitivity analysis, the standard deviation of individual percentage change was recovered as:

$$\left|\overline{\Delta\%}\right|\times\frac{CV}{100}$$

Where the mean percentage change was unavailable, it was approximated from the group preintervention and postintervention means. The bias-corrected log variability ratio was:

$$\ln\left(\frac{SD_R}{SD_C}\right)+\frac{1}{2(n_R-1)}-\frac{1}{2(n_C-1)}$$

The corresponding sampling variance was:

$$\frac{1}{2(n_R-1)}+\frac{1}{2(n_C-1)}$$

The primary five-study model used REML and Wald 95% confidence intervals. Conventional and modified Hartung-Knapp intervals were examined as small-k sensitivity analyses without changing the effect definition or primary model [35]. A sensitivity model excluded the two studies in which mean percentage change had to be derived from group means. For input transparency, the locked dataset also records the reserve and comparator mean percentage changes that form the denominators of the source CVs, whether these means were reported directly or derived from group preintervention and postintervention means, and the recovered SDs used for the complementary VR analysis. Across the five primary comparisons, group mean percentage changes ranged from 4.0% to 7.2%, so none of the source CV denominators was close to zero. These transparency fields do not enter the CVR calculation. CVR and VR were interpreted as observed dispersion and not as measurement-error-adjusted true interindividual response variance. Because dispersion was based on individual percentage change, both metrics also inherit the statistical dependence of percentage change on baseline values. The VR sensitivity removes the response mean from the variability metric but does not remove baseline normalization or measurement error.

S3.3 Stratified and comparator robustness analyses

Exploratory strata separated ASR studies in running and team sports from APR studies in rowing. Reserve construct and exercise modality were completely confounded, so pooled stratum estimates and Q-between statistics were treated only as descriptive robustness checks. When more than one eligible conventional comparator shared the same reserve prescription group, the primary model selected the comparator that most directly represented the conventional aerobic lower boundary anchor in that exercise mode. Alternative eligible comparators were evaluated by substitution rather than simultaneous inclusion to avoid duplicate weighting of the shared reserve prescription group. In Wang and Ye (2024) [14] and Dai and Xie (2023) [17], the comparator sensitivity replaced MAS with VIFT because both conventional anchors were eligible.

S3.4 Chronic high intensity and sport specific performance

Performance outcomes were not combined into a single global effect because constructs, units, and exercise modes differed. Only the three rowing studies reporting 2000 m time trial performance were pooled as a common outcome. Between-group differences in change were calculated with positive values oriented to favour reserve-based prescription. The primary exploratory change-score model assumed r = 0.5 for the unknown pre-post correlation and sensitivity analyses used r = 0.0, 0.3, 0.5, 0.7, and 0.9. Performance-change dispersion used the same bias-corrected log CVR framework. Other performance families were summarized within outcome and were not promoted to confirmatory pooled evidence when only one or two comparable studies were available.

S3.5 Outcome-domain evidence mapping and clustering

Sixty-eight matched chronic outcome-level CVR observations were classified into five descriptive domains: maximal aerobic/cardiorespiratory, threshold/submaximal aerobic, performance, reserve/component capacity, and hormonal/other. Multiple outcomes from the same study were considered statistically dependent and were not meta-analyzed as independent effects. A median CVR was calculated for each study by domain combination. For visualization, CVR < 0.90 was labelled reserve lower dispersion, 0.90 to 1.10 similar, and CVR > 1.10 comparator lower dispersion; this band was descriptive and was not treated as an equivalence margin.

For heatmap ordering, study-domain median CVRs were $\log_2$ transformed. Studies and domains were clustered separately. Pairwise root-mean-square Euclidean distance was calculated using only cells observed in both members of a pair; missing cells were not imputed. When a pair had no overlapping observed domain, a distance 1.05 times the largest observed pairwise distance was assigned solely to permit complete hierarchical ordering. Agglomerative clustering used average linkage followed by optimal leaf ordering [36]. The dendrogram structure was used only to reorder the matrix and was not interpreted as inferential clustering or predictive modelling.

S3.6 Longitudinal reserve dynamics and illustrative prescription drift

Reserve dynamics were examined in chronic study arms that reported a complete lower-anchor, upper-anchor, and reserve triad at preintervention and postintervention. Percentage change for each component was calculated as:

$$100 \times \frac{X_{post} - X_{pre}}{X_{pre}}$$

Active training arms were summarized separately from the control arm. The analysis recorded whether lower-anchor improvement exceeded upper-anchor improvement. Because the reserve is derived from the two component anchors, reserve decrease was treated as the mathematical consequence of their relative changes rather than as an independent outcome signal.

For reserve-based arms with an explicit baseline target p, letting L denote the lower anchor, R the reserve, and W0 the baseline absolute target, the baseline absolute target was calculated as:

$$W_0 = L_{pre} + p \times R_{pre}$$

The posttraining equivalent position was then:

$$p_{post} = 100 \times \frac{W_0 - L_{post}}{R_{post}}$$

This calculation asks where the original absolute workload would lie within the updated reserve if it had been retained unchanged. It does not reconstruct the actual progression used by the investigators and cannot establish that reserve-based training caused reserve contraction.

S3.7 Acute direct prescription and environmental calculations

Acute direct studies were retained as within-study structured comparisons because dispersion metrics were not commensurate across studies. For Julio et al. (2020) [21], participant-level paired data were used for paired mean contrasts and paired CV ratios, with bias-corrected and accelerated bootstrap confidence intervals based on 50,000 resamples and a fixed random seed of 20260807 [37]. For Collison et al. (2022) [23]and Di Gennaro et al. (2025 JSM) [12], residual-dispersion or condition-level CV estimates were normalized to the study-specific conventional reference. Environmental coefficients of variation were calculated from reported means and standard deviations only to describe observed dispersion across oxygen conditions. They were not treated as evidence that APR standardized responses under hypoxia or as comparative evidence of incremental prescription value.

S3.8 Predictor, profile, moderation, and longitudinal analyses

Predictor and longitudinal effects were not meta-analyzed because the studies addressed different predictors, outcomes, session formats, and causal questions, and several studies contributed multiple dependent effects. Pearson correlations, partial correlations, profile contrasts, treatment-by-profile interactions, coefficients of determination, and Spearman correlations were therefore synthesized as separate evidence types. Because

reserve variables are deterministic functions of their component anchors, adjusted prediction was interpreted relative to the baseline model used. A reserve added to one component may carry information from the omitted component, while including both component anchors and their difference in the same unconstrained model would create exact linear dependence. Fisher r-to-z confidence intervals were calculated when needed for isolated Pearson correlations. For Di Gennaro et al. (2025 APNM) [24], published partial-correlation coefficients were retained while two-sided P values and 95% confidence intervals were recalculated uniformly using n = 19 and one controlled variable; the published inferential values were not used when internally inconsistent. Interpretation emphasized effect magnitude, intensity dependence, adjusted incremental value, design limitations, and risk of bias. For the Fig. 7 evidence landscape, signal categories were assigned at each incremental-value stage using explicit operational rules. Association or profile description was supportive when reserve or profile information consistently characterized response heterogeneity within the reported context. Normalization was supportive when reserve scaling reduced an apparent between-profile or between-athlete difference in relative demand. Adjusted prediction was supportive only when reserve information retained predictive value after a conventional component anchor was included in the same model; mixed or conditional was assigned when this occurred only at selected intensities, outcomes, or model specifications. Treatment moderation required a profile-by-training interaction, profile-guided decision required prospective allocation evidence, and longitudinal monitoring required prospective associations between individualized thresholds and subsequent adaptation. No clear or negative was assigned when the relevant stage showed no clear incremental value or evidence favoured the conventional anchor. Statistical significance alone did not determine classification. Cell-level rationales are retained in the Figure 7 source data and predictor-profile analysis package. Published and review-recalculated inference for the exercise-tolerance partial correlations is shown in Table S4B.

S4 Chronic primary synthesis sensitivity analyses

Exploratory stratification was used only as a robustness check because the five studies in the chronic primary models separated into two fully confounded strata: ASR studies in running and team sports and APR studies in rowing. The analysis therefore cannot distinguish reserve construct effects from exercise modality effects. Between-stratum tests are descriptive and do not establish equivalence. The comparator robustness scenario replaced MAS with VIFT in Wang and Ye (2024) [14] and Dai and Xie (2023) [17], the two ASR studies in which both conventional anchors were available.

Table S3 Exploratory stratified sensitivity analyses

Outcome and scenario	Stratum	k	Pooled estimate (95% CI)	Q between	*P*
VO_2max MD, primary comparator	ASR running and team sport	3	1.146 (−0.292 to 2.584)	0.086	.770
VO_2max MD, primary comparator	APR rowing	2	0.838 (−0.641 to 2.317)		
CVR, primary comparator	ASR running and team sport	3	0.315 (0.211 to 0.470)	0.006	.939
CVR, primary comparator	APR rowing	2	0.307 (0.191 to 0.493)		
CVR, VIFT comparator where eligible	ASR running and team sport	3	0.527 (0.356 to 0.781)	2.963	.085
CVR, VIFT comparator where eligible	APR rowing	2	0.307 (0.191 to 0.493)		

Note. ASR and APR are completely confounded with exercise modality in these strata, so Q between and its P value are descriptive robustness checks only. CVR values below 1 indicate lower observed relative dispersion under reserve based prescription. In the VIFT comparator scenario, the APR rowing estimates are unchanged. MD mean difference; APR anaerobic power reserve; ASR anaerobic speed reserve; CI confidence interval; CVR coefficient of variation ratio; MAS maximal aerobic speed; VIFT final velocity reached in the 30 15 Intermittent Fitness Test.

