sports

MDPI

Review

Monitoring External Load and Internal Responses Across Selected Endurance Sports: A Scoping Review

Thomas Osdal, Terje Dalen, Tore Kristian Aune and Morten Andreas Aune *

Department of Sport Science, Faculty of Education and Arts, Nord University, 7600 Levanger, Norway; thomas.osdal@student.nord.no (T.O.); terje.dalen@nord.no (T.D.); tore.k.aune@nord.no (T.K.A.)
* Correspondence: morten.a.aune@nord.no

Abstract

Background: Endurance-sport monitoring combines external-load measures with physiological and subjective responses, but studies vary in their measurement techniques, interpretation, and links to action. This scoping review mapped how these indicators have been studied and the types of evidence linking them to training and athlete outcomes. Methods: PubMed, SPORTDiscus, Scopus, and Web of Science were searched on 24 October 2026. Eligible original research evaluated at least one non-invasive indicator on at least two days during a documented training period or endurance exposure of seven days or longer in six selected endurance sports. Age eligibility was a sample mean/median of at least 18 years or explicit adult recruitment. Findings were summarized descriptively without meta-analysis or formal critical appraisal. Results: Of 8398 screened records, 83 database-derived reports and six reports from other methods were assessed in full, and 56 reports were included. Of these, 15 reports concerned training-load quantification, 32 examined autonomic or physiological responses, and 9 studied subjective responses. By evidence type, 13 examined measurement properties or response behavior, 20 considered concurrent or longitudinal associations, 12 assessed prospective relationships, and 11 examined monitoring-informed action rules. Relationships varied with calibration, measurement protocol, time window, sport, and outcome. Conclusions: The mapped comparisons do not establish broad transferable action thresholds or comparative effectiveness. Future research should evaluate explicit monitoring-and-action rules against adaptations, performance, and health outcomes.

Keywords: training load; athlete monitoring; endurance sport; heart rate variability; rating of perceived exertion; recovery; performance; decision-making

Check for updates

Academic Editor: César Cavinato Cal Abad

Received: 7 October 2026
Revised: 3 November 2026
Accepted: 5 November 2026
Published: 9 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Training monitoring in endurance sport has expanded from time, distance, and coach observation to continuous physiological, biomechanical, and perceptual data. Power meters, global positioning systems, accelerometers, heart-rate monitors, wearable sensors, and short self-report scales can describe different parts of training exposure and response. Their availability, however, does not guarantee that the resulting values have the same meaning or support the same decision [1–5].

The distinction between external load and internal response is central [1,2,6]. External load describes the work completed, including duration, distance, speed, power, elevation, and mechanical exposure. Internal responses describe how the athlete responds under the prevailing conditions and are assessed through physiological measures, such as heart rate and heart-rate variability (HRV), and perceptual measures, including ratings of perceived

exertion and questionnaire-derived ratings of perceived recovery, wellness, and fatigue. The same external work can produce different responses between athletes or within the same athlete across sleep, illness, environmental, and training contexts. Conversely, similar measured internal responses can arise from different modes and mechanical demands [4]. Such similarity does not establish equivalent physiological demands: the selected indicators capture only part of the biological response, and their sensitivity depends on the measurement procedure.

A monitoring score can therefore answer several different questions. It may summarize exposure, represent a physiological or perceptual construct, show acceptable repeatability, correspond with another measure, associate with a later outcome, or improve outcomes when used to change training. These claims require different study designs. Reliability or correlation does not by itself establish adaptation, readiness, or an appropriate action [3–5,7,8].

Previous reviews have examined HRV-guided endurance training and perceived exertion in continuous sports [9,10]. Düking et al. [9] compared HRV-guided and predefined training, Torres et al. [10] mapped perceived exertion, and Bellenger et al. [11] examined autonomic changes with endurance-performance change. The present review maps evidence across external-load, physiological, and subjective domains. This cross-domain approach addresses a practical question for coaches and athletes: what can monitoring indicators tell us about training responses, and to what extent does the available evidence support their use in guiding training decisions?

Aim and Research Questions

This review aimed to map how external-load and internal-response indicators have been operationalized and evaluated in samples meeting the age criterion from six selected endurance sports, and to examine to what extent the available evidence supports their interpretation and use in monitoring-informed decisions. This descriptive map does not estimate common effects across sports, sexes or competitive levels.

The present review addresses two research questions:

1. Which external-load and internal-response indicators have been used across the selected endurance-sport contexts, how have they been measured, and what types of evidence link them to training and athlete outcomes?
2. What methodological characteristics, limitations, and evidence gaps in the included reports constrain the interpretation of these indicators and their use in training decisions?

2. Materials and Methods

2.1. Design

A scoping-review design was used to map concepts, methods, and gaps across a heterogeneous body of evidence rather than estimate a pooled intervention effect. Reporting was structured according to PRISMA-ScR [12], and methodological guidance informed the conduct for scoping reviews [13]. No protocol was prospectively registered.

Eligibility was organized using the Population–Concept–Context framework. The population comprised participants meeting the age criterion described in Section 2.2. The concept comprised non-invasive external-load or internal-response indicators evaluated during an ongoing training process. The context comprised running, cycling, swimming, triathlon, cross-country skiing, and rowing. These six contexts provide a pragmatic map across different mechanical demands, not an exhaustive definition of endurance sport. Road, trail and ultra-running were eligible; kayaking, biathlon and speed skating were outside the scope. Habitual sport and monitored exercise mode were distinguished.

2.2. Eligibility Criteria

Reports were included when all the following criteria were met:

- At least one external-load or internal-response indicator was evaluated, and at least one eligible indicator was non-invasive. Blood-derived measures could be reported alongside an eligible non-invasive indicator. The review focused on non-invasive indicators because these can be collected repeatedly in routine training settings and are therefore particularly relevant to ongoing monitoring-informed decisions.
- Participants trained or competed in running, cycling, swimming, triathlon, cross-country skiing, or rowing.
- The reported mean or median participant age was at least 18 years, or recruitment was explicitly restricted to adults. Studies with participants of unconfirmed age were excluded; elite or senior status alone was insufficient. This sample-level rule permits mixed-age samples.
- The report presented original data.
- The indicator was assessed on at least two separate days across a documented training period or repeated endurance exposure of at least seven days, and the study either related the indicator to adaptation, performance, fatigue, recovery, or injury or evaluated its measurement properties, comparability, or interpretability within that training process. A minimum duration of seven days was selected to distinguish monitoring across an ongoing training period from repeated measurements confined to a single acute exercise-recovery episode. This operational boundary is not a physiological cutoff; standalone repeated tests without intervening training did not qualify.

Clinical populations, acute single-session studies without longitudinal linkage, reports focused solely on agreement with a reference device, and ineligible publication types were excluded. The clarified criteria were applied retrospectively to every full-text decision. The final sample-level age rule was confirmed on 27 October 2026 and changed five decisions, including four exclusions for unconfirmed age and the inclusion of Dai et al. [14] (Supplementary Materials, Table S4).

2.3. Information Sources and Search Strategy

T.O. developed and ran database searches. The final search was run on 24 October 2026 in PubMed (NLM/NCBI), SPORTDiscus with Full Text (EBSCOhost), Scopus (Elsevier; TITLE-ABS-KEY), and Web of Science Core Collection (Clarivate; Topic fields), from database inception. No date, language, document-type, or automatic adult filters were applied. The complete verbatim strategies and interface settings are presented in Supplementary Table S1.

Search strings combined the endurance-sport context with external-load, physiological, subjective, wearable, and performance indicators. Specific monitoring indicators required a sport term; general terms additionally required training-load, monitoring, prescription, adaptation, or training-status terms to limit unrelated and single-session results. Syntax was adapted to each platform.

MeSH terms were added in PubMed. Scopus TITLE-ABS-KEY and Web of Science TS included indexed keywords; these fields are not controlled thesauri. SPORTDiscus included subject terms with equivalent-subject and related-word expanders enabled, without manually added thesaurus headings. Complete strategies and interface settings are provided in Supplementary Table S1.

The searches retrieved 4409 PubMed, 3409 SPORTDiscus, 6465 Scopus and 5452 Web of Science records (19,735 in total; Supplementary Table S2).

Known studies informed search-term development and refinement (Supplementary Section S1; Table S1b). Because they influenced the strategy, these development checks do not independently estimate search sensitivity.

A PubMed retrieval check on 28 October 2026 retrieved all 47 located PubMed records among the 53 database-included reports; six had no located record. The ten controls were a subset of these 47. Report-level results are provided in the Supplementary Materials, Table S1a. The 4430 rerun results were not an additional PRISMA batch. Known-item retrieval does not establish comprehensive sensitivity.

Cross-checking by DOI/title retrieved all eight reports in Düking et al. [9] and 147 of the 232 reports in Torres et al. [10]. The denominator included ineligible single-session studies, so the retrieval proportion was not interpreted as an estimate of search sensitivity. T.O. screened the unmatched reports; potentially eligible full texts were assessed by T.O. and M.A.A. using the same criteria (Supplementary Table S3). No systematic backward or forward citation search was performed.

2.4. Record Management and Screening

Records were managed in the Rayyan web application (Rayyan Systems, Inc., Cambridge, MA, USA). The build number was not recorded. The Systematic Auto Resolver was applied at an overall similarity threshold of 95%, removing 10,222 records. Further automatic rounds and manual inspection removed 1115 records; their criteria were not recorded. In total, 11,337 duplicate records were removed, leaving 8398 records for screening. All exported records were then reconciled with the screening set, and one eligible report removed in error was reinstated. Reconciliation details and 33 residual repeated DOI records are documented in Supplementary Section S1; the screened total is a record count.

T.O. and M.A.A. screened titles and abstracts. Records were advanced to full-text assessment when either reviewer judged them potentially eligible. T.O. advanced 80 records, and 5 records that T.O. had excluded were advanced after screening by M.A.A., leading to 85 reports being sought for retrieval and 8313 records being excluded at the title/abstract level. Independent duplicate title/abstract screening was not performed; paired decisions were unavailable for calculating inter-reviewer agreement.

M.A.A. sought unavailable full texts through SPORTDiscus and Oria (Nord University Library). Four were obtained, including the full article corresponding to a conference abstract, and two remained unavailable. Both reviewers independently assessed all 83 retrieved database reports; 30 were excluded and 53 were included. Any uncertainties were resolved through discussion concerning the age criterion. At the full-text stage, one primary reason for exclusion was recorded for each excluded report (Supplementary Tables S4 and S5). Paired decisions were unavailable for calculating agreement. The Spanish-language report [15] was translated with AI assistance (Section 2.6). The retained record does not separately document how non-English titles and abstracts were assessed.

2.5. Data Charting and Synthesis

Data were charted by T.O. using a structured form developed for the review. The charted information included author, publication year, sport or monitored exercise mode, participant characteristics and sample size, study design, observation period, indicators and measurement protocols, study purpose, principal findings, and relevant interpretive limitations. Where applicable, information concerning reference periods, monitoring-informed action rules, attrition, and potentially overlapping participant cohorts was also recorded. Each report was assigned to one primary analytical domain for descriptive counting: quantification of training load, autonomic and physiological responses, or subjective responses. Because several reports included indicators from more than one domain, indicator over-

laps were retained in the narrative synthesis. NR denotes information not reported in the examined full text.

Each report was also assigned one primary evidence type: measurement properties or response behavior, concurrent or longitudinal association, prospective relationship, or monitoring-informed action rule. Primary analytical domain and evidence-type assignments were initially made by T.O. as part of data charting and were subsequently reviewed and confirmed by M.A.A. and T.D. Any differences in interpretation were resolved through discussion. When a report contributed to more than one evidence type, the primary category prioritized monitoring-informed action rules, prospective relationships, associations, and measurement properties or response behavior, in that order. Change–change analyses and retrospective responder comparisons were classified as associations. Assigning one primary evidence type per report avoided double counting. The categories did not represent a hierarchy of study quality. Independent duplicate coding and calculation of agreement statistics were not performed (Supplementary Table S6).

