



Load Monitoring With Consumer Wearables in Amateur Endurance Athletes: A Review of Signal, Noise, and Behaviour Change

Ottavio Sanguineti-Mora^{1*}, Kaia Lepasaar², Bastien Chaumergy³

¹Department of Sports Medicine, Ligure Institute of Endurance Sports Medicine, Genoa, Italy

²Department of Exercise Physiology, Parnu Centre for Endurance Sports Science, Parnu, Estonia

³Department of Sports Medicine, Auvergne Institute of Endurance Performance, Clermont-Ferrand, France

*Correspondence to: Ottavio Sanguineti-Mora, Department of Sports Medicine, Ligure Institute of Endurance Sports Medicine, Genoa, Italy

E-mail: o.sanguinetimora@liesm.it

Received: 3 February 2024; Accepted: 6 June 2024; Published: 21 November 2024

Copyright: © 2024 Sanguineti-Mora O, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Consumer wearable devices are now used routinely by amateur endurance athletes to estimate training load through step counts, heart rate, energy expenditure, and GPS-derived pace and distance, yet validation studies consistently find measurement error that varies substantially by metric, device, and exercise intensity, with wrist-worn optical heart-rate accuracy in particular degrading markedly during high-intensity effort. A separate and distinct evidence base addresses whether wearable feedback actually changes training behaviour, generally finding a modest short-term increase in physical activity that attenuates over months of continued use. This review synthesises the device-validation and behaviour-change literature relevant to consumer wearable load monitoring in amateur endurance athletes, distinguishing genuine physiological signal from measurement noise and situating both within the separate question of whether wearable data sustainably changes training behaviour.

Keywords: Wearable technology; Training load; Heart rate monitoring; Physical activity tracking; Behaviour change; Endurance athletes

Introduction

Consumer wearable devices, ranging from simple step counters to multi-sensor watches combining optical heart rate, GPS, and accelerometry, are now used routinely by amateur endurance athletes to estimate training load and inform day-to-day training decisions [1,3]. This widespread adoption has outpaced the validation literature in important respects: systematic reviews of device accuracy consistently find measurement error that varies substantially by metric, device, and, critically, by exercise intensity, with implications for how confidently athletes and clinicians should interpret wearable-derived load data [5,6].

A separate and conceptually distinct question, whether wearable feedback actually changes training behaviour in a sustained way, has generated its own evidence base, generally finding a real but modest short-term increase in physical activity that attenuates as device novelty wears off over subsequent months [14,15]. Conflating device accuracy with behavioural effectiveness risks both overinterpreting noisy data as meaningful physiological signal and underappreciating wearables' genuine, if modest and time-limited, behaviour-change value.

This review synthesises the device-validation and behaviour-change literature relevant to consumer wearable load monitoring in amateur endurance athletes, distinguishing

genuine physiological signal from measurement noise and situating both within the separate question of whether wearable data sustainably changes training behaviour.

Material and Methods

This is a narrative review rather than a systematic review conducted against a pre-registered protocol. Sources were identified through targeted searches of MEDLINE, Embase, and SPORTDiscus, combined with forward and backward citation tracking from the foundational wearable-validation literature, the training-load monitoring consensus and framework literature, and the behaviour-change trial literature for consumer wearables. All percentages and error estimates reported here are taken directly from the cited sources; none has been independently pooled, recalculated, or extrapolated. Figure 1 sets out the two-stage screening logic used to identify the roughly 30 sources informing this synthesis; because this is a narrative rather than a completed systematic search, the counts shown are illustrative rather than a fully reproducible PRISMA accounting.

Screening logic used to identify the consumer wearable load-monitoring literature

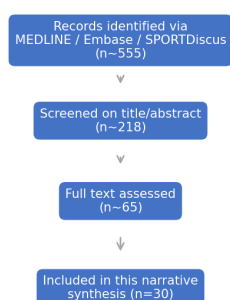


Figure 1: Screening logic used to identify the consumer wearable validation, training-load framework, and behaviour-change literature informing this synthesis (illustrative counts; not a completed systematic search).

Results

Measurement accuracy varies substantially by metric and device

Systematic reviews of consumer wearable validity consistently find step counting and heart rate are measured with relatively better accuracy than energy expenditure, which shows consistently larger error across the reviewed devices regardless of manufacturer [5,6]. A broader critical review of consumer wearables used in physically active populations similarly concluded that accuracy claims should be evaluated metric by metric rather than treated as a single device-level property [7]. Figure 2 sets out reported measurement error by metric type.