Table S4 Variability-ratio sensitivity analysis for chronic VO_2max observed response dispersion

Analysis	k	VR	95% CI	Q	df	P(Q)	I^2
Primary five-study VR analysis	5	0.38	0.29 to 0.51	2.02	4	0.732	0%
Exclude studies with derived mean percentage change	3	0.42	0.29 to 0.62	1.18	2	0.555	0%

Note. VR values below 1 indicate lower absolute observed dispersion of individual percentage change under reserve-based prescription. The exclusion sensitivity removes Wang and Ye (2024) and Luo et al. (2024), for which mean percentage change was derived from group preintervention and postintervention means. VR variability ratio.

Table S4A Input values underlying the primary chronic VO_2max dispersion synthesis

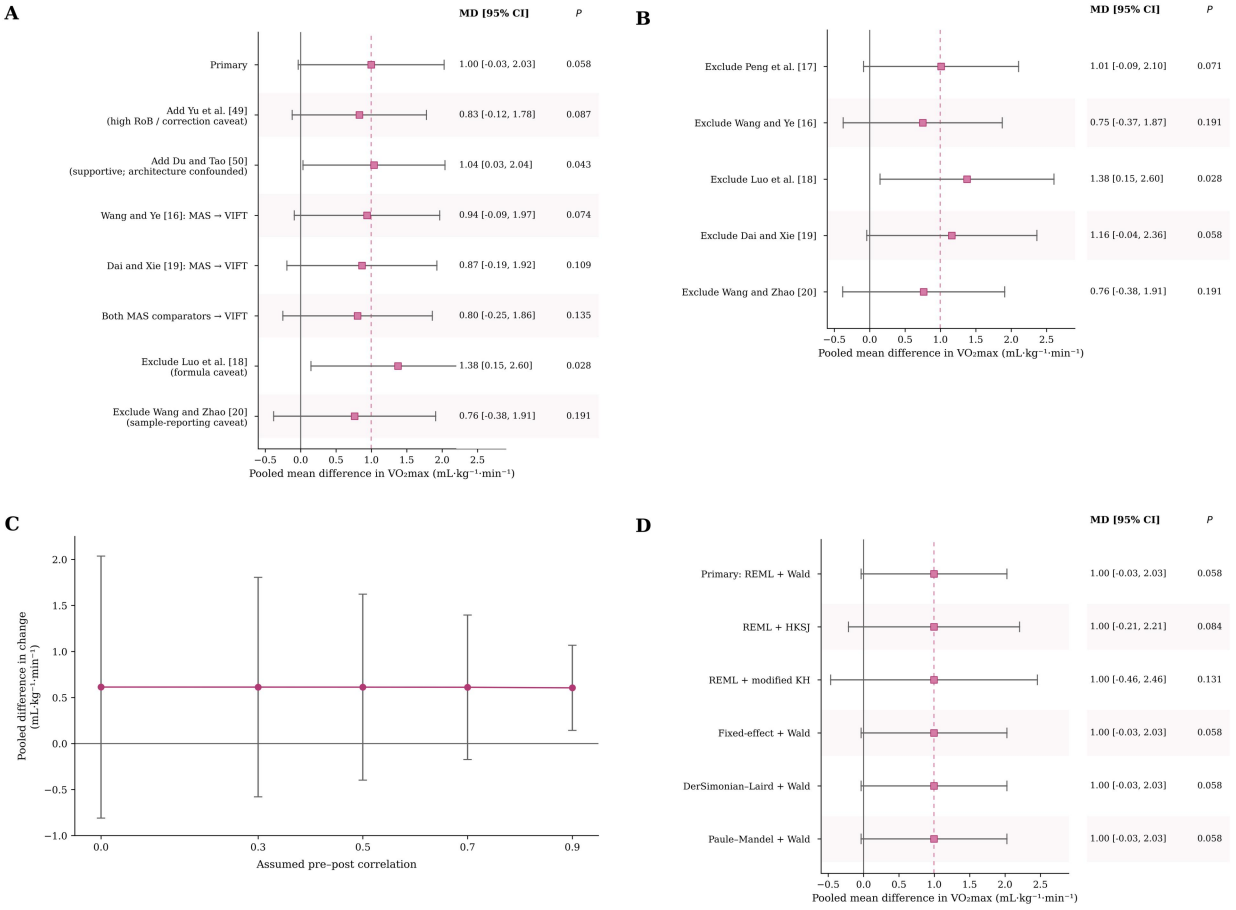
Study	n R/C	Mean change R/C, %	Source R/C	CV R/C, %	Recovered SD R/C, %
Peng 2025 [13]	10/10	5.50 / 4.00	Direct / direct	18.0 / 45.0	0.99 / 1.80
Wang & Ye 2024 [14]	10/10	7.07 / 5.92	Derived / derived	8.7 / 34.6	0.62 / 2.05
Luo 2024 [16]	12/12	6.22 / 5.85	Derived / derived	13.0 / 39.0	0.81 / 2.28
Dai & Xie 2023 [17]	10/10	7.20 / 5.40	Direct / direct	9.0 / 29.0	0.65 / 1.57
Wang & Zhao 2023 [22]	10/10	6.90 / 5.70	Direct / direct	12.0 / 44.0	0.83 / 2.51

Note. R reserve-based prescription group; C conventional comparator group. Mean percentage change was reported directly for Peng 2025, Dai & Xie 2023, and Wang & Zhao 2023, and derived from group preintervention and postintervention means for Wang & Ye 2024 and Luo 2024. Recovered SD is the SD of individual percentage change calculated from the reported CV and mean percentage change for the complementary VR sensitivity. These values are transparency inputs and do not alter the CVR calculation.

Table S4B Published and review-recalculated inference for Di Gennaro et al. (2025 APNM) [24] exercise-tolerance partial correlations

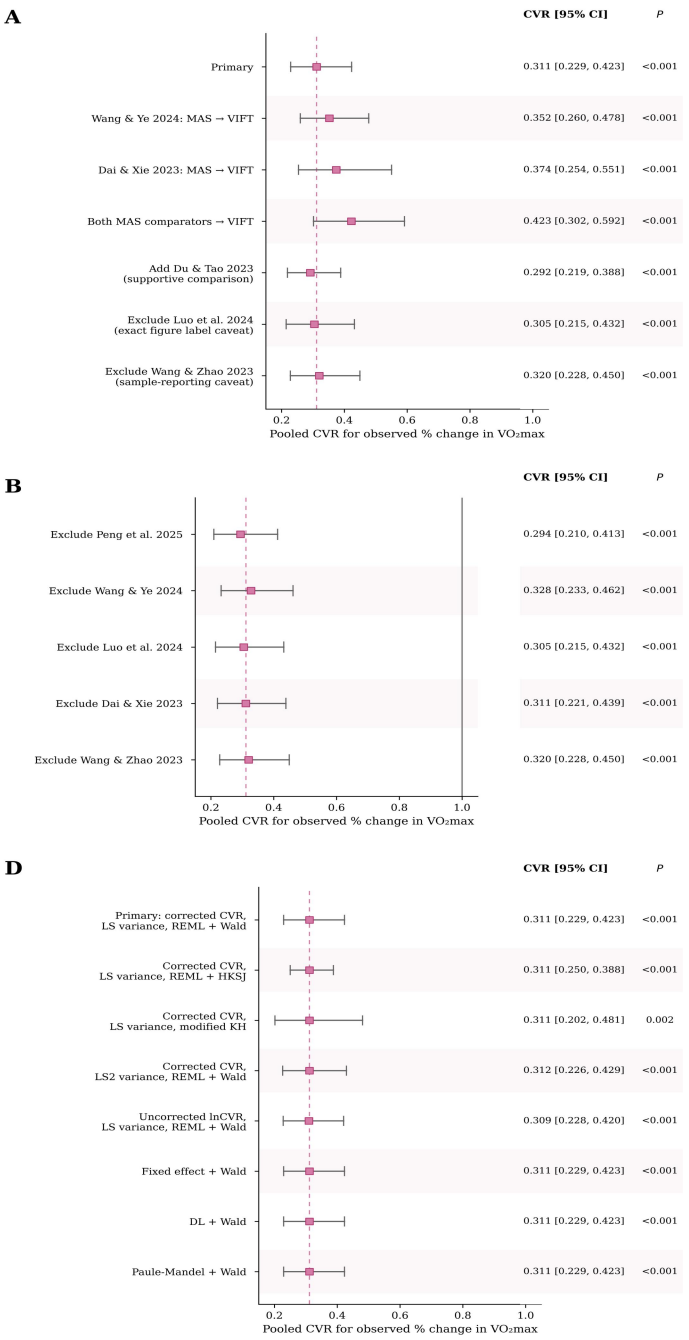
Adjustment	Intensity, % MAP	Partial r	Published 95% CI	Published *P*	Recalculated 95% CI	Recalculated *P*
APR | PPO	130	.467	−0.192 to 0.853	0.059	0.000 to 0.767	0.051
APR | PPO	115	.363	−0.241 to 0.759	0.152	−0.125 to 0.710	0.139
APR | PPO	100	−0.253	−0.637 to 0.197	0.328	−0.644 to 0.243	0.311
APR | PPO	85	−0.255	−0.716 to 0.293	0.323	−0.645 to 0.241	0.307
APR | PPO	80	−0.333	−0.684 to 0.026	0.192	−0.692 to 0.159	0.177
MPR | PPO	130	.606	0.055 to 0.884	0.010	0.194 to 0.836	0.008
MPR | PPO	115	.508	−0.021 to 0.811	0.037	0.054 to 0.788	0.031
MPR | PPO	100	−0.151	−0.519 to 0.276	0.564	−0.577 to 0.340	0.550
MPR | PPO	85	−0.261	−0.765 to 0.346	0.312	−0.649 to 0.234	0.296
MPR | PPO	80	−0.435	−0.768 to −0.139	0.081	−0.750 to 0.040	0.071
PPO | CP	130	.746	0.542 to 0.901	<0.001	0.428 to 0.900	<0.001
PPO | CP	115	.757	0.594 to 0.888	<0.001	0.449 to 0.904	<0.001
PPO | CP	100	.729	0.522 to 0.916	0.001	0.397 to 0.892	<0.001
PPO | CP	85	.348	−0.113 to 0.656	0.157	−0.142 to 0.701	0.157
PPO | CP	80	.279	−0.273 to 0.647	0.263	−0.216 to 0.660	0.262
PPO | MAP	130	.703	0.406 to 0.903	0.002	0.351 to 0.881	0.001
PPO | MAP	115	.728	0.505 to 0.873	0.001	0.396 to 0.892	<0.001
PPO | MAP	100	.792	0.610 to 0.930	<0.001	0.516 to 0.919	<0.001
PPO | MAP	85	.451	0.041 to 0.779	0.070	−0.020 to 0.758	0.060
PPO | MAP	80	.428	−0.277 to 0.774	0.086	−0.049 to 0.746	0.076