Publications were counted as reports, and reports arising from related intervention programs or potentially overlapping cohorts were not treated as independent replications. Funding and related reports are provided in Supplementary Table S7. Confirmed shared-study families A–C were combined; possible links P1–P3 remained separate, with sensitivity scenarios presented in Supplementary Table S8. Study groups do not establish independent participant cohorts.

Concurrent associations were not interpreted as prospective predictions, and evaluations of complete monitoring-informed action rules were not treated as validations of each individual indicator included in the intervention. Results were summarized descriptively across indicator domains, study designs, measurement protocols, time horizons, outcomes, and evidence types. No meta-analysis, formal critical appraisal, risk-of-bias assessment, certainty rating, or ranking of instruments was undertaken because the purpose of the review was to map the scope, characteristics, and type of available evidence rather than estimate pooled effects or determine comparative effectiveness. Charted limitations do not constitute formal appraisal [12].

RQ1 was synthesized by indicator, protocol, outcome, and evidence type, while RQ2 was synthesized by study design, attrition, overlap, analysis, and reporting uncertainty (Supplementary Tables S6 and S9). Extraction was not exhaustive. The PRISMA-ScR checklist is in Supplementary Table S10.

2.6. Use of Artificial Intelligence

ChatGPT-4 (OpenAI, San Francisco, CA, USA) was used to support language editing, consistency checks, restructuring, reference checks, document preparation, and translation of the included Spanish-language report. It did not replace the database search or author responsibility for eligibility decisions, extracted information, interpretation, or the final manuscript. The exact model/version for historical editing sessions could not be established from the retained revision records. The authors retain responsibility for all content.

3. Results

3.1. Study Selection

After removal of 11,337 duplicates, 8398 records were screened. Of these, 8313 were excluded at the title/abstract level and 85 reports were sought for retrieval. Two could not be obtained, leaving 83 reports for full-text assessment. Thirty were excluded for the following reasons: ineligible publication type ($n = 2$), age criterion ($n = 6$), ineligible sport or population ($n = 7$), absence of a monitored training period ($n = 13$), and no eligible

analytical link ($n = 2$). Fifty-three database-derived reports were included (Figure 1; Supplementary Tables S2, S4 and S5).

Six reports identified through the cross-check against Torres et al. [10] were also assessed in full; three were excluded and three were included [16–18], giving fifty-six included reports in total. Plews et al. [19] contributes only its eligible triathlete analysis.

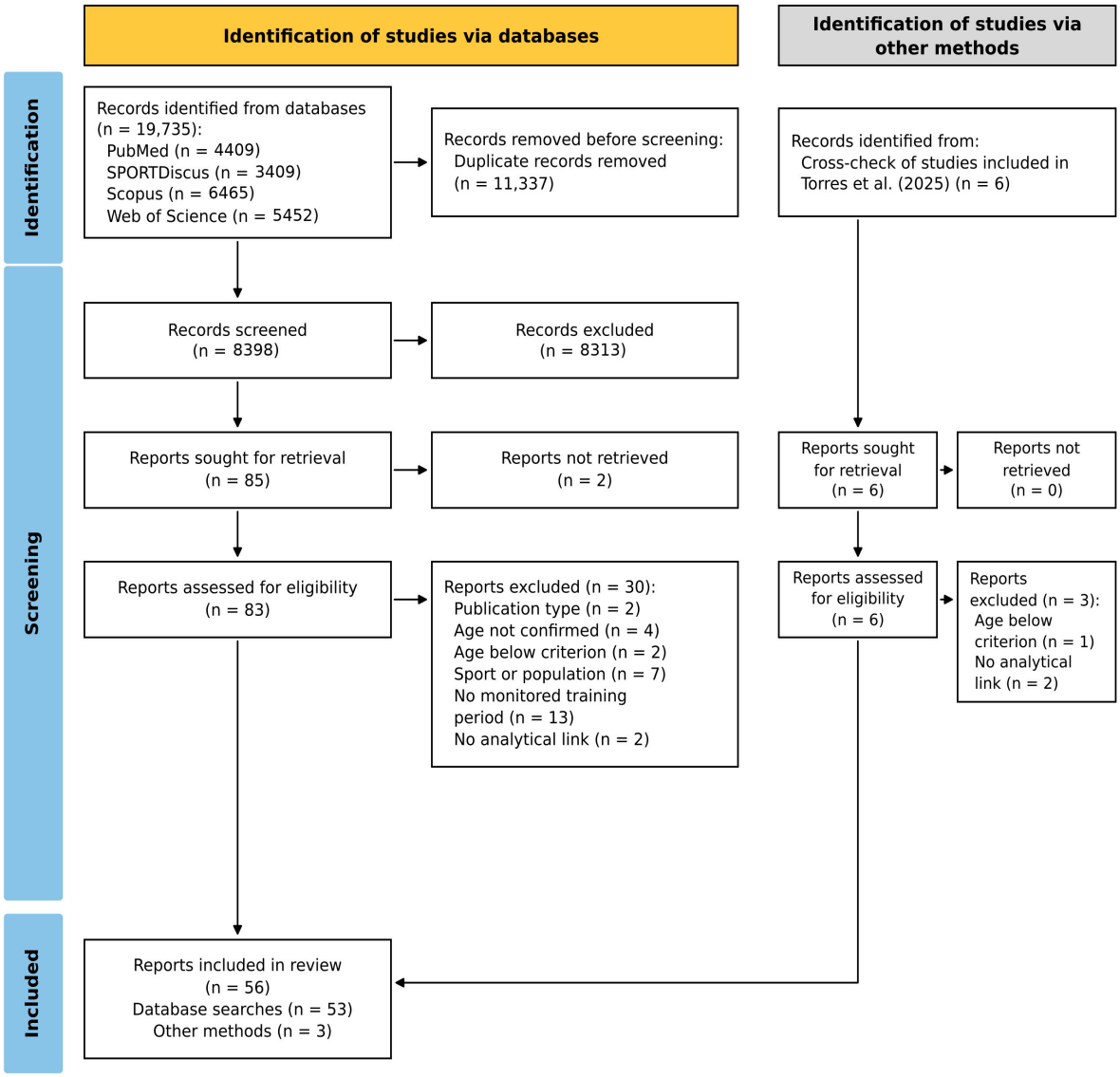


Figure 1. PRISMA 2020 flow diagram [20] for the four-database search and the cross-check of Torres et al. [10], within PRISMA-ScR reporting [12]. Counts are reported at the record or report level as labeled.

3.2. Characteristics of the Included Evidence

The 56 included reports were published from 1989 to 2026. Of these, 15 primarily concerned training-load quantification, 32 concerned autonomic or physiological responses, and 9 assessed subjective responses. The primary contexts were running (26), cycling (10), triathlon (6), rowing (5), swimming (5), cross-country skiing (2), and two mixed-sport contexts (one report each). Observation periods ranged from three weeks to competitive seasons and longer programs. Table 1 summarizes the included reports.

Table 1. Characteristics, age, design, selected indicators, and primary evidence type of the 56 included reports. Shared study and possible overlap markers are explained below.

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
Training-load quantification					
Dai X, Yan J, Bi X (2025a) [21] [P2]	Rowing, elite male rowers ($n = 30$)	22.1 ± 3.6	Longitudinal observational study; 48-week monitoring period (1196 sessions)	External: duration, stroke rate, stroke count, 500 m split time, % of boat speed; internal: TRIMP, session-RPE training load	Concurrent or longitudinal association
Gallo G, Bosio A, Martin M, Morelli A, Azzolini M, Guercilena L, Larrazabal J, Rampinini E (2023) [22]	Cycling; 46 male WorldTour professional cyclists	29.6 ± 5.0	Observational analysis; four seasons (2016–2019), with monthly and 8-week pre-race exposure windows	Time, Edwards TRIMP, Training Stress Score, time in power zones Z1–Z3, polarization index; record power outputs over 1, 5, 20 and 40 min	Prospective relationship
Manzi V, Iellamo F, Impellizzeri F, D'Ottavio S, Castagna C (2009) [23]	Running; 8 recreational long-distance runners	39.9 ± 6.5	Longitudinal observational dose–response study; 8 weeks	Individualized TRIMP (TRIMPi), conventional TRIMP, and individual HR–blood-lactate profiles; 5 km/10 km performance and lactate-threshold running speed	Prospective relationship
Manzi V, Bovenzi A, Castagna C, Sinibaldi Salimei P, Volterrani M, Iellamo F (2015) [24]	Running, long-distance recreational athletes ($n = 7$)	36.5 ± 3.8	Longitudinal observational study; 5 months culminating in a marathon	Training zones from HR at 2.0 and 4.0 mmol/L, TRIMPi, session-RPE	Concurrent or longitudinal association
Napier C, Ryan M, Menon C, Paquette MR (2020) [25]	Running, recreational runners ($n = 68$)	40.1 ± 10.4	Longitudinal observational case series; 7 months	Running time, steps, cumulative shock (accelerometry); session-RPE and the products timeRPE, stepsRPE, shockRPE	Measurement properties or response behavior
Rodríguez-Marroyo JA, Villa JG, Fernández G, Foster C (2013) [26]	Cycling, professional cyclists ($n = 12$)	25 ± 1 (mean ± SEM)	Observational repeated-measures study; two 21-day Vuelta a España editions	Heart rate, session-RPE, TRIMP-HR, TRIMP-RPE across time-trial, flat, medium- and high-mountain stages	Concurrent or longitudinal association
Sanders D, Abt G, Hesselink MKC, Myers T, Akubat I (2017a) [27] [P1]	Cycling, well-trained competitive cyclists ($n = 15$)	22 ± 2.5	Longitudinal observational study; 10 weeks	TSS (power-based); bTRIMP, eTRIMP, iTRIMP, luTRIMP, sRPE	Prospective relationship
Sanders D, Myers T, Akubat I (2017b) [28] [P1]	Cycling, well-trained road cyclists ($n = 15$)	22 ± 2.5	Longitudinal observational measurement comparison; 10 weeks	Training-intensity distribution in three zones from RPE, HR and power output	Measurement properties or response behavior

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Training-load quantification		
Schumann M, Botella J, Karavirta L, Häkkinen K (2017) [29]	Running; 30 men recruited; 24 matched into training groups (10 load-guided, 14 standardized)	34 ± 7	Matched-group intervention; 24 weeks (12 common + 12 guided/standardized); systematic matching documented; random allocation not documented	Polar proprietary cumulative-load score guides training; orthostatic HRV (5 min supine + 3 standing) is an outcome, not the guidance rule; TRIMP (training exposure) and running performance	Monitoring-informed action rule
Sixsmith et al. (2023) [16]	Swimming; 10 national-to-international swimmers (8 men, 2 women)	22.0 ± 2.3	Longitudinal observational study; 6 months of daily monitoring (January–June 2021), including three training camps	Time in HR zones (Polar OH1, PPG at the temple) versus planned volume at set intensities; coach-planned and athlete-reported sRPE (modified CR10)	Measurement properties or response behavior
Talsnes RK, van den Tillaar R, Cai X, Sandbakk Ø (2020) [30]	Cross-country skiing; 24 endurance-transfer athletes recruited, 23 with training data; high/low-responder comparison $n = 9/11$	19 ± 2	Single-cohort training study with retrospective responder groups; 6-month ski-specific program	Daily training volume, intensity and mode; session-RPE load; performance tests; injury/illness days	Concurrent or longitudinal association
Vanwanseele B, Op De Beéck T, Schütte K, Davis J (2020) [31]	Running; 96 outdoor participants stated; Table 2 demographics $n = 59$ (41 men, 18 women), denominator unclear †; 9 indoor validation participants	Outdoor recruitment 18–60. Table 2 demographics 29.5 ± 8.1 ($n = 59$; not established for all 96). Indoor age NR †	Measurement-validation study: laboratory validation ($n = 9$) plus 3-month observational field monitoring ($n = 96$; 498 sessions)	Trunk-mounted accelerometry (cumulative load as sum of peak vertical accelerations), GPS duration, distance, speed, step count; instrumented-treadmill GRF	Measurement properties or response behavior
Vermeire KM, Vandewiele G, Caen K, Lievens M, Bourgois JG, Boone J (2021) [32]	Cycling, recreational cyclists ($n = 11$)	40.1 ± 6.4	Longitudinal observational study with self-directed training; 12 weeks	Banister TRIMP, Edwards TRIMP, Lucia TRIMP, individualized TRIMP, training intensity distribution	Prospective relationship
Wallace LK, Slattery KM, Coutts AJ (2014) [33]	Running, trained ($n = 7$)	38.6 ± 9.4	Longitudinal observational modeling study; 15 weeks	rTSS; sRPE, Banister TRIMP, submaximal HR, HRV (SD1), POMS-fatigue	Concurrent or longitudinal association