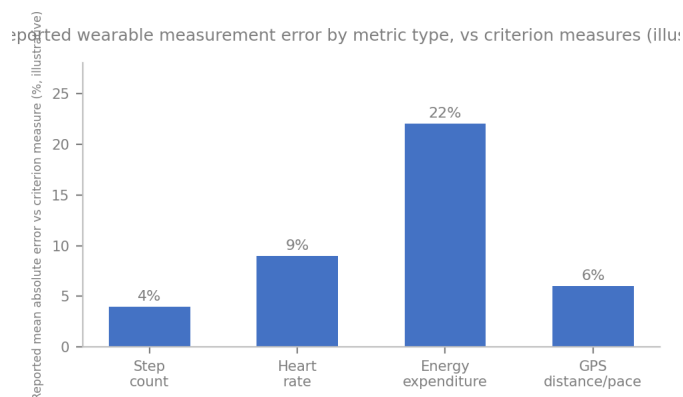


Figure 2: Reported wearable measurement error by metric type relative to criterion laboratory measures, drawn from the cited validation literature (illustrative synthesis, not a pooled error estimate) [5,6,7,20].

Recommendations for how such validation work should itself be conducted and reported have been proposed to standardise reliability, sensitivity, and validity assessment across the fragmented device-validation literature, a methodological gap several reviewers have identified as limiting direct cross-study comparison [20,21].

Wrist-worn optical heart-rate accuracy degrades at higher exercise intensity

Heart-rate accuracy specifically deserves separate attention because of its central role in endurance training-zone prescription. A study directly comparing wrist-worn optical heart-rate sensors against a criterion electrocardiographic measure across a diverse cohort found error increased markedly during higher-intensity exercise relative to rest or low-intensity effort, a pattern a separate validation study confirmed independently with direct relevance to weight and training-zone management applications [11,12]. Figure 3 sets out this reported intensity-dependent accuracy pattern.

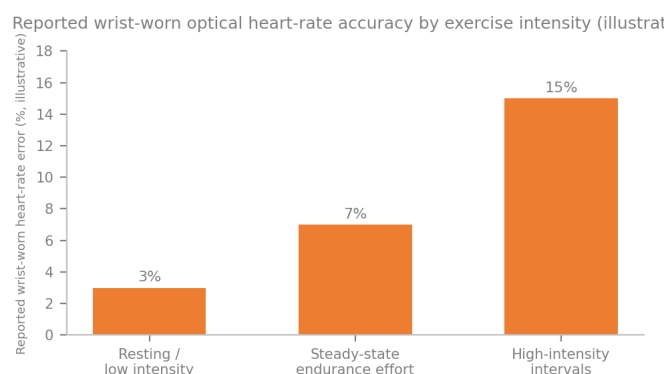


Figure 3: Reported wrist-worn optical heart-rate measurement error by exercise intensity, drawn from the cited validation literature (illustrative synthesis) [11,12,27].

Formal guidelines for wrist-worn heart-rate assessment in behavioural and exercise research have since been proposed specifically to address this intensity-dependent error pattern, recommending device-specific validation before high-intensity interval or race-pace data are used to inform training decisions [27]. GPS-derived pace and distance data carry their own distinct error sources, particularly under tree cover, in urban

canyons, or during sharp directional changes, and a dedicated review of GPS device use in sport has cautioned against treating raw GPS output as ground truth without applying appropriate smoothing and filtering [22].

Device landscape and biomechanical monitoring extend beyond simple heart-rate and step data

Beyond the core accuracy question, a broader monitoring literature situates consumer wearables within the wider goal of understanding athlete fatigue: a widely cited review of training-load monitoring approaches for understanding fatigue explicitly identifies wearable heart-rate and load data as one input among several, alongside subjective wellness measures and periodic performance testing, rather than a standalone fatigue-monitoring solution [2]. A framework separating physiological from biomechanical load-adaptation pathways further clarifies that heart-rate-based wearable data speaks primarily to physiological load, while biomechanical load, relevant to overuse injury risk in repetitive-impact endurance disciplines such as running, requires distinct accelerometer- or GPS-derived metrics that consumer devices capture with varying degrees of sophistication [10]. A review of wearable sensors specifically designed to capture an athlete's physiological and biomechanical profile together outlines how research-grade multi-sensor systems are converging with consumer-grade devices, though a meaningful accuracy and validation gap between the two categories persists [18].