Note. Source partial-correlation coefficients were retained. Review inference was recalculated uniformly using n = 19 and one controlled variable because several published confidence-interval and P-value combinations were internally inconsistent. Two-sided P values were obtained from the partial-correlation t statistic with 16 degrees of freedom; 95% confidence intervals were recalculated on the Fisher z scale. Values are rounded to three decimals for display. APR anaerobic power reserve; CP critical power; MAP maximal aerobic power; MPR maximal power reserve; PPO peak power output.



Supplementary Fig. S2 Sensitivity analyses for the chronic VO_2max mean-effect synthesis.

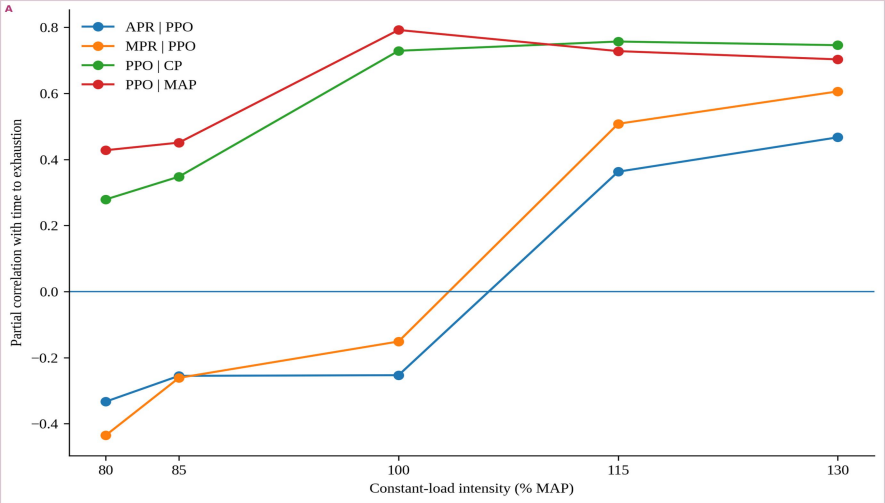
(a) Prespecified analytical sensitivities, including addition of Yu 2025 and Du and Tao 2023, alternative VIFT comparators, and prespecified exclusions. (b) Leave-one-out analyses. (c) Change-score sensitivity across assumed within-group preintervention-to-postintervention correlations from 0.0 to 0.9. (d) Alternative random-effects estimators and interval methods.

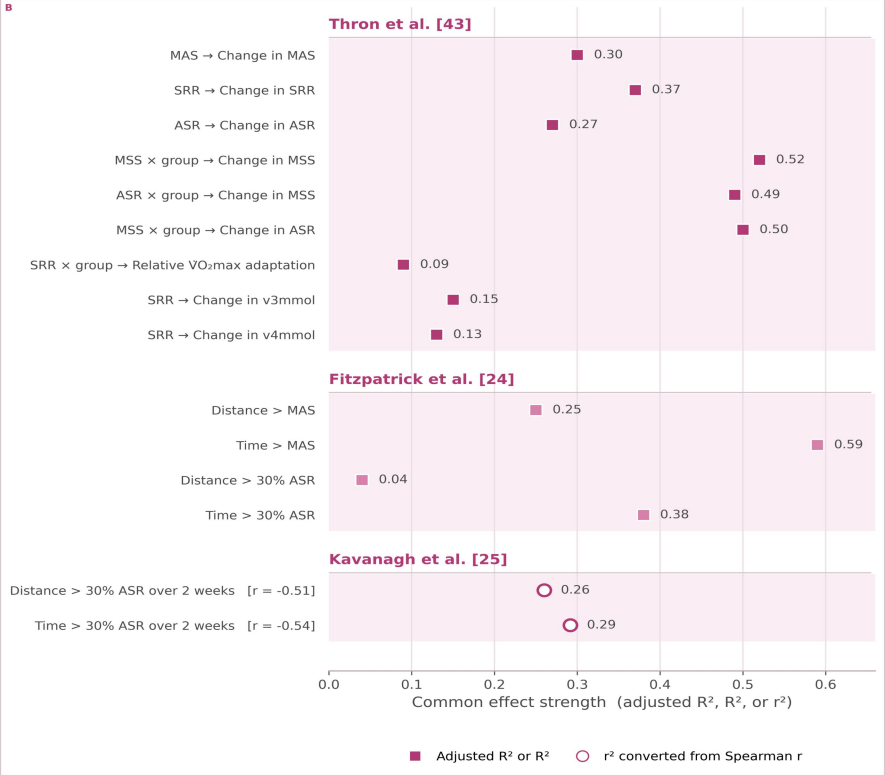


Supplementary Fig. S3 Sensitivity analyses for the chronic variability synthesis based on the coefficient of variation ratio.

(a) Prespecified analytical sensitivities, including alternative VIFT comparators, addition of Du 2023, and prespecified exclusions. (b) Leave-one-out analyses. (c) Alternative model and estimator sensitivity. Across 1024 plausible source-data rounding combinations, the pooled CVR ranged from 0.298 to 0.326, the overall 95% confidence-interval envelope ranged from 0.219 to 0.442, and the maximum τ^2 remained 0

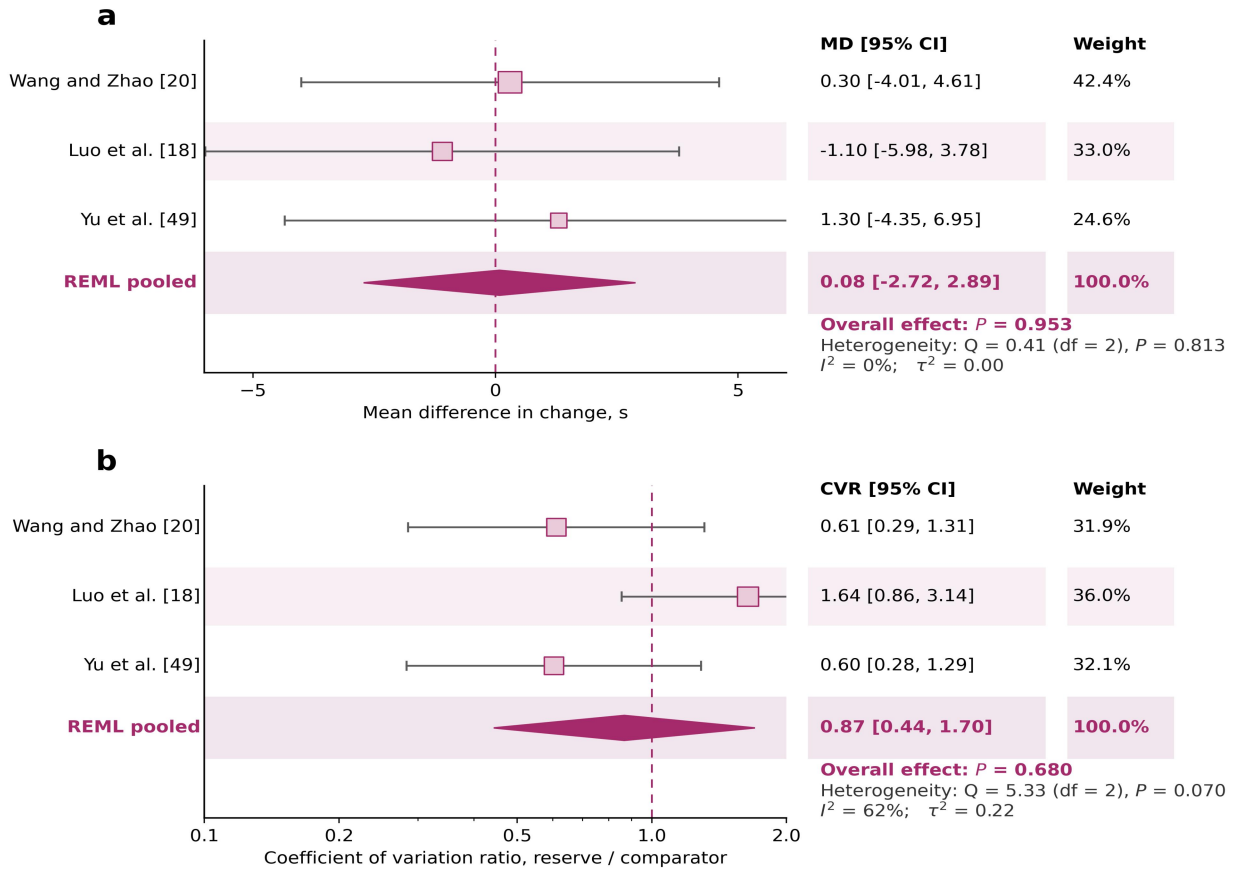
S5 Predictor, moderator, and longitudinal supporting evidence





Supplementary Fig. S4 Supporting evidence for predictor, moderator, and short-term profile signals. (a) Di Gennaro et al. partial correlations between reserve-related indices and time to exhaustion across constant-load intensities, showing how associations of APR, MPR, PPO controlled for critical power, and PPO controlled for MAP varied with intensity. (b) Common effect-strength display for the Thron 2026 moderation models, Fitzpatrick 2018 threshold comparisons, and Kavanagh 2025 longitudinal signals. Adjusted R^2 and R^2 are plotted directly; Kavanagh Spearman correlations are converted to r^2 for plotting while the signed r values remain shown in the row labels

S6 Chronic performance supporting analyses



Supplementary Fig. S5 Exploratory synthesis of chronic 2000 m rowing performance.