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Training-load quantification		
Dai et al. (2025c) [14] [P2]	Rowing, elite male open-class rowers (n = 30); recruitment > 16 y	21.6 ± 3.4; individual minimum/maximum NR	Longitudinal validation/reliability study; formal investigation 12 weeks; abstract states 12 months †	Borg CR10 session-RPE × duration (within 30 min); Polar Team Pro HR; Banister TRIMP; load measures across ergometer, strength and functional sessions	Concurrent or longitudinal association
			Autonomic and physiological responses		
Alfonso C, Clarke DC, Capdevila L (2025) [34]	Cycling; 116 eligible participants allocated; 28 male completers analyzed (G1 n = 8; G2 n = 12; G3 n = 8)	G1 39.6 ± 10.0; G2 47.7 ± 15.7; G3 49.5 ± 10.3. Ranges: 29–55; 27–69; 37–65	Three-arm randomized training intervention as reported; 40-day protocol; complete-case analysis with substantial attrition	Vagally mediated HRV, resting HR, subjective well-being scores; Pmax, 1, 5, and 20 min power, FTP	Monitoring-informed action rule
Anderson T, van Mourik RA, Martin KJ, Eijsvogels TMH, Longoria KA (2023) [35]	Triathlon; one male ultraendurance athlete	44 (individual age)	Observational case study; 100-day exposure; HR/HRV/SpO2 available on 74 days	Night-time HRV (PPG; exact metric/transform NR), resting HR, SpO2, sleep and daily performance	Prospective relationship
Bulte et al. (2025) [36] †	Swimming; 11 competitive swimmers (4 women, 7 men); 8 middle-distance, 2 distance, 1 sprint	Women 20 [2]; men 23 [2] (mean [SD])	Retrospective repeated-measures study; 5 months; Daily morning HRV and recorded training exposure	Daily waking 1 min supine HRV4Training PPG; raw rMSSD, AVNN and rMSSD/AVNN; 7-day means and 3-day taper averages; zone-weighted swimming distance	Concurrent or longitudinal association
Carrasco-Poyatos et al. (2022) [37]	Running, professional; 13 assessed, 12 randomized and completed (6/group); sex NR	HRV-guided 31.5 ± 8.36 (25–46); traditional 31.2 ± 10.3 (19–42)	Cluster-randomized intervention; 12-week schedule (4 preparation + 8 intervention weeks); tests in weeks 5 and 12	Daily morning 60 s supine HRV4Training/rMSSD; weekly lnRMSSD and CV; linked protocol reference mean ± 0.5 SD, updated weekly; training time/intensity, VO2max, velocity and HR.	Monitoring-informed action rule
Clark JR, McKune AJ, Wood PS (2025) [38]	Rowing; 47 recruited, 43 analyzed (30 men, 13 women); national/international level	22.9 ± 3.8	Observational repeated test reliability study; three consecutive weeks	Power output (PO), test time; submaximal HR response (HRex), RPE, heart-rate recovery (HRR), PO:HRex and PO:RPE indices	Measurement properties or response behavior

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Autonomic and physiological responses		
Coates et al. (2018) [39]	Cycling/triathlon; 21 recruited; 17 analyzed (7 control, 10 overload)	36 ± 10 (21 recruited)	Block-randomized controlled intervention; 4 weeks (1 reduced + 3 regular or overload weeks)	Training duration and HR-zone distribution (Polar A300); resting HRV (SDNN, RMSSD, LF, HF), cardiac baroreflex sensitivity, MSNA (microneurography); POMS-2; peak power output, VO2max	Measurement properties or response behavior
Davletyarova K, Vacher P, Nicolas M, Kapilevich LV, Mourot L (2022) [40]	Swimming, well-trained swimmers ($n = 14$)	18.5 ± 1.6	Longitudinal repeated-measures association study; 4 weeks with increasing training load; randomization not reported in abstract	Daily external/internal training load, summed weekly; Saturday-morning HRV: MeanRR, RMSSD, LF, HF, SD1, SD2, SD1/SD2; mean HR	Concurrent or longitudinal association
Du C, Xu H, Shi J, Wei J, Quan H, Zhu T (2026) [41]	Cross-country skiing; 12 men; mixed adolescent/adult sample	18.83 ± 2.72; observed range 16–22	Longitudinal observational study; 3-week summer training period	Banister TRIMP (Firstbeat); resting HR; 3 min supine morning quick recovery test (QRT%, proprietary HRV frequency-domain-derived index); VO2max	Concurrent or longitudinal association
Fuller et al. (2017) [42]	Running, trained (10 men)	36 ±8	Within-subject intervention; 7 days light, 14 days heavy and 10 days light training	Stride DFA; TRIMP; DALDA; test RPE	Concurrent or longitudinal association
Fuller JT, Doyle TLA, Doyle EW, Arnold JB, Buckley JD, Wills JA, Thewlis D, Bellenger CR (2024) [43]	Running, trained (15 enrolled; 9 analyzed: 8 men, 1 woman)	37.8 ± 4.0	Single-group overload-and-taper intervention; 7 days light training, 14 days heavy training and 10 days taper	DFA-alpha (stride-interval correlations via accelerometry); Short Recovery and Stress Scale (SRSS); exercise-test RPE	Concurrent or longitudinal association
Iguchi et al. (2025) [44]	Rowing, collegiate male rowers ($n = 11$)	21.2 ± 0.8; range 20–22	Longitudinal observational repeated-measures study; four 7-day monitoring windows from October to January	Session-RPE training load; accelerometry-based external load indicators (ABELI); actigraphy sleep (total sleep time, efficiency, latency, WASO, awakenings), PSQI, sleep log; salivary cortisol and testosterone	Concurrent or longitudinal association

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
Autonomic and physiological responses					
Javaloyes et al. (2019) [45] [P3]	Cycling, well-trained male road cyclists (n = 17; HRV-guided n = 9, traditional periodization n = 8)	HRV 39.22 ± 5.33; traditional 37.62 ± 7.09	Randomized two-group intervention; 4 baseline + 8 training weeks, plus 3 evaluation weeks (15-week schedule)	Morning LnRMSSD (chest strap and smartphone application, 90 s supine; 7-day rolling average vs smallest worthwhile change); weekly training time and time in zones; VO2max, peak power output, ventilatory thresholds, 40 min time trial	Monitoring-informed action rule
Javaloyes et al. (2020) [46] † [P3]	Cycling; 20 recruited; 15 completed; 1 injury and 4 non-compliance exclusions; group n conflict †	Recruited range 18–46. Table 1: BP 30.8 ± 10.5; HRV 28.1 ± 13.2; group n labels conflict †	Matched two-group intervention; random allocation not documented; 2 baseline + 8 training + 3 evaluation weeks	Morning supine PPG: 90 s, last 60 s analyzed; 7-day lnRMSSD versus mean ±0.5 SD; HR-zone TRIMP † and TrainingPeaks	Monitoring-informed action rule
Kiviniemi et al. (2007) [47]	Running, recreational male runners (n = 30 randomized; n = 26 analyzed: predefined n = 8, HRV-guided n = 9, control n = 9)	Predefined 32 ± 5 (26–38); HRV 31 ± 6 (22–40); control 35 ± 8 (24–45)	Three-group randomized controlled training intervention; 10-day baseline and 4 weeks of training	Daily morning HF power of R–R intervals (standing, Polar S810i) against a moving reference (10-day mean − SD); training HR and running velocity (Polar S625X); VO2peak, maximal running velocity (Loadmax), ventilatory threshold	Monitoring-informed action rule
Le Meur et al. (2013) [48] [C]	Triathlon; 24 men recruited; 21 analyzed (13 overreached, 8 controls)	Overload 30 ± 5; control 32 ± 8	Randomized overload versus usual training; five-week HRV window (1 baseline, 3 overload, 1 taper)	Daily orthostatic HR/HRV, HR-based training exposure, fatigue and running performance	Measurement properties or response behavior
Morinaga H, Takai Y (2025) [49]	Cycle-ergometer intervention; 12 endurance-trained men with a running background (De Pauw level 2)	20.4 ± 1.8	Two-group intervention; 5 weeks (1 common familiarization week + 4 weeks of HRV-guided or block-periodized training)	Wmax, WLT, WOBLA; lnRMSSD (5 min supine on waking; 7-day moving average), TRIMP, HRres	Monitoring-informed action rule
Nieto-Jiménez C, Ruso-Álvarez J, Pardos-Mainer E, Naranjo Orellana J (2020) [15]	Triathlon, single case, Ironman athlete (n = 1)	46	Longitudinal observational single-athlete case study; 7 months as reported, 195 recorded days; race on day 179	External training load; daily morning RMSSD and stress score (5 min HRV on waking)	Concurrent or longitudinal association

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Autonomic and physiological responses		
Nuuttila et al. (2017) [50] †	Running; 32 men recruited; 24 analyzed (13 HRV-guided, 11 predetermined) †	Reported range 19–37; HRV 29 ± 4; predetermined 31 ± 5	Matched-pair randomized two-group intervention, 3-week control + 8-week training period	Morning quick recovery test (3-day rolling average vs individual reference); nocturnal HRV (RMSSD, LF, HF, TP) and HR; training HR, GPS and time in zones; serum testosterone and cortisol; Vmax, 3000 m, CMJ	Monitoring-informed action rule
Nuuttila OP, Nummela A, Häkkinen K, Seipäjärvi S, Kyröläinen H (2021) [51]	Predominantly running; alternative endurance modes permitted; 42 recruited, 30 analyzed (16 men, 14 women; intensity n =13, volume n =17)	Intensity group 38 ± 4; volume group 36 ± 6	Matched non-randomized two-group intervention after a common control period; 20 weeks (10 habitual-training + 10 modified-training weeks)	HR–running-speed index, CMJ; nocturnal lnHF and HR (Firstbeat Bodyguard 2; 3 nights/week; 30 min–4 h 30 min after bedtime); daily session RPE and perceived recovery (0–10)	Concurrent or longitudinal association
Nuuttila OP, Nummela A, Korhonen E, Häkkinen K, Kyröläinen H (2022) [52]	Running; 40 recruited, 30 analyzed (15 men, 15 women; predefined 14, individualized 16); 10 km analysis n =29	Predefined 34 ± 7; individualized 37 ± 7	Matched-pair randomized two-group intervention; 15 weeks (3 preparation + 6 volume + 6 interval weeks)	HR–running-speed index; nocturnal 3-day lnRMSSD mean versus a 4-week reference; perceived fatigue/recovery and muscle soreness (1–7)	Monitoring-informed action rule
Nuuttila OP, Kyröläinen H, Kokkonen VP, Uusitalo A (2024a) [53] [B]	Running; 32 recruited, 24 analyzed (10 women)	Range 28–48	Within-participant measurement comparison during a single-group training intervention; 5-week analysis (3 baseline + 2 overload) within a 6-week protocol	Morning versus nocturnal lnRMSSD and HR (orthostatic test versus PPG-based watch); Lucia TRIMP; session-RPE load	Measurement properties or response behavior
Nuuttila OP, Uusitalo A, Kokkonen VP, Weerarathna N, Kyröläinen H (2024b) [54] [B]	Running, recreational (n = 24)	39 ± 6	Single-group training intervention with retrospective response groups; 6 weeks (3 baseline + 2 overload + 1 recovery)	HR–running-economy index; lnRMSSD (nocturnal, 4 h window), nocturnal HR, orthostatic test, subjective readiness, RPE; Lucia TRIMP	Measurement properties or response behavior