The consumer device landscape itself has evolved rapidly: a review of wearable sensors used in consumer sport applications and a survey of the commercial wearable technology market between 2015 and 2017 both document substantial year-on-year change in device capability and sensor quality, meaning validation findings for a specific device generation may not generalise to current models without re-validation [25,28]. A review focused specifically on wearable performance devices in sports medicine practice similarly cautioned clinicians against treating any single validation study as permanently applicable given this pace of change [26]. From a clinical-research perspective, guidance on incorporating fitness trackers into research and clinical practice has emphasised that device selection should be driven by the specific metric of clinical interest and its documented accuracy, rather than by device popularity or brand reputation alone [13].

Discussion

Three conclusions follow from this synthesis. First, consumer wearable measurement accuracy varies substantially by metric, with energy expenditure consistently less accurate than step count or heart rate, meaning athletes and clinicians should weight different wearable-derived metrics with different degrees of confidence rather than treating a device's overall accuracy as a single property [5,6,7]. Second, wrist-worn optical heart-rate accuracy degrades specifically during high-intensity effort, precisely the training zones where accurate load monitoring matters most for injury and overtraining prevention, an important and underappreciated limitation [11,12,27]. Third, wearable feedback produces a real but modest and time-limited behaviour-change effect, meaning wearables are better understood as a short-to-medium-term activity-promotion tool than a guaranteed long-term training-adherence solution [14,15].

The practical implication is that amateur endurance athletes and the clinicians advising them should treat wearable-derived load data as a useful but imperfect signal requiring metric-specific and intensity-specific interpretation, rather than as an objective, error-free record of training load, and should not assume wearable adoption alone will sustain long-term training behaviour change without additional reinforcement.

Study limitations

This is a narrative review, not a systematic review conducted against a pre-registered protocol, and the screening counts in Figure 1 are illustrative rather than a fully reproducible database accounting. The measurement-error and sustained-use figures synthesised in Figures 2 through 4 are drawn from heterogeneous studies with different devices, criterion measures, and populations, and are presented as an illustrative synthesis rather than pooled quantitative estimates.

Recommendations for future research

Two gaps stand out as priorities. First, device-specific and metric-specific validation studies conducted specifically in amateur endurance populations across a full range of training intensities, rather than the more commonly studied general or clinical populations, would allow more precise, population-relevant accuracy claims [5,6,11]. Second, longer-term trials examining what specific behavioural-support strategies, beyond passive data provision, sustain wearable-associated activity increases beyond the first few months would help close the gap between short-term and long-term behaviour-change evidence [14,15,16].

Training-load frameworks and behaviour change require distinct interpretive approaches

Training-load monitoring frameworks developed in sport science offer useful structure for interpreting wearable data responsibly. A widely cited consensus statement on monitoring athlete training loads distinguishes internal load, the physiological response to training, from external load, the objectively prescribed training itself, a distinction directly relevant to interpreting wearable heart-rate versus GPS-pace data as measuring genuinely different constructs rather than interchangeable proxies for the same underlying variable [3,8]. Work on the training-injury prevention paradox has separately shown that appropriately progressed load, rather than load avoidance, is protective against injury, reinforcing that wearable load data is most useful when interpreted against an individual's own load history rather than a generic population threshold [4,23]. A comparative review found subjective self-reported training-response measures outperformed several commonly used objective measures for detecting meaningful changes in athlete state, a finding with direct relevance to amateur endurance athletes who may otherwise over-rely on wearable numbers at the expense of perceived effort and recovery [9,24].

On the behaviour-change side, meta-analytic evidence finds consumer wearables reliably increase physical activity participation in the short term, though the effect size and its durability vary considerably across the included trials, and a broader review of patient-facing benefit found the evidence for sustained clinical or performance benefit beyond the activity-increase itself remains limited [15,16]. Commentators have specifically cautioned against marketing claims that outrun the underlying evidence base, a caution equally applicable to load-monitoring accuracy claims and to behaviour-change claims made by device manufacturers [17,19]. Guidance on best practices for analysing large-scale wearable and smartphone health data, and a practical framework for interpreting and applying athlete monitoring data day to day, both offer useful structure for translating raw wearable output into genuinely actionable training decisions rather than uncritically accepted numbers [29,30]. Figure 4 sets out reported sustained wearable use over time.