(a) Random-effects meta-analysis of the between-group difference in change of 2000 m time trial performance. Positive values favour reserve-based prescription. The pooled change-score estimate assumes a preintervention-to-postintervention correlation of r = 0.5; sensitivity analyses across r = 0.0 to 0.9 yielded the same qualitative interpretation. (b) Random-effects meta-analysis of the coefficient of variation ratio for change in 2000 m time trial performance. CVR values below 1 indicate lower observed relative dispersion under reserve-based prescription. The pooled CVR was calculated using the bias-corrected log CVR with REML estimation. Studies included were Wang and Zhao (2023), Luo et al. (2024), and Yu Tongwu et al. (2025).

Table S5 Structured synthesis of chronic high intensity and sport specific performance outcomes

Outcome family	Studies	Mean adaptation	Observed dispersion	Interpretation
2000 m rowing TT	Wang and Zhao 2023; Luo 2024; Yu 2025	MD 0.08 s, 95% CI −2.72 to 2.89; I^2 = 0%	CVR 0.87, 95% CI 0.44 to 1.70; I^2 = 62%	No clear mean advantage; dispersion mixed
Rowing MSP, 60 s power, critical power	Luo 2024; Yu 2025; Wang and Zhao 2023 for MSP	No clear pooled reserve advantage for comparable k = 2 families	P60 CVR 1.40; CP CVR 1.12; MSP evidence mixed	Outcome-dependent; small k
Lower-body Wingate PPO	Peng 2025; Wang and Ye 2024	MD 1.30 W, 95% CI −44.31 to 46.92 at r = 0.5	CVR 0.60, 95% CI 0.37 to 0.99	Supportive dispersion signal; k = 2 and HKSJ interval wide
Lower-body Wingate APO	Peng 2025; Wang and Ye 2024	MD −33.49 W, 95% CI −104.09 to 37.11; I^2 = 75% at r = 0.5	CVR 0.73, 95% CI 0.45 to 1.19	Inconsistent; Wang reporting caveat
Running MSS	Dai and Xie 2023; Peng 2025	MD −0.07 km/h, 95% CI −0.94 to 0.80 at r = 0.5	Insufficient common CVR synthesis	No clear ASR mean advantage
Sprint kayak 500 m and 1000 m TT	Du and Tao 2023	ASR improved within group; between-group differences were not significant	Outcome CVs favoured ASR where reported	Supportive only because protocols differed in work and recovery
All-out SIT benchmark	Peng 2025; Yu 2025	Several sprint, power, or performance outcomes improved as much or more under SIT	Not a common effect family	Reserve prescription should not be equated with maximal mean performance gain

Note. Mean change effects for k = 2 families use the illustrative r = 0.5 change-score model; conclusions were checked across r = 0.0 to 0.9. Values are exploratory because most families contain only one or two directly comparable studies. CVR coefficient of variation ratio; MD mean difference; MSP maximal sprint power; MSS maximal sprint speed; PPO peak power output; SIT sprint interval training; TT time trial.

S7 Outcome-specific observed response dispersion

Sixty-eight matched chronic outcome-level CVR observations were available. Because outcomes within a study were dependent, the analysis was summarized at the study by domain level and used for descriptive evidence mapping rather than meta-analysis. The main manuscript presents the clustered heatmap as Fig. 4.

Table S6 Outcome-domain summary of chronic observed response dispersion

Domain	Studies	Median of study-domain medians	Reserve lower	Similar	Comparator lower
Maximal aerobic/cardiorespiratory	6	0.386	6	0	0
Threshold/submaximal aerobic	6	0.428	5	0	1
Performance	5	0.792	4	0	1
Reserve/component capacity	5	0.692	5	0	0
Hormonal/other	2	1.032	0	1	1

Note. CVR < 0.90 was classified as reserve lower dispersion, 0.90 to 1.10 as similar, and CVR > 1.10 as comparator lower dispersion for visualization only. The neutral band is not an equivalence margin.

Table S7 Study by outcome-domain data underlying the chronic dispersion heatmap

Study	Outcome domain	n outcomes	Median CVR	Cell signal
Peng 2025	Maximal aerobic/cardiorespiratory	3	0.44	Reserve lower dispersion
Peng 2025	Threshold/submaximal aerobic	2	0.43	Reserve lower dispersion
Peng 2025	Performance	5	0.79	Reserve lower dispersion
Wang & Ye 2024	Maximal aerobic/cardiorespiratory	4	0.31	Reserve lower dispersion
Wang & Ye 2024	Threshold/submaximal aerobic	2	0.33	Reserve lower dispersion
Wang & Ye 2024	Performance	2	0.59	Reserve lower dispersion
Wang & Ye 2024	Hormonal/other	2	0.92	Similar
Dai & Xie 2023	Maximal aerobic/cardiorespiratory	6	0.58	Reserve lower dispersion
Dai & Xie 2023	Threshold/submaximal aerobic	2	0.43	Reserve lower dispersion
Dai & Xie 2023	Reserve/component capacity	2	0.77	Reserve lower dispersion
Wang & Zhao 2023	Maximal aerobic/cardiorespiratory	4	0.49	Reserve lower dispersion
Wang & Zhao 2023	Threshold/submaximal aerobic	2	0.61	Reserve lower dispersion
Wang & Zhao 2023	Performance	1	0.61	Reserve lower dispersion
Wang & Zhao 2023	Reserve/component capacity	2	0.89	Reserve lower dispersion
Wang & Zhao 2023	Hormonal/other	2	1.15	Comparator lower dispersion
Du & Tao 2023	Maximal aerobic/cardiorespiratory	2	0.25	Reserve lower dispersion
Du & Tao 2023	Threshold/submaximal aerobic	1	0.28	Reserve lower dispersion
Du & Tao 2023	Reserve/component capacity	2	0.45	Reserve lower dispersion
Luo 2024	Maximal aerobic/cardiorespiratory	5	0.33	Reserve lower dispersion
Luo 2024	Threshold/submaximal aerobic	8	1.67	Comparator lower dispersion
Luo 2024	Performance	3	1.64	Comparator lower dispersion
Luo 2024	Reserve/component capacity	2	0.69	Reserve lower dispersion
Yu 2025	Performance	3	0.81	Reserve lower dispersion
Yu 2025	Reserve/component capacity	1	0.50	Reserve lower dispersion

Note. Each row corresponds to one populated cell in main manuscript Fig. 4. The complete 68 matched outcome-level CVR observations are retained in the reproducibility workbook ASR_APR_outcome_domain_CV_evidence_map_20260813.xlsx. CVR coefficient of variation ratio.

S8 Longitudinal reserve dynamics

The analysis included four chronic studies with complete lower-anchor, upper-anchor, and reserve triads. Nine active training arms and one control arm were available. The main manuscript presents the percentage-change pattern as Fig. 5.

Table S8 Arm-level changes in lower anchor, upper anchor, and reserve

Study	Arm	Arm type	Lower change	Upper change	Reserve change	Lower minus upper, pp
Dai and Xie 2023	ASR	reserve based	7.35%	0.64%	-5.08%	6.71
Dai and Xie 2023	VIFT	conventional	8.09%	0.65%	-5.81%	7.44
Dai and Xie 2023	MAS	conventional	6.57%	0.97%	-2.92%	5.60
Du and Tao 2023	ASR-HIIT	reserve based	6.92%	0.24%	-23.75%	6.67
Du and Tao 2023	MAS-HIIT	conventional	5.16%	0.15%	-16.57%	5.01
Du and Tao 2023	Control	control	0.65%	0.24%	0.40%	0.40
Wang and Zhao 2023	APR	reserve based	7.58%	2.17%	-4.75%	5.41
Wang and Zhao 2023	MAP	conventional	6.44%	2.57%	-2.83%	3.87
Luo et al. 2024	APR-HIIT	reserve based	6.04%	0.49%	-4.85%	5.55
Luo et al. 2024	WVO2max-HIIT	conventional	5.73%	1.42%	-3.58%	4.31

Note. Lower, upper, and reserve changes are calculated within each arm from reported preintervention and postintervention means. Across the nine active training arms, lower-anchor improvement exceeded upper-anchor improvement in 9/9. Because the reserve is derived from these components, this pattern was accompanied mathematically by a reserve decrease in 9/9. Median changes were +6.57%, +0.65%, and −4.85%, respectively. pp percentage points.

Table S9 Illustrative drift in reserve-based workload position

Study	Reserve arm	Baseline target	Posttraining equivalent
Dai and Xie 2023	ASR	20%	15.1%
Du and Tao 2023	ASR-HIIT	20%	-5.2%
Wang and Zhao 2023	APR	30%	21.0%
Luo et al. 2024	APR-HIIT	30%	25.4%

Note. Posttraining equivalent is the position within the updated reserve that would be occupied by the original baseline absolute workload if that workload were retained unchanged. It is an illustrative calculation and does not reconstruct the actual training progression.