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Autonomic and physiological responses		
Nuuttila OP, Schäfer Olstad D, Martinmäki K, Uusitalo A, Kyröläinen H (2025) [55] [B]	Running; 32 recruited, 24 analyzed; abstract/Methods: 10 women; Table 1 sex labels: 14 women/10 men †; HR/HRV n =23	38.6 ± 6.5	Analysis of sleep and recovery during a single-group intervention; 6 weeks (3 baseline + 2 overload + 1 recovery)	Wearable (Polar Vantage V2); nocturnal HRV (RMSSD), sleep parameters, ANS charge, subjective recovery; Lucia TRIMP	Concurrent or longitudinal association
Plews et al. (2012) [56]	Triathlon, elite triathletes, case comparison (n = 2)	Man 22; woman 20	Longitudinal observational case comparison; 77 days	Daily morning LnRMSSD (7-day rolling average and coefficient of variation), resting HR; weekly training volume; psychometric questionnaire	Prospective relationship
Plews et al. (2014a) [19] [C]	Triathlon; 20 trained triathletes (13 overload, 7 control). Only this separable analysis is included; the 10-runner analysis is excluded for unverified age.	Triathletes 32 ± 8	Secondary methodological analysis of a 5-week intervention: 1 baseline, 3 overload, 1 taper week; random resampling of HRV measurement days	Daily morning lnRMSSD averaged across 1–7 randomly selected days per week; incremental track running performance	Concurrent or longitudinal association
Ranieri LE, Casado A, Martin D, Trujillo-Colmena D, Gil-Arias A, Kenneally M, Jiménez A (2025) [57]	Running; allocation involved 33 participants; 28 analyzed: HR 9, race pace 9, HRV 10; 5 excluded for <90% adherence (3 HR, 2 race pace)	Analyzed: 22 men 40.95 ± 6.35; 6 women 30.83 ± 7.49; explicit adult inclusion	Three-arm randomized training intervention; 4 preparation + 6 intervention weeks	Morning 1 min supine HRV4Training PPG; rMSSD/7-day lnRMSSD and app score as described; HR/GPS zones; Banister TRIMP; Borg 6–20 RPE × minutes	Monitoring-informed action rule
Stanley et al. (2015) [58]	Triathlon, elite male long-course triathlete (n = 1)	28	Longitudinal observational single-athlete case study; 32-week competitive season	Daily morning resting HR, LnRMSSD and LnRMSSD:RR; training load from power meter and GPS	Prospective relationship

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Autonomic and physiological responses		
Vesterinen et al. (2013) [59]	Running; 28 male recreational runners recruited, 25 analyzed	36 ± 6 (recruited sample)	Single-group training intervention; 28 weeks (14 basic + 14 intensive weeks)	Nocturnal HRV (SDNN, RMSSD, LF, HF, TP) and HR; TRIMP, HR zones, session RPE, GPS distance; serum testosterone and cortisol; Vpeak, VO2max, marathon performance	Prospective relationship
Vesterinen et al. (2016a) [60]	Running; 40 recreational endurance runners recruited (20 women, 20 men); 35 completed	20–55 (reported range)	Single-cohort training intervention with retrospective cluster analysis; 18-week study (8 low-intensity + 8 intensive weeks; maximal tests at weeks 0, 9 and 18)	Weekly submaximal running test (running speed at 70, 80 and 90% HRmax; post-exercise HR recovery); HR monitoring of all training; VO2max, peak running speed, lactate-threshold speeds	Concurrent or longitudinal association
Vesterinen et al. (2016b) [61]	Running; 40 recruited (20 women, 20 men); 31 completed (13 HRV-guided, 18 traditional)	Women 34 ± 8; men 35 ± 7	Matched-pair randomized intervention; 4-week preparation + 8-week intensive training	Morning RMSSD, 7-day average and individual SWC; HR/GPS/RPE training records	Monitoring-informed action rule
Barsumyan et al. (2026) [62]	Cycling, seven male competitive endurance cyclists; 590 usable athlete-days	32.6 ± 6.8	Observational longitudinal forecasting study; 16 consecutive weeks; expanding-window within-athlete one-step-ahead validation	Nocturnal native-scale rMSSD (3 WHOOP, 4 Garmin); power-meter mechanical work (kJ); Borg CR10 RPE; 1/4/7/10/14-day lags; ARX, SVR, XGBoost	Prospective relationship
Rothschild et al. (2024) [63]	Mixed triathlon/running/cycling/rowing; 55 recruited, 43 analyzed; 8 excluded for <6 h/week, 4 for <85% logs	Recruitment ≥18; recruited 42.6 ±9.1; analyzed-subset summary NR	Longitudinal observational modeling; planned 12 weeks; 11 analyzed participants completed 6 to <12 weeks	Daily lnRMSSD from waking apps/nocturnal wearables; CR100 RPE × minutes/10; 100-point recovery rating; sleep/diet; no external power/pace	Prospective relationship
			Subjective responses		
Bertochi GFA, Mota GR, Haile L, Marocolo M, Sasaki JE (2023) [64] [A]	Running; 90 screened, 71 eligible, 53 analyzed (22 men/31 women); distance analyses n =52	37.92 ± 8.95; recruitment 20–59	Longitudinal observational validity study; 4 weeks	Weekly/monthly modified CR10 rating × corresponding duration; aggregated Banister TRIMP; own optical/chest-belt HR devices	Concurrent or longitudinal association

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Subjective responses		
Bertochi GFA, Sasaki JE (2025) [65] [A]	Running, recreational (n = 52)	20–59 (recruitment criterion); sample age summary NR	Secondary methodological analysis of the parent dataset reported by Bertochi et al. [64]; 4 weeks	Running distance (W-DIST); weekly perceived load (W-RPE), TRIMP	Measurement properties or response behavior
Crowcroft S, McCleave E, Slattery K, Coutts AJ (2017) [66]	Swimming, nationally competitive swimmers (n = 14; self-report n = 7, combined n = 7)	21.2 ± 3.2	Longitudinal observational study of measurement sensitivity and diagnostic characteristics; 15 months	General health, energy, motivation, stress, recovery, soreness, wellness%; sleep quality, perceived fatigue, TQR; R-R interval, LnRMSSD, LnRMSSD:RR	Measurement properties or response behavior
Dai X, Yan J, Bi X (2025b) [67] [P2]	Rowing; 28 recruited; 26 analyzed male elite rowers	21.7 ± 3.6	Observational longitudinal study; six weeks as reported †; stated dates 4 Nov–2 Dec 2024	CR10 RPE about 30 min post-session × duration; ACWR, monotony and strain; fasting CK, urea, testosterone and cortisol	Concurrent or longitudinal association
Garcin et al. (2002) [17]	Running; 8 high-level male middle-distance runners	20.1 ± 4.0	Single-group pre–post training study; 8 weeks (5 sessions/week, 2 interval sessions)	RPE (Borg 6–20) and estimated time limit during incremental and constant-load track tests; HR, VO2; HLa:RPE ratio (blood lactate)	Measurement properties or response behavior
Matos S, Clemente FM, Brandão A, Pereira J, Rosemann T, Nikolaidis PT, Knechtle B (2019) [68]	Running/trail, male recreational trail runners (n = 47)	34.85 ± 8.88	Longitudinal observational study; 4-week mesocycle	RPE, training time, session-RPE, distance covered; Hooper index (sleep, stress, fatigue, muscle soreness); Cooper test VO2max	Concurrent or longitudinal association
Myers et al. (2017) [69]	Swimming; 10 national/international swimmers (8 men, 2 women)	19.7 ± 2.91	Observational longitudinal monitoring; assessments 6, 4 and 2 weeks before competition; 21-day taper; coach-recorded group training volume	Training volume (group level); BRUMS, RESTQ-76; resting muscle tension, skin conductance and respiration rate (biofeedback); submaximal 2 × 200 m swim (speed, HR); Rusko orthostatic HR test; competition performance	Prospective relationship

Table 1. *Cont.*

Study	Sport/Sample	Age Years	Design/Period	Indicators	Primary Evidence Type
			Subjective responses		
Sanchez R, Nieto C, Leppe J, Gabbett T, Besomi M (2025) [70]	Running/trail-running club sample (n =15; 80% male)	Median 36.5 (IQR 30–38)	Prospective observational cohort; 12-week training mesocycle	TSS (TrainingPeaks); sRPE (CR-10 × min); lnRMSSD (5 min supine on waking; chest strap/EliteHRV); wellness (5-item), readiness (11-point), sleep (PROMIS 8a)	Prospective relationship
Wittig et al. (1989) [18]	Running; 10 well-trained adult male distance runners	32 ± 2.6 (mean ± SE)	Single-group training study; 4 weeks of baseline training followed by 3 weeks of 70% reduced training volume	Weekly POMS, Physical Self-Efficacy Scale and overall/leg RPE at two fixed treadmill speeds (~65% and 85% VO2max); pre-race POMS, PSES and competition anxiety; 5 km race performance	Measurement properties or response behavior

Note: Each report is counted once by its primary analytical domain. Indicators can span domains. Reports arising from related intervention programs should not be interpreted as independent replications. Ages: years, mean ± SD unless labeled; NR = not reported; † = source discrepancy. Confirmed A–C and possible P1–P3 study links are detailed in Supplementary Table S8.

Table 2. Primary evidence categories: report counts and study-group counts after confirmed links are combined.

Primary Evidence Type	Reports *n*	Share %	Study Groups *n*
Measurement properties or response behavior	13	23.2	12
Concurrent or longitudinal association	20	35.7	20
Prospective relationship	12	21.4	12
Monitoring-informed action rule	11	19.6	11
Total	56	100.0	52

Note: Each report was assigned one primary evidence type according to the design feature most directly related to the review questions. The categories describe the type of inference supported and do not represent a hierarchy of study quality. Percentages use 56 reports. Groups may enter several categories; 52 is their union, not the sum of the categories. Possible links remain separate.

Samples ranged from single-athlete case studies to 116 randomized participants in Alfonso: 28 completed and 24 contributed daily-data analyses [34]. Ranieri allocated 33 and analyzed 28 [57]. Vanwanseele reported 96 outdoor participants but 59 in their demographic table [31]. Table 1 retains recruitment/analysis denominators and source discrepancies; this is not an exhaustive attrition appraisal.

Participant level ranged from recreational to professional/elite, with male-only and mixed-sex samples. Participant background and monitored exercise mode differed in some reports (Table 1; Supplementary Table S6). Only the age-eligible triathlete analysis of Plews et al. [19] was included. No pooled subgroup effects were estimated.

Combining confirmed shared studies maps 56 reports to 52 study groups: 15 load, 29 autonomic/physiological, and 8 subjective. Possible Sanders, Dai and Javaloyes links remain separate. Grouping all possible links would yield 48 groups. Sensitivity totals are provided in Supplementary Table S8. These counts do not establish independent participant cohorts.