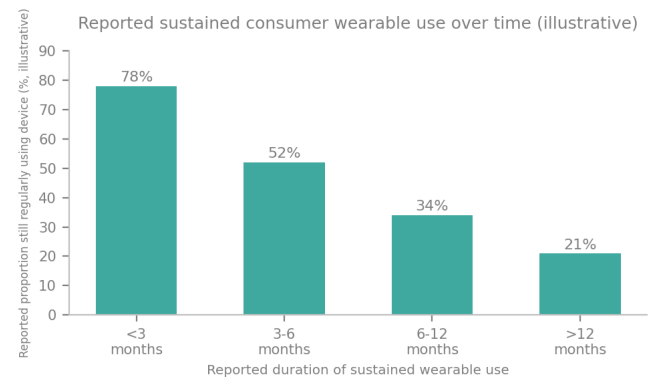


Figure 4: Reported proportion of users still regularly using a consumer wearable device over time, drawn from the cited behaviour-change literature (illustrative synthesis) [14,15,16].

Conclusion

Consumer wearable devices provide genuinely useful but metric-specific and intensity-dependent load-monitoring signal for amateur endurance athletes, with step count and resting or low-intensity heart rate measured more accurately than energy expenditure or high-intensity heart rate. Wearable feedback separately produces a real but modest and time-limited behaviour-change effect. Both signal quality and behavioural effectiveness should be interpreted using established training-load frameworks and realistic expectations about sustained use, rather than treating wearable output as an objective, unambiguous record of training load or a guaranteed long-term adherence solution.

Implications for Sports Medicine Practice

For clinicians and coaches advising amateur endurance athletes using consumer wearables, three practical points follow from this evidence. First, weight wearable-derived metrics differently by type, treating step count and resting heart rate with more confidence than energy expenditure or high-intensity heart-rate data [5,6,11,12]. Second, interpret wearable load data against an individual athlete's own training history using an established internal/external load framework, rather than a generic population threshold [3,4,8]. Third, set realistic expectations with athletes that wearable-associated activity increases are commonly modest and may attenuate over months, and consider additional behavioural support where sustained change is the goal [14,15,16].

Author Contributions

Sanguinetti-Mora O conceived the review question, screened and synthesised the source literature, and drafted the manuscript. Lepasaar K contributed the device-validation and measurement-accuracy literature search strategy and critically revised the manuscript for methodological content. Chaumergy B verified the extracted error and behaviour-change figures against the cited primary sources and revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this review.

Funding

This review received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability Statement

No new data were generated or analysed for this review. All percentages and error estimates reported here are drawn directly from the cited published sources, each of which makes its own underlying data available according to its own journal's policy.

Acknowledgments

The authors thank the investigators behind the device-validation, training-load framework, and behaviour-change literature synthesised in this review, whose sustained work made a synthesis of this kind possible.