S9 Component anchor measurement appraisal

Component anchor measurement was examined as a post hoc contextual methodological appraisal because reserve values depend on how both performance boundaries are operationalized. Its purpose was to document whether the included reserve constructs were based on measurements that aligned with available mode-specific validity and reliability evidence and to identify where that alignment was indirect, protocol-dependent, or incomplete. External sources were selected from the cited measurement review and other available mode-specific evidence; no dedicated systematic search of measurement studies was undertaken. The appraisal was therefore not treated as a second systematic review. It was deliberately descriptive and did not change study eligibility, risk of bias judgements, statistical weighting, meta-analytic estimates, or GRADE certainty. Lower and upper anchors were judged separately as supported, supported with caveat, limited support, or not assessable, and no numerical or composite quality score was assigned. The 24 included reports represented 23 independent studies. Girard and Racinais (2023) was a companion report to Soo et al. (2021) and shared the same component anchor measurements, so no additional independent anchor row was created.

Table S10 Criteria for appraisal of component anchor measurement evidence alignment

Domain	Classification	Operational criterion
Anchor evidence alignment	Supported	Direct or closely matched external validity evidence supports the measurement method and operational definition in a comparable exercise mode and population.
Anchor evidence alignment	Supported with caveat	Evidence generally supports the method, but a relevant difference exists in protocol, population, device, distance, reference criterion, or operational definition.
Anchor evidence alignment	Limited support	Related reliability or construct evidence exists, but direct validity evidence for the specific operational definition is lacking, or available evidence indicates limited interchangeability with the intended anchor.
Anchor evidence alignment	Not assessable	Reporting in the included study is insufficient to determine how the anchor was measured or operationally defined.
Mode and dimensional compatibility	Compatible	Lower and upper anchors represent the same exercise mode and compatible speed or power dimensions.
Mode and dimensional compatibility	Compatible with caveat	The broad exercise mode is comparable, but a meaningful cross setting or equipment difference could affect interchangeability of the constructed reserve.
Mode and dimensional compatibility	Not compatible	Exercise modes or dimensions are incompatible for direct reserve construction.
Reliability evidence	Directly demonstrated	Reliability or repeat measurement consistency was quantified within the included report for the exact or closely matching anchor or test.
Reliability evidence	Externally supported	The included report did not provide its own repeatability statistic, but a closely matched published protocol or device has external reliability evidence.
Reliability evidence	Not demonstrated	No sufficiently matched within study or published external reliability evidence was identified for the exact operational definition.
Reliability evidence	Not assessable	The measurement method was reported too incompletely for a reproducible reliability judgement.

Note. Evidence alignment addresses contextual measurement support for the specific lower or upper performance boundary and is distinct from study risk of bias. Running MAS and MSS judgements were anchored primarily to Thron et al. (2024) and supporting criterion or reliability studies; cycling judgements considered incremental protocol dependence, critical power methods, and maximal sprint power evidence; rowing, kayak, and judo judgements used available mode-specific validity and reliability studies. Reliability was appraised separately for lower and upper anchors. Because this appraisal was added after data extraction and the external measurement literature was not identified through a dedicated systematic search, the classifications should be interpreted as contextual methodological judgements rather than a systematic synthesis of measurement evidence. They did not alter the primary synthesis.

Table S11 Component anchor measurement appraisal of included studies

Study	Construct	Lower anchor and measurement	Lower evidence alignment	Upper anchor and measurement	Upper evidence alignment	Reliability evidence	Methodological note
Thron et al., 2025 [11]	ASR / relative ASR	MAS. Treadmill CPET; first interpolated velocity at VO_2 plateau	Supported. Thron et al. 2024 [38]	MSS. Maximal sprint; radar gun with speed time curve fitting	Supported. Thron et al. 2024 [38]	Lower: Externally supported (Thron et al. 2024) [38]. Upper: Externally supported (Thron et al. 2024) [38].	Both running anchors were directly assessed in the same mode. The MAS definition is physiologically explicit and the MSS method uses a reference measurement technology.
Buchheit et al., 2012 [2]	ASR	vVO_2max. Vam Eval incremental field test with breath by breath gas analysis; lowest speed eliciting VO_2max for at least 30 s	Supported with caveat. Thron et al. 2024 [38]	MSS. Three maximal 40 m sprints; fastest 10 m split using dual beam timing gates	Supported with caveat. Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]	Lower: Externally supported (Midgley et al. 2007b [41]; Thron et al. 2024 [38]). Upper: Externally supported; the included report also cites prior CV = 1.4% for the same MSS method.	Both anchors were directly assessed in running. The lower anchor used a physiologically explicit vVO_2max definition; the timing gate MSS method is supported with a device and split distance caveat.
Di Gennaro et al., 2025 JSM [12]	APR	MAP. Cycle CPET; 25 W every 3 min; fractional final stage equation	Supported with caveat. Michalik et al. 2019 [42]; Balmer et al. 2000 [43]	PPO. 30 s Wingate; maximal instantaneous power	Supported. Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]	Lower: Externally supported (Balmer et al. 2000 [43]; Bentley and McNaughton 2003 [46]). Upper:Not demonstrated for the exact operational definition.	The lower MAP anchor is clearly operationalized but protocol dependent. The cited studies concern short maximal cycling sprints and peak power; they do not directly establish the reliability of 30 s mean Wingate power or its validity as a maximal upper boundary.
Di Gennaro et al., 2025 JSM[12]	GPR	MAP. Cycle CPET; 25 W every 3 min; fractional final stage equation	Supported with caveat. Michalik et al. 2019 [42]; Balmer et al. 2000 [43]	30 s mean Wingate power. Mean power output across 30 s Wingate	Limited support. Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]	Lower: Externally supported (Balmer et al. 2000 [43]; Bentley and McNaughton 2003 [46]). Upper: Limited supported (Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]).	The lower MAP anchor is clearly operationalized but protocol dependent. Mean Wingate power is a reproducible performance measure, yet it is less directly supported as a maximal upper boundary than peak sprint power.
Peng et al., 2025 [13]	ASR	MAS. Graded treadmill test with gas analysis; minimal speed eliciting VO_2max	Supported. Thron et al. 2024 [38]	MSS. Two 40 m sprints; fastest 10 m timing split	Supported with caveat. Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]	Lower: Externally supported (Thron et al. 2024 [38]). Upper: Externally supported (Thron et al. 2024 [38];	The lower anchor directly represents a physiologically defined aerobic speed. Ten metre timing splits are externally supported but are not fully interchangeable with

Study	Construct	Lower anchor and measurement	Lower evidence alignment	Upper anchor and measurement	Upper evidence alignment	Reliability evidence	Methodological note
						Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]).	continuous radar based MSS.
Wang and Ye, 2024 [14]	ASR	MAS. Graded treadmill test with gas analysis; minimal velocity eliciting VO_2max	Supported. Thron et al. 2024 [38]	MSS. 40 m sprint; fastest 10 m electronic timing split	Supported with caveat. Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]	Lower: Externally supported (Thron et al. 2024 [38]). Upper: Externally supported (Thron et al. 2024 [38]; Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]).	The lower anchor is directly physiologically defined. The upper anchor is a practical and supported estimate of MSS, with known individual level differences relative to radar.
Zhu et al., 2025 [15]	ASR locomotor profile	MAS. 5 min continuous maximal field run; distance based estimate	Supported with caveat. Berthon et al. 1997 [47]; Thron et al. 2024 [38]	MSS. 30 m sprint; Photo Finish timing with 25 to 30 m split	Supported with caveat. Roe et al. 2017[39]; Zabaloy et al. 2024 [40]	Lower: Externally supported (Dabonneville et al. 2003 [48]). Upper: Directly demonstrated (within study).	A five minute field test has external support for estimating maximal aerobic velocity, but population and protocol dependence remain. Sprint timing was directly checked for intra observer reliability, although this is not athlete test retest reliability.
Luo et al., 2024 [16]	APR	Power at VO_2max. Graded rowing test; regression derived power at VO_2max	Supported with caveat. Huntsman et al. 2011 [49]	MSP. 3 min all out rowing; mean power from 0 to 10 s	Limited support. Metikos et al. 2015 [50]; Cheng et al. 2012 [51]	Lower: Not demonstrated (no sufficiently matched external evidence). Upper: Not demonstrated (no sufficiently matched external evidence).	Rowing specific aerobic testing is supported, although the exact regression derived lower anchor differs from directly measured incremental MAP. External validation of the three minute test concerns end test power and critical power, not the first ten second mean as MSP.
Dai and Xie, 2023 [17]	ASR	MAS. Incremental treadmill test with gas analysis; minimal velocity eliciting VO_2max	Supported. Thron et al. 2024 [38]	MSS. Two 40 m sprints; fastest 10 m split	Supported with caveat. Roe et al. 2017[39]; Zabaloy et al. 2024 [40]	Lower: Externally supported (Thron et al. 2024 [38]). Upper: Externally supported (Thron et al. 2024 [38]; Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]).	The lower anchor is physiologically explicit. The upper anchor follows a widely used timing gate method with external support and a known caveat versus continuous radar measurement.
Yu Tongwu et al., 2025	APR	MAP. Progressive RowErg	Not assessable. Huntsman et al.	MSP. 3 min all out rowing; mean	Limited support. Metikos et al.	Lower: Not assessable (not assessable). Upper:	The report describes the progressive cardiorespiratory test