3.2.1. Types of Evidence Examined

When reports were classified according to the ordered allocation rule, 13 primarily examined measurement properties or response behavior, 20 examined concurrent or longitudinal associations, 12 evaluated prospective relationships with later outcomes, and 11 evaluated monitoring-informed action rules. Thus, 33 of the 56 reports primarily described how indicators behaved or were associated with other measures, while fewer evaluated prospective relationships or monitoring-informed actions. The study group counts in Table 2 are non-additive and do not establish quality or independent replication.

3.2.2. Indicator Use, Measurement Protocols and Outcomes

Chart-entry counts identified TRIMP variants in 21 reports, RPE/session-RPE in 29, and HRV or explicitly HRV-derived metrics in 31 (Table 3). Families are non-exclusive. Supplementary Table S9 identifies each contributing report and category. Counts reflect charted fields rather than exhaustive extraction of every original variable.

Table 3. Nonexclusive indicator-family counts by evidence category, with study-group totals after confirmed links are combined.

Charted Family	Measurement/Response	Association	Prospective	Action Rule	Reports *n*	Study Groups *n*
TRIMP variants	3	9	5	4	21	18
RPE/session-RPE	9	13	5	2	29	27
HRV/HRV-derived metrics	5	8	7	11	31	28

Note: Families are non-exclusive. Counts concern charted fields; operational definitions and report-level entries are provided in Supplementary Table S9.

Protocols varied in timing, posture, sensors used, processing, and aggregation; outcomes included performance, physiological change, recovery, and injury (Supplementary Table S9).

3.3. Quantification of Training Load

The 15 reports examined how training exposure was quantified. Heart-rate-derived TRIMP variants, session rating of perceived exertion (sRPE), Training Stress Score, power-zone distributions, distance, duration, accelerometry, stroke metrics, and modeled fitness-fatigue responses were common [14,16,21–33]. The choice of measure changed the apparent magnitude and distribution of load. Subjective, heart-rate-, and power-based intensity measures showed different distributions and mode-specific associations [24,26,28]. Relationships also varied by training mode in rowers [14] and between coach-planned volume and heart-rate-zone exposure in swimmers [16]. Association or validity for a specific purpose does not establish numerical interchangeability or benefit from using the measure to guide training.

Individual calibration improved correspondence in some settings. Individually weighted TRIMP related more closely to selected performance changes than generic group-based weighting in runners ($n = 8$) and cyclists ($n = 15$) [23,27]. However, associations differed by outcome and context, and Vermeire et al. [32] found no simple linear dose–response relationship between weekly load and threshold or maximal power changes. In professional cyclists, short-term relationships between training dose and record power outputs were accompanied by substantial intra- and inter-individual variability [22]. These associations do not establish that prescription using individualized TRIMP improves performance.

Mechanical and spatiotemporal exposure added information not captured by time or distance alone. In habitual runners, GPS-derived duration, distance, speed, and step count explained only part of accelerometer-derived cumulative load [31]. Napier et al. [25] similarly showed that week-to-week changes depended on whether exposure was expressed as time, steps, cumulative shock, or an external measure multiplied by sRPE.

In elite rowers, individualized internal and external variables strengthened some relationships, but correlation did not establish that either measure represented adaptation [21].

Talsnes et al. [30] observed higher weekly sRPE load and larger performance improvements among high responders during a six-month ski-specific program. The groups also differed in compliance, injury/illness days, motivation, and coach–athlete relationships. These findings describe a multidimensional response pattern and do not isolate training load as a causal mechanism. Schumann et al. [29] compared a proprietary load-guided program with standardized training and found broadly similar improvements, illustrating that evaluating a complete prescription rule is different from validating the score used within it.

3.4. Autonomic and Physiological Responses

The 32 reports focused primarily on autonomic or physiological responses [15,19,34–63]. HRV, especially RMSSD or log-transformed RMSSD, was common, sometimes combined with resting or nocturnal heart rate, sleep, perceived recovery, submaximal exercise responses, or performance tests. Measurement timing, posture, sensor type, aggregation window, and artifact processing varied [5,7]. Pulse-rate variability from photoplethysmography can approximate ECG-derived HRV under suitable resting conditions but is not automatically interchangeable with it [71].

The protocol used influenced interpretation. Morning and nocturnal HRV showed related but non-identical responses during intensified training [53]. Repeated-measures correlation produced different conclusions from pooled Pearson correlations when HRV and load were measured repeatedly within swimmers [40]. In elite rowers, submaximal exercise

heart rate and power-based indices were more reliable than RPE or heart-rate recovery under the tested protocol [38]. These findings support protocol-specific measurement claims rather than universal thresholds.

The 11 reports across the full evidence map evaluated monitoring-informed action rules [29,34,37,45–47,49,50,52,57,61]. Nuuttila et al. [52] reported a greater 10 km improvement with a multi-indicator individualized rule than with predefined training. Alfonso et al. [34] compared rules based on HRV alone or HRV combined with well-being and resting heart rate, but severe attrition and limited between-group findings constrained conclusions about the added value of individual components. Studies using HRV-based action rules used different signals, comparators, and training adjustments and produced outcome-dependent patterns. Table 4 documents the operational components and outcomes of all 11 reports. These comparisons do not establish general superiority of HRV-only guidance across contexts.

Swimming and cross-country skiing were represented by few reports. Du et al. [41] found variation in bTRIMP between session types but no cumulative fatigue signal across three weeks. Davletyarova et al. [40] showed that within-athlete analytical methods affected estimated relationships between HRV and load. Case studies in an ultra-endurance triathlete and an Ironman athlete described detailed temporal patterns [15,35], but their transferability is necessarily limited.

Additional reports examined overload/taper responses and training guidance and forecast HRV or daily recovery using internal validation [36,37,48,62,63]. Their protocols and analytical limitations are detailed in Supplementary Table S6.

Table 4. Operational components of the 11 recorded monitoring-informed action-rule reports.

Report	Signal and Reference	Criterion and Permitted Action	Comparator	Outcome and Interpretation
Schumann et al. [29]	Polar cumulative training-load color code; daily review during weeks 13–24 after 12 common training weeks. Numerical algorithm not public.	Red: rest ≥2 days or until yellow/green. Yellow/green: rotate prescribed sessions; yellow also requires rest in weeks 16, 20, 24. One extra weekly rest day permitted regardless of code.	Standardized endurance program; matched groups.	VO2max and 1000 m running: similar overall improvements; no established superiority of the load-guided program.
Alfonso et al. [34]	9-day initial baseline; then daily RMSSD against rolling 10-day mean − SD. Programs add well-being (WB), then an HRV–RHR composite.	HRV below reference: low; otherwise, high. WB below its own 10-day mean − SD downgrades high→low or low→rest. Program 3 first applies the HRV–RHR composite threshold, then WB. Low after two high days; rest after nine consecutive training days; persistently low HRV permits ≤ 2 rest days, then low.	HRV alone vs HRV + WB vs HRV + WB + RHR.	Selected 1/5 min power differences; 116 randomized, 28 completed, 24 daily-data analysis. No unguided comparator; substantial attrition.
Morinaga and Takai [49]	Morning supine lnRMSSD; 7-day moving average against individual familiarization mean ±0.5 SD.	Reference band mean ± 0.5 SD. Values within the band permitted high-intensity exercise; deviations prompted low-intensity exercise or rest.	Block-periodized cycling training.	Submaximal and maximal aerobic power; outcome-dependent differences under the tested five-week protocol.

Table 4. *Cont.*

Report	Signal and Reference	Criterion and Permitted Action	Comparator	Outcome and Interpretation
Nuuttila et al. [52]	3-day nocturnal lnRMSSD mean vs 4-week rolling mean ±0.5 SD; fatigue/soreness (1–7); HR–running-speed index vs prior 2 weeks.	HRV outside band, symptoms > 5, or index fall > 0.50 flagged. Monday/Thursday integrated review adjusted next 3–4 days. Volume progression +5 percentage points of baseline; recovery −25% current volume, with HIT removed in interval phase.	Predefined program scaled to usual training.	Greater 10 km improvement with individualized guidance; 40 recruited; 30 analyzed, 29 for 10 km. Multiple indicators were changed together.
Ranieri et al. [57]	Published full text: 1 min morning supine HRV4Training; 7-day lnRMSSD moving average against app-defined individual normal range; 4 preparation +6 intervention weeks. Exact numerical band NR.	Within individual range: planned moderate/high intensity. Outside: 45–60 min below VT1 or rest; two consecutive non-optimal days → rest. Numerical band must not be assumed to equal ± 0.5 SD.	HR/ventilatory-threshold-based vs race-pace-based prescription.	33 involved in allocation; 5 adherence exclusions (3 HR, 2 race pace); 28 analyzed (HR9/pace9/HRV10). No group-by-time interactions; improvements over time do not establish HRV-guided superiority.
Kiviniemi et al. [47]	Daily morning standing HF power; moving 10-day mean − SD.	HF below reference or decline for 2 days: low intensity or rest; otherwise high intensity.	Predefined running and non-training control.	Maximal running velocity gain favored HRV; between-group VO2peak change nonsignificant. 30 randomized, 26 analyzed.
Nuuttila et al. [50]	Morning 3 min supine QRT; 3-day mean vs individual 3-week control mean.	Begin next preprogrammed HIT block when QRT 3-day mean exceeds control mean; otherwise low-intensity training.	Predetermined block periodization.	Larger relative Vmax and CMJ improvements with HRV guidance; n = 24 analyzed (13/11); withdrawal categories inconsistent†.
Javaloyes et al. [45]	Morning 90 s supine lnRMSSD; 7-day rolling average. Four-week baseline reference mean ± 0.5 SD; reference updated after four training weeks.	Within reference band: planned moderate/high session permitted. Outside band: low intensity or rest. Follow the source sequence and restrictions; threshold is not transferable by assumption.	Traditional predefined periodization.	17 cyclists (9/8). Selected within-group improvements; between-group differences nonsignificant.
Javaloyes et al. [46]	Smartphone PPG: 90 s supine, last 60 s analyzed. 7-day lnRMSSD mean; reference mean ± 0.5 SD, updated after 4 training weeks.	Outside reference: low/rest; at most 2 consecutive HIT or rest days.	Block periodization.	20 recruited, 15 completed. Methods: HRV8/BP7; source article [46], Table 1: HRV7/BP8 †. No established between-group fitness/performance advantage.

Table 4. *Cont.*

Report	Signal and Reference	Criterion and Permitted Action	Comparator	Outcome and Interpretation
Vesterinen et al. [61]	4 min morning supine RMSSD; 7-day mean. Preparation mean ± 0.5 SD; reference updated after 4 intensive weeks.	Outside band: low/rest; resume MOD/HIT once the mean reference level is regained.	Predefined endurance training.	40 recruited, 31 completed (13/18). Fewer MOD/HIT sessions; between-group 3000 m change nonsignificant (ES 0.42).
Carrasco-Poyatos et al. [37]; protocol [72]	Morning 60 s supine HRV4Training; 7-day lnRMSSD mean. Linked protocol: preparation mean ± 0.5 SD; normal range updated weekly.	Protocol rule: within range → moderate/high intensity; below lower limit → low intensity/rest. Maximum two consecutive moderate/high sessions or rest sessions. Above-upper-limit handling is not explicitly clarified.	Traditional training; 12 randomized/completed (6/group); no post-allocation loss.	No established general VO2max/performance superiority. Within-group Vmax gain is distinct from between-group effects; adjusted HR contrasts depend on model covariates (Carrasco-Poyatos et al. [37]).

Note: Components describe complete rules, not independently validated indicators. Carrasco-Poyatos includes linked protocol details [37,72]. Alfonso's first calculation and composite normalization, and Ranieri's exact bandwidth, were unspecified. Javaloyes et al. [46] has conflicting group labels. Source-specific qualifications are retained in the table.