Study	Construct	Lower anchor and measurement	Lower evidence alignment	Upper anchor and measurement	Upper evidence alignment	Reliability evidence	Methodological note
[18]		cardiorespiratory test; exact MAP derivation not stated	2011 [49];	power from 0 to 10 s	2015 [50]; Cheng et al. 2012 [51]	Directly demonstrated (within study).	but does not define exactly how MAP was derived, preventing a reproducible lower anchor validity judgement. The zero to ten second MSP definition lacks directly matched validation.
Du and Tao, 2023 [19]	ASR	MAS. Graded kayak ergometer test with gas analysis; lowest speed eliciting VO_2peak	Supported with caveat. Winchcombe et al. 2019 [52]	MSS. Two maximal kayak ergometer sprints; highest second by second velocity	Limited support. Winchcombe et al. 2019 [52]	Lower: Externally supported [52]. Upper: Not demonstrated (no sufficiently matched external evidence).	Kayak specific graded testing has external support, but the exact MAS operational definition is not identical to the available validation studies. No closely matched criterion validation was identified for the very short kayak ergometer MSS definition.
Aspin et al., 2024 [20]	ASR	MAS. 1600 m track time trial; average speed estimate	Supported with caveat. Bellenger et al. 2015 [53]; Lundquist et al. 2021 [54]; Thron et al. 2024 [38]	MSS. Three 40 m sprints; fastest 10 m dual beam timing split	Supported with caveat. Roe et al. 2017[39]; Zabaloy et al. 2024 [40]	Lower: Directly demonstrated (within study). Upper: Directly demonstrated (within study).	Fixed distance trials can estimate MAS in field settings, but the optimal distance depends on population. The study itself reported acceptable test retest reliability for MAS, MSS and ASR.
Julio et al., 2020 [21]	ASR	MAS. Incremental field track test; final stage speed or fractional stage adjustment	Limited support. Thron et al. 2024 [38]; Thron et al. 2022 [55]	MSS. Two 40 m sprints; shortest 10 m timing split	Supported with caveat. Roe et al. 2017[39]; Zabaloy et al. 2024 [40]	Lower: Not demonstrated (no sufficiently matched external evidence). Upper: Externally supported (Thron et al. 2024 [38]; Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]).	Incremental field final speed cannot be assumed to equal criterion MAS on an individual basis. The timing gate MSS procedure has useful external support with an individual level caveat.
Wang and Zhao, 2023 [22]	APR	MAP. Incremental RowErg with gas analysis; minimal power eliciting VO_2max	Supported. Huntsman et al. 2011 [49]	MSP. All out rowing; maximum within ten strokes; mean of final five strokes	Supported with caveat. Metikos et al. 2015 [50]	Lower: Externally supported (Huntsman et al. 2011 [49]). Upper: Externally supported (Metikos et al. 2015 [50]).	The lower anchor is directly and physiologically defined on a rowing ergometer. The short maximal stroke protocol is closely related to validated rowing peak power and seven stroke tests, although it is not identical.
Soo et al., 2021; Girard	APR	MAP. Incremental cycling; last fully	Supported with caveat. Michalik	PPO. Three 10 s maximal cycling	Supported with caveat. Mendez-	Lower: Externally supported (Balmer et al.	Both anchors are cycling specific. The lower anchor is a protocol

Study	Construct	Lower anchor and measurement	Lower evidence alignment	Upper anchor and measurement	Upper evidence alignment	Reliability evidence	Methodological note
and Racinais, 2023 companion report [9, 10]		completed stage	et al. 2019 [42]; Balmer et al. 2000 [43]	sprints; highest peak power	Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]	2000 [43]; Bentley and McNaughton 2003 [46]). Upper: Externally supported (Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]).	dependent final stage definition. Ten second sprint peak power is strongly supported by nearby maximal sprint evidence, but the exact protocol differs from the most directly validated six second and Wingate tests. The companion report uses the same cohort, project, APR prescription, and component anchors and therefore shares this appraisal.
Collison et al., 2022 [23]	ASR	MAS. 2000 m time trial; average speed prediction	Supported with caveat. Bellenger et al. 2015 [53]	MSS. Two sprint trials; shortest 10 m timing split	Supported with caveat. Roe et al. 2017[39]; Zabaloy et al. 2024 [40]	Lower: Externally supported (Thron et al. 2024 [38]). Upper: Externally supported (Thron et al. 2024 [38]; Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]).	The 2000 m time-trial method is supported in a highly comparable Australian football population, but the key Bellenger validation used UMTT as the reference rather than CPET; it is therefore retained with a measurement-criterion caveat. The timing-gate MSS estimate is externally supported but is not fully interchangeable with continuous radar measurement.
Di Gennaro et al., 2025 APNM [24]	APR	MAP. MAP as final power output achieved during the ramp test	Supported with caveat. Michalik et al. 2019 [42]; Balmer et al. 2000 [43]	PPO. 30 s Wingate; highest 1 s power	Supported with caveat. Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]	Lower: Externally supported (Michalik et al. 2019 [42]; Balmer et al. 2000 [43]; Bentley and McNaughton 2003 [46]). Upper: Externally supported (Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]).	MAP is clearly defined from the ramp test, but external evidence shows that peak incremental cycling power is protocol dependent. PPO was obtained directly from a maximal cycling sprint.
Di Gennaro et al., 2025 APNM [24]	MPR	Critical power. CP estimated from five severe intensity time to exhaustion trials	Supported. Simpson and Kordi 2017 [56]; Karsten et al. 2015 [57]	PPO. 30 s Wingate; highest 1 s power	Supported with caveat. Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006	Lower: Externally supported (Simpson and Kordi 2017 [56]; Karsten et al. 2015 [57]). Upper: Externally supported (Mendez-	Critical power was estimated from multiple severe intensity trials, matching established cycling CP methodology. PPO was obtained directly from a maximal cycling sprint.

Study	Construct	Lower anchor and measurement	Lower evidence alignment	Upper anchor and measurement	Upper evidence alignment	Reliability evidence	Methodological note
					[45]	Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]).	
Salter et al., 2025 [25]	ASR locomotor profile	MAS. 1000 m track time trial; average speed estimate	Limited support. Bellenger et al. 2015 [53]; Lundquist et al. 2021 [54]; Thron et al. 2024 [38]	MSS. Two 40 m maximal sprints; radar peak velocity	Supported. Thron et al. 2024 [38]; Roe et al. 2017 [39]	Lower: Externally supported [53]. Upper: Externally supported (Thron et al. 2024 [38]).	The authors justify the shorter time trial for youth females and report reliability, but direct criterion validity for 1000 m as MAS in this population remains limited. Radar based MSS is strongly supported.
Thron et al., 2026 [26]	ASR / SRR	MAS. Treadmill CPET; first interpolated velocity at VO_2 plateau with alternate VO_2 criterion	Supported. Thron et al. 2024 [38]	MSS. 40 m sprint; radar peak speed	Supported. Thron et al. 2024 [38]; Roe et al. 2017 [39]	Lower: Externally supported (Thron et al. 2024 [38]). Upper: Externally supported (Thron et al. 2024 [38]; Roe et al. 2017 [39]).	Both anchors are directly measured using the best supported running specific approaches and are expressed in the same speed domain.
Belechris et al., 2025 [27]	ASR	vVO_2max. Incremental treadmill test; speed corresponding to highest VO_2 by Douglas bag	Supported with caveat. Thron et al. 2024 [38]	vmax. Two maximal 60 m sprints; high speed video based sprint velocity analysis	Supported. Thron et al. 2024 [38]	Lower: Externally supported (Thron et al. 2024 [38]). Upper: Externally supported (Thron et al. 2024 [38]).	The lower anchor is directly tied to measured VO_2 but uses the highest observed VO_2 stage speed rather than plateau onset. The 60 m sprint distance and high speed video approach are strongly aligned with supported MSS assessment.
Girard et al., 2021 [28]	APR	MAP. Incremental cycling test; 25 W per min to exhaustion	Supported with caveat. Michalik et al. 2019 [42]; Balmer et al. 2000 [43]	PPO. Three 10 s maximal cycling sprints	Supported with caveat. Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]	Lower: Externally supported (Balmer et al. 2000 [43]; Bentley and McNaughton 2003 [46]). Upper: Externally supported (Mendez-Villanueva et al. 2007 [44]; Del Coso and Mora-Rodríguez 2006 [45]).	The cycling mode is consistent. Incremental peak power remains protocol dependent, and the ten second maximal sprint is closely but not exactly matched to the strongest external sprint validation protocols.
Fitzpatrick et al., 2018 [29]	ASR high speed threshold	MAS. 1500 m time trial on artificial pitch; average speed estimate	Limited support. Thron et al. 2024 [38]; Thron et al. 2022 [55]	MSS. Two 40 m sprints; 30 to 40 m timing split	Supported with caveat. Roe et al. 2017[39]; Zabaloy et al. 2024 [40]	Lower: Not demonstrated (unpublished MDC only). Upper: Externally supported	A 1500 m time trial can be related to MAS, but direct evidence in trained soccer players indicates important individual overestimation. The study reported

Study	Construct	Lower anchor and measurement	Lower evidence alignment	Upper anchor and measurement	Upper evidence alignment	Reliability evidence	Methodological note
						(Thron et al. 2024 [38]; Roe et al. 2017 [39]; Zabaloy et al. 2024 [40]).	its own minimum detectable changes for both anchors.
Franchini et al., 2022 [30]	Judo specific ASR	Judo specific MAS. Incremental hikidashi test; last completed stage	Supported. Shiroma et al. 2019 [58]	Judo specific MSS. Five all out hikidashi repetitions timed with digital timer	Limited support.	Lower: Externally supported (Shiroma et al. 2019 [58]). Upper: Directly demonstrated (within study).	The lower anchor has directly matched judo specific validity and reliability evidence. Related maximal all out hikidashi tests are reproducible, but the exact five repetition speed definition lacks directly matched criterion validity.
Kavanagh et al., 2025 [31]	ASR threshold	MAS. 1200 m Bronco shuttle; correction equation for change of direction time	Limited support. Thron et al. 2024 [38]; Thron et al. 2022 [55]	MSS. Highest speed during repeated linear peak speed exposures; 10 Hz STATSports Apex GPS	Supported with caveat. Beato et al. 2018 [59]; Thron et al. 2024 [38]	Lower: Not demonstrated (no sufficiently matched external evidence). Upper: Externally supported (Beato et al. 2018 [59]; Thron et al. 2024 [38]).	Both lower and upper anchors are running-speed measures and are dimensionally compatible. The Bronco-derived MAS remains a proxy with limited criterion support, while the STATSports Apex 10 Hz upper anchor has device-specific external support but was observed during repeated peak-speed exposures rather than a standardized dedicated sprint test.

Note. The table is intended to document the measurement basis of the included reserve constructs and to make uncertainty in anchor operationalization visible without changing the review question. All construct rows met the review eligibility requirement for dimensional and exercise mode compatibility. Supported indicates direct or closely matched external evidence; supported with caveat indicates generally supportive evidence with a relevant protocol, population, device, reference criterion, or definition difference; limited support indicates that related evidence exists but direct support for the exact operational definition is lacking; not assessable indicates insufficient reporting. Directly demonstrated reliability refers to evidence reported in the included study, whereas externally supported reliability refers to a closely matched published protocol or device. APR anaerobic power reserve; ASR anaerobic speed reserve; CP critical power; CPET cardiopulmonary exercise testing; GPR glycolytic power reserve; MAP maximal aerobic power; MAS maximal aerobic speed; MPR maximal power reserve; MSP maximal sprint power; MSS maximal sprint speed; PPO peak power output; SRR speed reserve ratio; VO_2max maximal oxygen uptake; vVO_2max velocity at maximal oxygen uptake.

S10 GRADE certainty of evidence

Certainty was assessed separately for each formal review question using the five GRADE domains reported in the main manuscript [60]. Randomized intervention evidence began at high certainty, while nonrandomized or observational predictor and exposure evidence began at low certainty. The table below provides the domain profile underlying the final certainty ratings. Domain judgements summarize the magnitude of

concern for each evidence body; publication bias was not rated down when the evidence base was too small for formal assessment and no independent evidence justified an additional downgrade. Environmental-modifier reports were retained as exploratory contextual evidence and were not assigned a separate GRADE rating because they did not directly compare reserve-based with conventional prescription.

Table S12 GRADE evidence profile by review question

Review question	Evidence base	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Final certainty	Key rationale
Chronic mean VO_2max adaptation	5 randomized parallel studies; N = 104	Serious	Not serious	Not serious	Serious	Not rated down	Low	All five primary studies had some risk of bias concerns. Point estimates were directionally similar and heterogeneity was low, but the pooled interval crossed no difference and remained compatible with a potentially meaningful benefit.
Chronic observed response dispersion	5 randomized parallel studies; N = 104	Serious	Not serious	Not serious	Not serious	Not rated down	Moderate	The same five studies had some risk of bias concerns, but all primary contrasts favoured lower observed dispersion. The pooled interval remained well below CVR = 1 and all rounding, comparator, and leave one out sensitivities retained the same direction. The Moderate rating applies specifically to the observed dispersion contrast and does not imply moderate certainty that reserve-based prescription reduces true biological response variability or improves athlete-level training decisions.
Chronic high intensity and sport specific performance adaptation	Several chronic direct studies; 3 rowing studies in exploratory 2000 m synthesis	Serious	Not rated down separately	Not serious	Serious	Not rated down	Low	The evidence base was small and one study in the 2000 m synthesis was at high risk of bias. Mean performance estimates were imprecise and dispersion was outcome dependent; heterogeneity across performance constructs was handled by restricting pooling to comparable outcomes rather than by an additional inconsistency downgrade.
Acute direct standardization	3 direct acute crossover or repeated-measures studies	Serious	Serious	Not serious	Serious	Not rated down	Very low	All three studies had some risk of bias concerns. Running studies were supportive, whereas the cycling study did not reproduce lower dispersion under reserve conditions; samples were small and no common dispersion metric could be pooled.
General predictor or profile evidence	Observational acute predictor and profile studies	Very serious	Serious	Not serious	Serious	Not rated down	Very low	Concerns centred on confounding, multiple analyses, profile cut points, attrition, and reporting or model instability. Associations varied by outcome, session format, sport, and profile definition, and several studies were small.
Incremental prediction beyond	Di Gennaro 2025 APNM and Belechris	Very serious	Serious	Not serious	Serious	Not rated down	Very low	Both adjusted prediction studies were small and had higher risk of bias concern. Selected MPR

Review question	Evidence base	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Final certainty	Key rationale
conventional anchors	2025							associations persisted after adjustment, whereas ASR did not add independent prediction in the Belechris model, and several adjusted intervals were wide.
Profile-guided allocation	1 randomized parallel trial; N = 46	Serious	Not serious	Not serious	Serious	Not rated down	Low	The single randomized trial directly tested profile guided allocation but had some risk of bias concerns and limited precision for a treatment selection effect.
Treatment moderation	1 nonrandomized training study; n = 27 completers	Serious	Not serious	Not serious	Serious	Not rated down	Very low	The study had serious risk of bias because of confounding and missing data. Numerous interaction tests were supported by a small completer sample and uncertainty was incompletely available for synthesis.
Longitudinal individualized thresholds	2 prospective observational cohorts; n = 14 and n = 7	Very serious	Serious	Not serious	Serious	Not rated down	Very low	Both studies had higher risk of bias concern, one with only seven analyzable athletes. Fitzpatrick favoured MAS thresholds whereas Kavanagh highlighted selected ASR thresholds, and multiple correlated thresholds and time windows were evaluated.

Note. Exploratory contextual environmental evidence was not assigned a separate GRADE rating. Abbreviations: CVR coefficient of variation ratio; GRADE Grading of Recommendations Assessment, Development and Evaluation; MAP maximal aerobic power; MAS maximal aerobic speed; MPR maximal power reserve; VO_2max maximal oxygen uptake.

Supplementary References

1. Sandford GN, Laursen PB, Buchheit M. Anaerobic Speed/Power Reserve and Sport Performance: Scientific Basis, Current Applications and Future Directions. Sports Med. 2021 Oct;51(10):2017-28.
2. Buchheit M, Hader K, Mendez-Villanueva A. Tolerance to high-intensity intermittent running exercise: do oxygen uptake kinetics really matter? Front Physiol. 2012;3:406.
3. Blondel N, Berthoin S, Billat V, Lensel G. Relationship between run times to exhaustion at 90, 100, 120, and 140% of vVO2max and velocity expressed relatively to critical velocity and maximal velocity. Int J Sports Med. 2001 Jan;22(1):27-33.
4. Barnett C, Jenkins D, MacKinnon L, Green S. A new method for the calculation of constant supra-VO2peak power outputs. Med Sci Sports Exerc. 1996 Dec;28(12):1505-9.
5. Heaney NE. THE EFFECT OF A FOUR WEEK HIGH INTENSITY INTERVAL TRAINING INTERVENTION UTILISING MAXIMAL AEROBIC SPEED AND THE ANAEROBIC SPEED RESERVE ON AEROBIC FITNESS IN ELITE FEMALE HOCKEY PLAYERS. Journal of Australian Strength & Conditioning. 2014;22(5):170-5.
6. Sterne JAC, 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 Aug 28;366:l4898.
7. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016 Oct 12;355:i4919.
8. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013 Feb 19;158(4):280-6.
9. Girard O, Racinais S. Exhaustive Intermittent Cycling Preferentially Decreases Explosive Over Maximal Voluntary Torque in the Knee Extensors, With No Difference Between Normoxia and Moderate to Severe Hypoxia. Int J Sports Physiol Perform. 2023 Nov 1;18(11):1352-6.
10. Soo J, Racinais S, Fairchild TJ, Ihsan M, Buchheit M, Girard O. Effects of graded hypoxia during exhaustive intermittent cycling on subsequent exercise performance and neuromuscular responses. Eur J Appl Physiol. 2021 Dec;121(12):3539-49.
11. Thron M, Ruf L, Buchheit M, Hartel S, Woll A, Altmann S. Anaerobic speed reserve and acute responses to a short-format high-intensity interval session in runners. J Sci Med Sport. 2025 May;28(5):408-17.
12. S DIG, Filipas L, Panasci M, Gallo G, Meloni A, Codella R, et al. Maximal aerobic power and anaerobic power reserves to prescribe cycling interval training sessions. J Sports Med Phys Fitness. 2025 Aug;65(8):1046-54.
13. Peng Q, Zhang Z, Gao L, Zhang H, Liu Y. Modified Sprint Interval Intervention Produces Lower Inter-Subject Variability in Physiological and Performance Adaptations Across Collegiate Soccer Players. J Sports Sci Med. 2025 Dec;24(4):839-50.
14. Wang C, Ye M. Individualizing Basketball-Specific Interval Training Using Anaerobic Speed Reserve: Effects on Physiological and Hormonal Adaptations. Int J Sports Physiol Perform. 2024 Apr 1;19(4):365-74.
15. Zhu D, Song D, Huang Z. Comparing Individualized vs. Non-Individualized Locomotor Profiling on High-Intensity Interval Training Adaptations in Soccer Players: A Randomized Parallel Study. J Sports Sci Med. 2025 Sep;24(3):503-12.
16. Luo X, Zhang D, Yu W. Uniform Homeostatic Stress Through Individualized Interval Training Facilitates Homogeneous Adaptations Across Rowers With Different Profiles. Int J Sports Physiol Perform. 2024 Mar 1;19(3):232-41.
17. Dai L, Xie B. Adaptations to Optimized Interval Training in Soccer Players: A Comparative Analysis of Standardized Methods for Individualizing Interval Interventions. J Sports Sci Med. 2023 Dec;22(4):760-8.
18. Tongwu Y, Jinghui Z, Chuanwei D, Zijian Z, Yuxiong X. Supramaximal interval training using anaerobic speed reserve or sprint interval training in rowers. Front Physiol. 2025;16:1516268.
19. Du G, Tao T. Effects of a paddling-based high-intensity interval training prescribed using anaerobic speed reserve on sprint kayak performance. Front Physiol. 2022;13:1077172.
20. Aspin GL, Graham MK, Franklin JD, Hicks KM, Taylor JM. The Relationship Between the Anaerobic Speed Reserve and Acute Responses to High-Intensity Interval Training in Female Soccer Players. J Strength Cond Res. 2024 Nov 1;38(11):1911-6.
21. Julio UF, Panissa VLG, Paludo AC, Alves ED, Campos FAD, Franchini E. Use of the anaerobic speed reserve to normalize the prescription of high-intensity interval exercise intensity. Eur J Sport Sci. 2020 Mar;20(2):166-73.