3.5. Subjective Responses

Nine reports primarily examined subjective indicators [17,18,64–70]. Weekly and monthly perceived-load estimates correlated with TRIMP, and weekly perceived exertion appeared more sensitive to week-to-week load variation than TRIMP or running distance in recreational runners [64,65]. Aggregation increased feasibility but reduced session-level temporal resolution and introduced a longer recall window. Myers et al. [69] described psychological and physiological changes across six weeks before competition, including a 21-day taper.

In national-level swimmers, several monitoring variables were sensitive to week-to-week change, but only general health met the authors' criterion for a good diagnostic tool for decreased performance, and no single measure discriminated performance change consistently [66]. In trail runners, relationships between RPE, wellness components, and load were statistically significant but small [68]. Among endurance athletes followed for 12 weeks, sleep-related impairment differed between injury and injury-free weeks, but the sample was small and did not support a stable injury-prediction rule [70]. In rowers, sRPE-derived load showed selective relationships with blood urea and creatine kinase but not with testosterone or cortisol [67]. The added single-group studies described RPE/mood changes during intensified or reduced training without establishing between-group intervention effects [17,18].

3.6. Methodological Characteristics and Evidence Gaps

Small samples, attrition, and participant overlap constrained transfer and independence [15,23,27,34,35,57]. Recruitment and analysis denominators are separated in Table 1; confirmed and possible shared datasets are in Supplementary Table S8.

Analytical approaches addressed different claims: pooled and repeated-measures correlations differed [40]; forecasting validation varied [62,63]; and intervention comparators and rule components differed (Table 4). Association, prediction, and action effects require separate interpretation.

Reporting uncertainties concerned denominators, group/sex labels and action-rule calculations (Tables 1 and 4). Attrition and confounder extraction were not exhaustive, so their prevalence was not quantified. These study characteristics are distinct from review-process limitations.

4. Discussion

4.1. Principal Findings

Across the 56 included reports, 13 primarily examined measurement properties or response behavior, 20 examined concurrent or longitudinal associations, 12 evaluated prospective relationships, and 11 evaluated monitoring-informed action rules. More reports described indicator behavior or associations than evaluated prospective relationships or monitoring-informed actions. This distinction reflects a signal-to-action gap: monitoring variables are frequently measured and compared, but the consequences of using them remain incompletely tested across the mapped settings. In relation to RQ1, the reports used heterogeneous measurement protocols, time windows, outcomes, and evidence types.

Four cross-domain considerations help organize interpretation of this descriptive map. First, a representation problem arises because indicators with similar labels do not necessarily quantify the same construct. Second, a temporal problem arises when session responses, next-day recovery, weekly load, and seasonal adaptation are treated as if they occurred on one time scale. Third, an inference problem arises when reliability or association is interpreted as a prediction or evidence of benefit. Fourth, an implementation problem arises when a signal is studied without specifying how it should alter training. These are connected issues concerning measurement, interpretation, action, and outcome; their frequency and causal contribution to inconsistent findings were not formally assessed. In relation to RQ2, participant characteristics, attrition, dataset overlap, analytical choices, and reporting uncertainties constrained interpretation and transferability.

The mapped comparisons do not establish a best indicator or broad transferable action threshold. Interpretation requires alignment. The construct, protocol, analytical window, individual reference, intended decision, and outcome must refer to the same practical question. TRIMP depends on the heart-rate model and individual calibration; sRPE depends on rating timing, scale, and aggregation; HRV depends on recording conditions and processing; and external-load summaries depend on mode, terrain, velocity, power, and mechanical exposure. Disagreement between measures may consequently reveal different dimensions or time courses rather than measurement failure [2,5,24,25,28,31,53,73]. The relevant question is not simply whether two indicators agree, but whether each provides information needed for the decision under consideration.

4.2. Synthesis Across Monitoring Domains

When the load, physiological and subjective domains are considered together, they form a sequence rather than a hierarchy. External load describes the work performed; internal load reflects the acute cost of performing it; physiological and subjective responses describe aspects of the athlete's state; and performance or health outcomes indicate whether the intended longer-term result occurred [1–4]. No single stage can represent the entire sequence. A training-load score cannot by itself establish adaptation, and a recovery signal cannot by itself identify which training stimulus caused the change.

The response domain adds a second layer rather than a definitive verdict on the first. Heart-rate-derived load and sRPE summarize session dose, HRV reflects autonomic regulation under a specified protocol, subjective measures capture the athlete's experienced state, and submaximal or maximal tests provide information about functional capability. The fact that these measures sometimes diverged is consistent with their different roles. Subjective

indicators changed in some studies when physiological or external-load summaries did not [64–66,68,70], while selected physiological and subjective indicators did not consistently classify performance change in the studied settings [38,66]. These reports illustrate the complementary questions addressed by different indicators; they do not establish that collecting more data or combining indicators improves a training decision.

Interpretation depends on matching the indicator to the relevant time scale. Same-session strain, next-morning recovery, accumulated weekly load, and adaptation across a training block are different outcomes. A daily deviation may warrant remeasurement without implying a chronic loss of readiness, whereas a weekly average can obscure brief but meaningful responses. Morning and nocturnal HRV, for example, showed related but non-identical patterns during intensified training [53]. Similarly, weekly and monthly perceived-load summaries may improve feasibility and reveal broader change, while sacrificing session-level resolution and increasing the recall interval [64,65]. The analytical window should therefore be selected based on the decision that the measure is intended to inform.

Individual baselines and calibration can improve the relevance of a measure, but individualization is not sufficient by itself. Individually weighted TRIMP showed stronger relationships with selected performance outcomes than generic weighting in some runner and cyclist samples [23,27]. Repeated-measures analysis also produced different conclusions from pooled correlations when load and HRV were observed repeatedly within swimmers [40]. These findings support within-athlete interpretation and suitable repeated-measures methods. They do not establish that an individualized score is a valid proxy for every form of adaptation, or that a threshold developed in one cohort transfers to another.

Direction of change cannot always be interpreted uniformly. HRV increased during functional overreaching in triathletes [48], whereas lower pre-race HRV accompanied better performance in one triathlete [58]. A value outside an individual reference interval should consequently be treated as a signal requiring context, rather than as a diagnosis. The same caution applies to perceived recovery, fatigue, and wellness scores. Taken together, the evidence indicates that individualization is most useful when it aligns the reference value, measurement window, and intended outcome; it cannot correct a poorly defined construct, inconsistent measurement conditions, or a mismatch in time scale.

4.3. Monitoring-Informed Action Rules and Multi-Indicator Systems

The 11 reports evaluating monitoring-informed action rules are the most direct evidence for the final link in the monitoring chain, but they evaluate different complete prescription systems (Table 4). Nuuttila et al. [52] reported greater 10 km improvement with a multi-indicator individualized approach than with predefined training, whereas Schumann et al. [29] found broadly similar improvements between a proprietary load-guided program and standardized training. HRV-guided studies used different signals, comparators, training adjustments, and outcomes, producing outcome-dependent patterns [34,37,45–47,49,50,52,57,61]. Severe attrition in Alfonso et al. [34] further limited conclusions about whether combining HRV with well-being and resting heart rate added value.

These interventions evaluate complete decision systems. Their effects may depend on signal acquisition, data quality, the rule used to interpret the signal, the permitted training adjustment, athlete and coach adherence, and the comparator program. Unless these elements are isolated, an observed effect cannot be attributed to the indicator alone. Javaloyes et al. [45,46] reported within-group gains without established between-group superiority, and Ranieri et al. [57] found no group-by-time interactions. Nonsignificant comparisons in small samples do not demonstrate equivalence, and within-group improvement alone does not establish between-group superiority.

Using several indicators is appealing because it can represent more of the load–response process, but additional variables do not automatically improve a decision. More complex systems create more opportunities for missing data, conflicting signals, and undocumented overrides. Their value should be judged by whether they change a defined decision and improve a relevant outcome enough to justify the measurement burden. The present map describes multi-indicator approaches, but it does not establish one combination as superior across endurance sports.

Future work should distinguish four questions: whether a measure is reliable under a specified protocol; whether it agrees with another measure; whether it prospectively predicts a relevant outcome; and whether acting on it improves that outcome. Concurrent associations cannot establish prospective predictions, and prospective predictions cannot establish that monitoring-informed action rules are beneficial. Historical model fit can support description without demonstrating transfer to a new athlete, season, sport, or outcome. These distinctions should be reflected in study aims, analyses, and conclusions.

4.4. Practical Interpretation and Research Priorities

Le Meur et al. [48] documented performance loss during overload followed by improvement after taper. Studies of daily readiness need temporal analyses that respect repeated observations within athletes, while studies of adaptation require sufficiently long follow-up and standardized performance or health outcomes. Direct comparison of complete monitoring-informed action rules, independent replication, and transparent reporting of overlapping cohorts would strengthen the evidence base.

We propose a practical interpretation framework for evaluation, as follows: define the decision; measure exposure and response under standardized conditions; compare the observations with an individual reference and a relevant outcome; examine convergence and discordance; choose and record an action; and reassess the result. Repeated use of this sequence allows coaches and athletes to examine whether a signal consistently preceded a relevant change, whether the planned action was followed, and whether the intended outcome improved. This is an author-proposed interpretive framework derived from the mapped evidence, not a validated diagnostic or prescriptive algorithm.

Intervention studies should predefine the signal, reference period, decision threshold, permitted action, and outcome used to judge benefit. Reporting should include adherence to both measurement and training, missing observations, device or algorithm changes, coach or athlete overrides, and unintended consequences. When a multi-indicator rule is studied, authors should report which component triggered each action and how frequently the rule changed the planned training. Such information is necessary to determine whether an apparently ineffective system failed at measurement, interpretation, implementation, or the training action itself.

4.5. Strengths and Limitations

A strength of this review is the reproducible search across PubMed, SPORTDiscus, Scopus, and Web of Science, together with explicit eligibility criteria and charting of 56 included reports. The broad scope made it possible to compare load quantification, autonomic and physiological responses, subjective responses, and monitoring-informed action rules within one evidence map. This breadth is consistent with the purpose of a scoping review and highlights where the literature is concentrated and where direct decision-outcome evidence remains limited. The frequency of an indicator or evidence type does not establish its validity, certainty, or effectiveness.

Several limitations affect completeness and interpretation. No systematic backward or forward citation searching, gray literature, Embase, or CINAHL search was performed. Re-

quiring monitoring terminology for general indicators may have led to studies being missed. The Torres cross-check found three eligible reports missed by the database searches; known-item checks do not establish comprehensive sensitivity. Two reports could not be retrieved. Unrecorded second-stage deduplication criteria, residual repeated DOI records and unavailable paired screening/classification decisions limit reproducibility. AI-assisted translation was used for the Spanish-language full text, but handling of non-English titles/abstracts was not separately documented.

No prospective protocol, formal risk-of-bias assessment, or certainty rating was undertaken; eligibility refinements were retrospective. Their absence limits evaluation of methodological decisions and prevents quality-weighted conclusions about effects. The reports included were heterogeneous, several had small samples or substantial attrition, and related publications may share participants. Although primary-domain and evidence-type allocation required judgment, all assignments were reviewed and confirmed by T.O., M.A.A. and T.D., with any differences resolved through discussion. Protocol, demographic and confounder extraction was incomplete. The sample-level age rule permits mixed-age samples and excludes unconfirmed ages. Transfer across sports, training modes, participant levels and sexes requires caution. Accordingly, the 56 reports and 52 mapped study groups describe the extent and character of the available evidence; they should not be interpreted as independent participant cohorts or as support for pooled effectiveness conclusions.