22. Wang X, Zhao L. Adaptive responses of cardiorespiratory system and hormonal parameters to individualized high-intensity interval training using anaerobic power reserve in well-trained rowers. Front Physiol. 2023;14:1177108.
23. Collison J, Debenedictis T, Fuller JT, Gerschwitz R, Ling T, Gotch L, et al. Supramaximal Interval Running Prescription in Australian Rules Football Players: A Comparison Between Maximal Aerobic Speed, Anaerobic Speed Reserve, and the 30-15 Intermittent Fitness Test. J Strength Cond Res. 2022 Dec 1;36(12):3409-14.
24. Di Gennaro S, Panasci M, Gallo G, Filipas L, Ruggeri P, Faelli E. Insights into anaerobic power reserve on relationships with exercise tolerance, work above critical power, and accumulated oxygen deficit in endurance-trained male cyclists: a pilot study. Appl Physiol Nutr Metab. 2025 Jan 1;50:1-11.
25. Salter J, Forsdyke D, Dawson Z, Rymer J, Walsh L, Mundy P. The Influence of Athletic Profile on Acute Neuromuscular Response to a Repeated Sprint Protocol in Youth Females: A Pilot Study. J Strength Cond Res. 2025 Jul 1;39(7):760-9.
26. Thron M, Sondermann Y, Ruf L, van Hooren B, Duking P, Wehdemeier L, et al. Locomotor Profile Influences Physiological and Locomotor Adaptations After 6 Weeks of High-Intensity or Sprint Interval Training in Male and Female Runners. Int J Sports Physiol Perform. 2026 Mar 1;21(3):472-81.
27. Belechris NP, Bogdanis GC, Zacharogiannis E, Tsoukos A, Paradisis G. Impact of Aerobic Capacity on Mechanical Variables in Track Sprinters and Middle-Distance Runners: A Comparative Study. J Funct Morphol Kinesiol. 2025 Sep 9;10(3).
28. Girard O, Buchheit M, Goodall S, Racinais S. Oxygen availability affects exercise capacity, but not neuromuscular fatigue characteristics of knee extensors, during exhaustive intermittent cycling. Eur J Appl Physiol. 2021 Jan;121(1):95-107.
29. Fitzpatrick JF, Hicks KM, Hayes PR. Dose-Response Relationship Between Training Load and Changes in Aerobic Fitness in Professional Youth Soccer Players. Int J Sports Physiol Perform. 2018 Nov 1;13(10):1365-70.
30. Franchini E, Kons RL, Panissa VLG, Detanico D. Assessment of the Anaerobic Speed Reserve during Specific High-Intensity Exercise in Judo Athletes. Journal of Science in Sport and Exercise. 2021;4(3):266-74.
31. Kavanagh R, McDaid K, Rhodes D, Alexander J, Harper D, Oliveira R, et al. The relationship between individualised speed thresholds and changes in aerobic fitness in elite professional youth soccer players. A case study. Baltic Journal of Health and Physical Activity. 2025:Article1-Article.
32. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002 Jun 15;21(11):1539-58.
33. Senior AM, Viechtbauer W, Nakagawa S. Revisiting and expanding the meta-analysis of variation: The log coefficient of variation ratio. Res Synth Methods. 2020 Jul;11(4):553-67.
34. Nakagawa S, Poulin R, Mengersen K, Reinhold K, Engqvist L, Lagisz M, et al. Meta-analysis of variation: ecological and evolutionary applications and beyond. Methods in Ecology and Evolution. 2014;6(2):143-52.
35. Rover C, Knapp G, Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. BMC Med Res Methodol. 2015 Nov 14;15:99.
36. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020 Mar;17(3):261-72.
37. Efron B. Better Bootstrap Confidence Intervals. Journal of the American Statistical Association. 1987;82(397):171-85.
38. Thron M, Duking P, Ruf L, Hartel S, Woll A, Altmann S. Assessing anaerobic speed reserve: A systematic review on the validity and reliability of methods to determine maximal aerobic speed and maximal sprinting speed in running-based sports. PLoS One. 2024;19(1):e0296866.
39. Roe G, Darrall-Jones J, Black C, Shaw W, Till K, Jones B. Validity of 10-HZ GPS and Timing Gates for Assessing Maximum Velocity in Professional Rugby Union Players. Int J Sports Physiol Perform. 2017 Jul;12(6):836-9.
40. Zabaloy S, Freitas TT, Carlos-Vivas J, Giraldez JC, Loturco I, Pareja-Blanco F, et al. Estimation of maximum sprinting speed with timing gates: greater accuracy of 5-m split times compared to 10-m splits. Sports Biomech. 2024 Mar;23(3):262-72.
41. Midgley AW, McNaughton LR, Carroll S. Reproducibility of time at or near VO2max during intermittent treadmill running. Int J Sports Med. 2007 Jan;28(1):40-7.
42. Michalik K, Danek N, Zaton M. Assessment of the physical fitness of road cyclists in the step and ramp protocols of the incremental test. J Sports Med Phys Fitness. 2019 Aug;59(8):1285-91.
43. Balmer J, Davison RC, Bird SR. Peak power predicts performance power during an outdoor 16.1-km cycling time trial. Med Sci Sports Exerc. 2000 Aug;32(8):1485-90.
44. Mendez-Villanueva A, Bishop D, Hamer P. Reproducibility of a 6-s maximal cycling sprint test. J Sci Med Sport. 2007 Oct;10(5):323-6.
45. Coso JD, Mora-Rodriguez R. Validity of cycling peak power as measured by a short-sprint test versus the Wingate anaerobic test. Appl Physiol Nutr Metab. 2006 Jun;31(3):186-9.

46. Bentley DJ, McNaughton LR. Comparison of W(peak), VO2(peak) and the ventilation threshold from two different incremental exercise tests: relationship to endurance performance. J Sci Med Sport. 2003 Dec;6(4):422-35.
47. Berthon P, Fellmann N, Bedu M, Beaune B, Dabonneville M, Coudert J, et al. A 5-min running field test as a measurement of maximal aerobic velocity. Eur J Appl Physiol Occup Physiol. 1997;75(3):233-8.
48. Dabonneville M, Berthon P, Vaslin P, Fellmann N. The 5 min running field test: test and retest reliability on trained men and women. Eur J Appl Physiol. 2003 Jan;88(4-5):353-60.
49. Huntsman HD, DiPietro L, Drury DG, Miller TA. Development of a rowing-specific VO2max field test. J Strength Cond Res. 2011 Jun;25(6):1774-9.
50. Metikos B, Mikulic P, Sarabon N, Markovic G. Peak Power Output Test on a Rowing Ergometer: A Methodological Study. J Strength Cond Res. 2015 Oct;29(10):2919-25.
51. Cheng CF, Yang YS, Lin HM, Lee CL, Wang CY. Determination of critical power in trained rowers using a three-minute all-out rowing test. Eur J Appl Physiol. 2012 Apr;112(4):1251-60.
52. Winchcombe CE, Binnie MJ, Doyle MM, Hogan C, Peeling P. Development of an On-Water Graded Exercise Test for Flat-Water Sprint Kayak Athletes. International Journal of Sports Physiology and Performance. 2019;14(9):1244-9.
53. Bellenger CR, Fuller JT, Nelson MJ, Hartland M, Buckley JD, Debenedictis TA. Predicting maximal aerobic speed through set distance time-trials. Eur J Appl Physiol. 2015 Dec;115(12):2593-8.
54. Lundquist M, Nelson MJ, Debenedictis T, Gollan S, Fuller JT, Larwood T, et al. Set distance time trials for predicting maximal aerobic speed in female Australian Rules Footballers. J Sci Med Sport. 2021 Apr;24(4):391-6.
55. Thron M, Woll A, Klos L, Hartel S, Ruf L, Kloss C, et al. Overestimation of maximal aerobic speed by the Universite de Montreal track test and a 1500-m-time trial in soccer. Front Physiol. 2022;13:1023257.
56. Simpson LP, Kordi M. Comparison of Critical Power and W' Derived From 2 or 3 Maximal Tests. Int J Sports Physiol Perform. 2017 Jul;12(6):825-30.
57. Karsten B, Jobson SA, Hopker J, Stevens L, Beedie C. Validity and reliability of critical power field testing. Eur J Appl Physiol. 2015 Jan;115(1):197-204.
58. Shiroma SA, Julio UF, Franchini E. Criterion Validity, Reliability and Usefulness of a Judo-Specific Maximal Aerobic Power Test. Int J Sports Physiol Perform. 2019 Jul 1;14(7):987-93.
59. Beato M, Coratella G, Stiff A, Iacono AD. The Validity and Between-Unit Variability of GNSS Units (STATSports Apex 10 and 18 Hz) for Measuring Distance and Peak Speed in Team Sports. Frontiers in Physiology. 2018 2018-September-21;Volume 9 - 2018.
60. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924-6.