5. Conclusions

This scoping review maps external-load and internal-response indicators across six selected endurance-sport contexts. The included reports varied substantially in measurement protocols, time windows, outcomes, and evidence types. Participant characteristics, attrition, dataset overlap, analytical choices, and reporting uncertainties further limited interpretation and transferability. The evidence therefore supports different uses of monitoring indicators, with distinct requirements for describing training responses and guiding training decisions. Some monitoring-informed training programs produced favorable outcomes, but these findings apply to the complete prescription systems and comparisons evaluated. They do not establish broad transferable action thresholds or a preferred combination of indicators. Future studies should test explicit, reproducible monitoring-and-action rules in defined populations and assess their effects on training exposure, adaptation, performance, and health. The proposed practical framework offers a structure for designing and evaluating these studies.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/sports14110503/s1: Table S1: Databases, interface settings and records retrieved; Table S1a: Post-screening PubMed retrieval register for the 53 database-included reports; Table S1b: Identified search-development references and separately recorded later retrieval outcomes; Table S2: Selection counts by identification route; Table S3a: Studies included by Düking et al. (2021) [9]; Table S3b: Eligibility assessment of potentially relevant reports identified through cross-checking Torres et al. (2025) [10]; Table S4: Excluded reports and primary reasons; Table S5: Reports not retrieved; Table S6: Included-report characteristics, protocols, evidence classifications and findings; Table S7: Funding and material support; Table S8: Source-based study-link assessment; Table S9: Indicator families, evidence type and participant context; Table S9a: Indicator-family frequencies by primary evidence type; Table S10: Complete PRISMA-ScR reporting checklist and locations. Tables S3a and S3b document cross-checks against Düking et al. [9] and Torres et al. [10], respectively. Table S8 cites Vesterinen's doctoral thesis [74] as a source for distinguishing related reports.

Author Contributions: Conceptualization, T.O., T.D., T.K.A. and M.A.A.; methodology, T.O., T.D., T.K.A. and M.A.A.; literature search, T.O.; title and abstract screening, T.O. and M.A.A.; full-text retrieval support, M.A.A.; independent full-text eligibility assessment, T.O. and M.A.A.; data curation and initial evidence classification, T.O.; verification of evidence classification, M.A.A. and T.D.; refinement of the methodology and research questions, M.A.A., T.D. and T.K.A.; writing—original draft preparation, T.O. and M.A.A.; writing—review and editing, T.O., T.D., T.K.A. and M.A.A.; supervision, M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study is a scoping review of published literature and did not involve human participants or animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: The complete database search strategies, selection counts, report-level eligibility decisions, and study-level evidence map are provided in the Supplementary Materials. The underlying screening and charting files are available from the corresponding author on reasonable request. Report-level eligibility decisions and recorded retrieval results are provided in the Supplementary Materials (Tables S1a, S4 and S5).

Acknowledgments: The authors acknowledge the institutional library resources used to access the literature included in this review. During manuscript preparation, ChatGPT-4 (OpenAI) supported language editing, consistency checks, restructuring, reference checks, document preparation, and translation of the included Spanish-language report. The exact models and versions used in the historical sessions could not be recovered from the retained records. The authors take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bourdon, P.C.; Cardinale, M.; Murray, A.; Gastin, P.; Kellmann, M.; Varley, M.C.; Gabbett, T.J.; Coutts, A.J.; Burgess, D.J.; Gregson, W.; et al. Monitoring athlete training loads: Consensus statement. *Int. J. Sports Physiol. Perform.* **2017**, *12*, S2-161–S2-170. [CrossRef] [PubMed]
2. Impellizzeri, F.M.; Marcora, S.M.; Coutts, A.J. Internal and external training load: 15 years on. *Int. J. Sports Physiol. Perform.* **2019**, *14*, 270–273. [CrossRef] [PubMed]
3. Borresen, J.; Lambert, M.I. The quantification of training load, the training response and the effect on performance. *Sports Med.* **2009**, *39*, 779–795. [CrossRef] [PubMed]
4. Halson, S.L. Monitoring training load to understand fatigue in athletes. *Sports Med.* **2014**, *44*, S139–S147. [CrossRef] [PubMed]
5. Buchheit, M. Monitoring training status with HR measures: Do all roads lead to Rome? *Front. Physiol.* **2014**, *5*, 73. [CrossRef] [PubMed]
6. Paquette, M.R.; Napier, C.; Willy, R.W.; Stellingwerff, T. Moving beyond weekly "distance": Optimizing quantification of training load in runners. *J. Orthop. Sports Phys. Ther.* **2020**, *50*, 564–569. [CrossRef] [PubMed]
7. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation* **1996**, *93*, 1043–1065. [CrossRef]
8. Meeusen, R.; Duclos, M.; Foster, C.; Fry, A.; Gleeson, M.; Nieman, D.; Raglin, J.; Rietjens, G.; Steinacker, J.; Urhausen, A.; et al. Prevention, diagnosis, and treatment of the overtraining syndrome: Joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Med. Sci. Sports Exerc.* **2013**, *45*, 186–205. [CrossRef] [PubMed]
9. Düking, P.; Zinner, C.; Trabelsi, K.; Reed, J.L.; Holmberg, H.-C.; Kunz, P.; Sperlich, B. Monitoring and adapting endurance training on the basis of heart rate variability monitored by wearable technologies: A systematic review with meta-analysis. *J. Sci. Med. Sport* **2021**, *24*, 1180–1192. [CrossRef] [PubMed]
10. Torres, G.; Maia, F.; Nakamura, F.Y.; Neiva, H.P.; Sousa, A. Rating of perceived exertion in continuous sports: A scoping review with evidence gap map. *Front. Sports Act. Living* **2025**, *7*, 1553998. [CrossRef] [PubMed]
11. Bellenger, C.R.; Fuller, J.T.; Thomson, R.L.; Davison, K.; Robertson, E.Y.; Buckley, J.D. Monitoring athletic training status through autonomic heart rate regulation: A systematic review and meta-analysis. *Sports Med.* **2016**, *46*, 1461–1486. [CrossRef] [PubMed]
12. Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.J.; Horsley, T.; Weeks, L.; et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [CrossRef] [PubMed]

13. Peters, M.D.J.; Marnie, C.; Tricco, A.C.; Pollock, D.; Munn, Z.; Alexander, L.; McInerney, P.; Godfrey, C.M.; Khalil, H. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid. Synth.* **2020**, *18*, 2119–2126. [CrossRef] [PubMed]
14. Dai, X.; Yan, J.; Bi, X. Concurrent validity and reliability of the session rating of perceived exertion scale among high-trained rower during training sessions. *BMC Sports Sci. Med. Rehabil.* **2025**, *17*, 196. [CrossRef] [PubMed]
15. Nieto-Jiménez, C.; Ruso-Álvarez, J.; Pardos-Mainer, E.; Naranjo Orellana, J. La variabilidad de la frecuencia cardiaca en el control del entrenamiento en un corredor de Ironman. *Retos* **2020**, *37*, 339–343. [CrossRef]
16. Sixsmith, H.; Crowcroft, S.; Slattery, K. Assessing the use of heart-rate monitoring for competitive swimmers. *Int. J. Sports Physiol. Perform.* **2023**, *18*, 1321–1327. [CrossRef] [PubMed]
17. Garcin, M.; Fleury, A.; Billat, V. The ratio HLa:RPE as a tool to appreciate overreaching in young high-level middle-distance runners. *Int. J. Sports Med.* **2002**, *23*, 16–21. [CrossRef] [PubMed]
18. Wittig, A.F.; Houmard, J.A.; Costill, D.L. Psychological effects during reduced training in distance runners. *Int. J. Sports Med.* **1989**, *10*, 97–100. [CrossRef] [PubMed]
19. Plews, D.J.; Laursen, P.B.; Le Meur, Y.; Hausswirth, C.; Kilding, A.E.; Buchheit, M. Monitoring training with heart rate-variability: How much compliance is needed for valid assessment? *Int. J. Sports Physiol. Perform.* **2014**, *9*, 783–790. [CrossRef] [PubMed]
20. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
21. Dai, X.; Yan, J.; Bi, X. The association between internal and external measures of training load in elite rowers during ergometer training. *Sci. Rep.* **2025**, *15*, 44923. [CrossRef] [PubMed]
22. Gallo, G.; Bosio, A.; Martin, M.; Morelli, A.; Azzolini, M.; Guercilena, L.; Larrazabal, J.; Rampinini, E. Relationships between training dose and record power outputs in professional road cyclists: Insights and threats to validity. *Biol. Sport* **2023**, *40*, 485–495. [CrossRef] [PubMed]
23. Manzi, V.; Iellamo, F.; Impellizzeri, F.; D'Ottavio, S.; Castagna, C. Relation between individualized training impulses and performance in distance runners. *Med. Sci. Sports Exerc.* **2009**, *41*, 2090–2096. [CrossRef] [PubMed]
24. Manzi, V.; Bovenzi, A.; Castagna, C.; Sinibaldi Salimei, P.; Volterrani, M.; Iellamo, F. Training-load distribution in endurance runners: Objective versus subjective assessment. *Int. J. Sports Physiol. Perform.* **2015**, *10*, 1023–1028. [CrossRef] [PubMed]
25. Napier, C.; Ryan, M.; Menon, C.; Paquette, M.R. Session rating of perceived exertion combined with training volume for estimating training responses in runners. *J. Athl. Train.* **2020**, *55*, 1285–1291. [CrossRef] [PubMed]
26. Rodríguez-Marroyo, J.A.; Villa, J.G.; Fernández, G.; Foster, C. Effect of cycling competition type on effort based on heart rate and session rating of perceived exertion. *J. Sports Med. Phys. Fit.* **2013**, *53*, 154–161.
27. Sanders, D.; Abt, G.; Hesselink, M.K.C.; Myers, T.; Akubat, I. Methods of monitoring training load and their relationships to changes in fitness and performance in competitive road cyclists. *Int. J. Sports Physiol. Perform.* **2017**, *12*, 668–675. [CrossRef] [PubMed]
28. Sanders, D.; Myers, T.; Akubat, I. Training-intensity distribution in road cyclists: Objective versus subjective measures. *Int. J. Sports Physiol. Perform.* **2017**, *12*, 1232–1237. [CrossRef] [PubMed]
29. Schumann, M.; Botella, J.; Karavirta, L.; Häkkinen, K. Training-load-guided vs standardized endurance training in recreational runners. *Int. J. Sports Physiol. Perform.* **2017**, *12*, 295–303. [CrossRef] [PubMed]
30. Talsnes, R.K.; van den Tillaar, R.; Cai, X.; Sandbakk, Ø. Comparison of high- vs. Low-responders following a 6-month XC ski-specific training period: A multidisciplinary approach. *Front. Sports Act. Living* **2020**, *2*, 114. [CrossRef] [PubMed]
31. Vanwanseele, B.; Op De Beéck, T.; Schütte, K.; Davis, J. Accelerometer based data can provide a better estimate of cumulative load during running compared to GPS based parameters. *Front. Sports Act. Living* **2020**, *2*, 575596. [CrossRef] [PubMed]
32. Vermeire, K.M.; Vandewiele, G.; Caen, K.; Lievens, M.; Bourgois, J.G.; Boone, J. Training progression in recreational cyclists: No linear dose-response relationship with training load. *J. Strength Cond. Res.* **2021**, *35*, 3500–3505. [CrossRef] [PubMed]
33. Wallace, L.K.; Slattery, K.M.; Coutts, A.J. A comparison of methods for quantifying training load: Relationships between modelled and actual training responses. *Eur. J. Appl. Physiol.* **2014**, *114*, 11–20. [CrossRef] [PubMed]
34. Alfonso, C.; Clarke, D.C.; Capdevila, L. Individual training prescribed by heart rate variability, heart rate and well-being scores in experienced cyclists. *Sci. Rep.* **2025**, *15*, 34023. [CrossRef] [PubMed]
35. Anderson, T.; van Mourik, R.A.; Martin, K.J.; Eijsvogels, T.M.H.; Longoria, K.A. 100 long-distance triathlons in 100 days: A case study on ultraendurance, biomarkers, and physiological outcomes. *Int. J. Sports Physiol. Perform.* **2023**, *18*, 444–453. [CrossRef] [PubMed]
36. Bulte, K.R.; Bruce, L.; Hammond, K.; Corrigan, S.L.; Main, L.C. Use of heart-rate variability to examine readiness to perform in response to overload and taper in swimmers. *Int. J. Sports Physiol. Perform.* **2025**, *20*, 918–924. [CrossRef] [PubMed]
37. Carrasco-Poyatos, M.; González-Quílez, A.; Altini, M.; Granero-Gallegos, A. Heart rate variability-guided training in professional runners: Effects on performance and vagal modulation. *Physiol. Behav.* **2022**, *244*, 113654. [CrossRef] [PubMed]

38. Clark, J.R.; McKune, A.J.; Wood, P.S. Reliability and usefulness of a self-regulated 6-km submaximal ergometer training test set in elite rowers. *Int. J. Sports Sci. Coach.* **2025**, *20*, 2613–2623. [CrossRef]
39. Coates, A.M.; Incognito, A.V.; Seed, J.D.; Doherty, C.J.; Millar, P.J.; Burr, J.F. Three weeks of overload training increases resting muscle sympathetic activity. *Med. Sci. Sports Exerc.* **2018**, *50*, 928–937. [CrossRef] [PubMed]
40. Davletyarova, K.; Vacher, P.; Nicolas, M.; Kapilevich, L.V.; Mourot, L. Associations between heart rate variability-derived indexes and training load: Repeated measures correlation approach contribution. *J. Strength Cond. Res.* **2022**, *36*, 2005–2010. [CrossRef] [PubMed]
41. Du, C.; Xu, H.; Shi, J.; Wei, J.; Quan, H.; Zhu, T. Internal load and fatigue characteristics in male cross-country skiers during summer training. *J. Strength Cond. Res.* **2026**, *40*, e727–e735. [CrossRef] [PubMed]
42. Fuller, J.T.; Bellenger, C.R.; Thewlis, D.; Arnold, J.; Thomson, R.L.; Tsiros, M.D.; Robertson, E.Y.; Buckley, J.D. Tracking performance changes with running-stride variability when athletes are functionally overreached. *Int. J. Sports Physiol. Perform.* **2017**, *12*, 357–363. [CrossRef] [PubMed]
43. Fuller, J.T.; Doyle, T.L.A.; Doyle, E.W.; Arnold, J.B.; Buckley, J.D.; Wills, J.A.; Thewlis, D.; Bellenger, C.R. The cumulative impacts of fatigue during overload training can be tracked using field-based monitoring of running stride interval correlations. *Sensors* **2024**, *24*, 5538. [CrossRef] [PubMed]
44. Iguchi, J.; Takimoto, M.; Kuzuhara, K.; Hojo, T.; Fujisawa, Y. Seasonal variations in training load, sleep parameters, and hormonal markers in collegiate male rowers during the off-season. *Sports* **2025**, *13*, 407. [CrossRef] [PubMed]
45. Javaloyes, A.; Sarabia, J.M.; Lamberts, R.P.; Moya-Ramon, M. Training prescription guided by heart-rate variability in cycling. *Int. J. Sports Physiol. Perform.* **2019**, *14*, 23–32. [CrossRef] [PubMed]
46. Javaloyes, A.; Sarabia, J.M.; Lamberts, R.P.; Plews, D.; Moya-Ramon, M. Training prescription guided by heart rate variability vs. Block periodization in well-trained cyclists. *J. Strength Cond. Res.* **2020**, *34*, 1511–1518. [CrossRef] [PubMed]
47. Kiviniemi, A.M.; Hautala, A.J.; Kinnunen, H.; Tulppo, M.P. Endurance training guided individually by daily heart rate variability measurements. *Eur. J. Appl. Physiol.* **2007**, *101*, 743–751. [CrossRef] [PubMed]
48. Le Meur, Y.; Pichon, A.; Schaal, K.; Schmitt, L.; Louis, J.; Gueneron, J.; Vidal, P.P.; Hausswirth, C. Evidence of parasympathetic hyperactivity in functionally overreached athletes. *Med. Sci. Sports Exerc.* **2013**, *45*, 2061–2071. [CrossRef] [PubMed]
49. Morinaga, H.; Takai, Y. Heart rate variability-guided aerobic training without moderate-intensity enhances submaximal and maximal aerobic power with less training load. *J. Hum. Sport Exerc.* **2025**, *20*, 366–380. [CrossRef]
50. Nuuttila, O.P.; Nikander, A.; Polomoshnov, D.; Laukkanen, J.A.; Häkkinen, K. Effects of HRV-guided vs. Predetermined block training on performance, HRV and serum hormones. *Int. J. Sports Med.* **2017**, *38*, 909–920. [CrossRef] [PubMed]
51. Nuuttila, O.P.; Nummela, A.; Häkkinen, K.; Seipäjärvi, S.; Kyröläinen, H. Monitoring training and recovery during a period of increased intensity or volume in recreational endurance athletes. *Int. J. Environ. Res. Public Health* **2021**, *18*, 2401. [CrossRef] [PubMed]
52. Nuuttila, O.P.; Nummela, A.; Korhonen, E.; Häkkinen, K.; Kyröläinen, H. Individualized endurance training based on recovery and training status in recreational runners. *Med. Sci. Sports Exerc.* **2022**, *54*, 1690–1701. [CrossRef] [PubMed]
53. Nuuttila, O.P.; Kyröläinen, H.; Kokkonen, V.P.; Uusitalo, A. Morning versus nocturnal heart rate and heart rate variability responses to intensified training in recreational runners. *Sports Med. Open* **2024**, *10*, 120. [CrossRef] [PubMed]
54. Nuuttila, O.P.; Uusitalo, A.; Kokkonen, V.P.; Weerarathna, N.; Kyröläinen, H. Monitoring fatigue state with heart rate-based and subjective methods during intensified training in recreational runners. *Eur. J. Sport Sci.* **2024**, *24*, 857–869. [CrossRef] [PubMed]
55. Nuuttila, O.P.; Schäfer Olstad, D.; Martinmäki, K.; Uusitalo, A.; Kyröläinen, H. Monitoring sleep and nightly recovery with wrist-worn wearables: Links to training load and performance adaptations. *Sensors* **2025**, *25*, 533. [CrossRef] [PubMed]
56. Plews, D.J.; Laursen, P.B.; Kilding, A.E.; Buchheit, M. Heart rate variability in elite triathletes, is variation in variability the key to effective training? A case comparison. *Eur. J. Appl. Physiol.* **2012**, *112*, 3729–3741. [CrossRef] [PubMed]
57. Ranieri, L.E.; Casado, A.; Martin, D.; Trujillo-Colmena, D.; Gil-Arias, A.; Kenneally, M.; Jiménez, A. Performance and physiological effects of race pace-based versus heart rate variability-guided training prescription in runners. *Med. Sci. Sports Exerc.* **2025**, *57*, 1510–1522. [CrossRef] [PubMed]
58. Stanley, J.; D'Auria, S.; Buchheit, M. Cardiac parasympathetic activity and race performance: An elite triathlete case study. *Int. J. Sports Physiol. Perform.* **2015**, *10*, 528–534. [CrossRef] [PubMed]
59. Vesterinen, V.; Häkkinen, K.; Hynynen, E.; Mikkola, J.; Hokka, L.; Nummela, A. Heart rate variability in prediction of individual adaptation to endurance training in recreational endurance runners. *Scand. J. Med. Sci. Sports* **2013**, *23*, 171–180. [CrossRef] [PubMed]
60. Vesterinen, V.; Nummela, A.; Ayramo, S.; Laine, T.; Hynynen, E.; Mikkola, J.; Häkkinen, K. Monitoring training adaptation with a submaximal running test under field conditions. *Int. J. Sports Physiol. Perform.* **2016**, *11*, 393–399. [CrossRef] [PubMed]
61. Vesterinen, V.; Nummela, A.; Heikura, I.; Laine, T.; Hynynen, E.; Botella, J.; Häkkinen, K. Individual endurance training prescription with heart rate variability. *Med. Sci. Sports Exerc.* **2016**, *48*, 1347–1354. [CrossRef] [PubMed]

62. Barsumyan, A.; Saukkonen, A.; Soost, C.; Graw, J.A.; Burchard, R. Predicting next day heart rate variability based on training load in cyclists using machine learning. *Sports* **2026**, *14*, 271. [CrossRef] [PubMed]
63. Rothschild, J.A.; Stewart, T.; Kilding, A.E.; Plews, D.J. Predicting daily recovery during long-term endurance training using machine learning analysis. *Eur. J. Appl. Physiol.* **2024**, *124*, 3279–3290. [CrossRef] [PubMed]
64. Bertochi, G.F.A.; Mota, G.R.; Haile, L.; Marocolo, M.; Sasaki, J.E. Validity of weekly and monthly perceived exertion in runners. *Int. J. Sports Med.* **2023**, *44*, 763–777. [CrossRef] [PubMed]
65. Bertochi, G.F.A.; Sasaki, J.E. Weekly perceived exertion is more sensitive to detecting variations in training load in runners than TRIMP or running distance. *Meas. Phys. Educ. Exerc. Sci.* **2025**, *29*, 59–65. [CrossRef]
66. Crowcroft, S.; McCleave, E.L.; Slattery, K.; Coutts, A.J. Assessing the measurement sensitivity and diagnostic characteristics of athlete-monitoring tools in national swimmers. *Int. J. Sports Physiol. Perform.* **2017**, *12*, S2-95–S2-100. [CrossRef] [PubMed]
67. Dai, X.; Yan, J.; Bi, X. Weekly fluctuations in subjective and objective measures of internal training load and their relationships in male elite rowers. *J. Sports Sci. Med.* **2025**, *24*, 269–276. [CrossRef] [PubMed]
68. Matos, S.; Clemente, F.M.; Brandão, A.; Pereira, J.; Rosemann, T.; Nikolaidis, P.T.; Knechtle, B. Training load, aerobic capacity and their relationship with wellness status in recreational trail runners. *Front. Physiol.* **2019**, *10*, 1189. [CrossRef] [PubMed]
69. Myers, V.R.; McKillop, A.L.; Fraser, S.J.; Abel, J.M.; Wells, G.D. Physiological and psychological adaptations during taper in competitive swimmers. *Int. J. Sports Sci. Coach.* **2017**, *12*, 481–494. [CrossRef]
70. Sanchez, R.; Nieto, C.; Leppe, J.; Gabbett, T.; Besomi, M. Associations between training load, heart rate variability, perceptual fatigue, sleep, and injury in endurance athletes during a 12-week training mesocycle. *Int. J. Sports Sci. Coach.* **2025**, *20*, 1918–1928. [CrossRef]
71. Schäfer, A.; Vagedes, J. How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram. *Int. J. Cardiol.* **2013**, *166*, 15–29. [CrossRef] [PubMed]
72. Carrasco-Poyatos, M.; González-Quílez, A.; Martínez-González-Moro, I.; Granero-Gallegos, A. HRV-guided training for professional endurance athletes: A protocol for a cluster-randomized controlled trial. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5465. [CrossRef] [PubMed]
73. Scheltinga, B.L.; Buurke, J.H.; Kok, J.N.; Reenalda, J. Do training load metrics agree? A comparison of session rate of perceived exertion, physiological and biomechanical load in outdoor running. *Sports Med. Open* **2026**, *12*, 18. [CrossRef] [PubMed]
74. Vesterinen, V. Predicting and Monitoring Individual Endurance Training Adaptation and Individualizing Training Prescription with Endurance Performance, Cardiac Autonomic Regulation and Neuromuscular Performance. Ph.D. Thesis, University of Jyväskylä, Jyväskylä, Finland, 2016. Available online: https://jyx.jyu.fi/handle/123456789/51870 (accessed on 1 October 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.