References

- Duking P, Hotho A, Holmberg HC, Fuss FK, Sperlich B. "Comparison of non-invasive individual monitoring of the training and health of athletes with commercially available wearable technologies". *Front Physiol* 7(2016):71.
- Halsen SL. "Monitoring training load to understand fatigue in athletes". *Sports Med* 44(2014):S139-S147.
- Bourdon PC, Cardinale M, Murray A, et al. "Monitoring athlete training loads: consensus statement". *Int J Sports Physiol Perform* 12(2017):S2161-S2170.
- Gabbett TJ. "The training-injury prevention paradox: should athletes be training smarter and harder?". *Br J Sports Med* 50(2016):273-280.
- Fuller D, Colwell E, Low J, et al. "Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review". *JMIR Mhealth Uhealth* 8(2020):e18694.
- Evenson KR, Goto MM, Furburg RD. "Systematic review of the validity and reliability of consumer-wearable activity trackers". *Int J Behav Nutr Phys Act* 12(2015):159.
- Peake JM, Kerr G, Sullivan JP. "A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations". *Front Physiol* 9(2018):743.
- Impellizzeri FM, Marcora SM, Coutts AJ. "Internal and external training load: 15 years on". *Int J Sports Physiol Perform* 14(2019):270-273.
- Saw AE, Main LC, Gastin PB. "Monitoring the athlete training response: subjective self-reported measures trump commonly used objective measures: a systematic review". *Br J Sports Med* 50(2016):281-291.
- Vanrenterghem J, Nedergaard NJ, Robinson MA, Drust B. "Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways". *Sports Med* 47(2017):2135-2142.
- Shcherbina A, Mattsson CM, Waggott D, et al. "Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort". *J Pers Med* 7(2017):3.
- Wallen MP, Gomersall SR, Keating SE, Wisloff U, Coombes JS. "Accuracy of heart rate watches: implications for weight management". *PLoS One* 11(2016):e0154420.
- Cadmus-Bertram L. "Using fitness trackers in clinical research: what nurse practitioners need to know". *J Nurse Pract* 13(2017):34-40.
- Ferguson T, Olds T, Curtis R, et al. "Effectiveness of wearable activity trackers to increase physical activity and reduce sedentary time: a systematic review and meta-analysis". *Lancet Digit Health* 4(2022):e615-e626.
- Brickwood KJ, Watson G, O'Brien J, Williams AD. "Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis". *JMIR Mhealth Uhealth* 7(2019):e11819.
- Jo A, Coronel BD, Coakes CE, Mainous AG 3rd. "Is there a benefit to patients using wearable devices such as Fitbit or health apps on mobiles? A systematic review". *Am J Med* 132(2019):1394-1400.
- Sperlich B, Holmberg HC. "Wearable, yes, but able...? It is time for evidence-based marketing claims!". *Br J Sports Med* 51(2017):1240.
- Seshadri DR, Li RT, Voos JE, et al. "Wearable sensors for monitoring the physiological and biomechanical profile of the athlete". *NPJ Digit Med* 2(2019):71.
- Piwek L, Ellis DA, Andrews S, Joinson A. "The rise of consumer health wearables: promises and barriers". *PLoS Med* 13(2016):e1001953.
- Case MA, Burwick HA, Volpp KG, Patel MS. "Accuracy of smartphone applications and wearable devices for tracking physical activity data". *JAMA* 313(2015):625-626.
- Duking P, Fuss FK, Holmberg HC, Sperlich B. "Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity". *JMIR Mhealth Uhealth* 6(2018):e102.
- Malone JJ, Lovell R, Varley MC, Coutts AJ. "Unpacking the black box: applications and considerations for using GPS devices in sport". *Int J Sports Physiol Perform* 12(2017):S218-S226.
- Windt J, Gabbett TJ. "How do training and competition workloads relate to injury? The workload-injury aetiology model". *Br J Sports Med* 51(2017):428-435.
- Foster C, Florhaug JA, Franklin J, et al. "A new approach to monitoring exercise training". *J Strength Cond Res* 15(2001):109-115.
- Aroganam G, Manivannan N, Harrison D. "Review on wearable technology sensors used in consumer sport applications". *Sensors (Basel)* 19(2019):1983.

26. Li RT, Kling SR, Salata MJ, Cupp SA, Sheehan J, Voos JE. "Wearable performance devices in sports medicine". *Sports Health* 8(2016):74-78.
27. Nelson BW, Low CA, Jacobson N, Arian P, Torous J, Allen NB. "Guidelines for wrist-worn consumer wearable assessment of heart rate in biobehavioral research". *NPJ Digit Med* 3(2020):90.
28. Bunn JA, Navalta JW, Fountaine CJ, Reece JD. "Current state of commercial wearable technology in physical activity monitoring 2015-2017". *Int J Exerc Sci* 11(2018):503-515.
29. Hicks JL, Althoff T, Sosic R, et al. "Best practices for analyzing large-scale health data from wearables and smartphone apps". *NPJ Digit Med* 2(2019):45.
30. Gabbett TJ, Nassis GP, Oetter E, et al. "The athlete monitoring cycle: a practical guide to interpreting and applying training monitoring data". *Br J Sports Med* 51(2017):1451-1452.

Cite this article: Sanguinetti-Mora O, Lepasaar K, Chaumergy B. Load Monitoring With Consumer Wearables in Amateur Endurance Athletes: A Review of Signal, Noise, and Behaviour Change. *J Sports Med Pain Manag*. 2024; 5: 031-036.

Ethics Approval and Consent to Participate

Not applicable. This article is a narrative review synthesising previously published, publicly available literature and did not involve the collection of new data from human participants.

Search Strategy

Sources were identified through targeted searches of MEDLINE, Embase, and SPORTDiscus, combined with manual forward and backward citation tracking from the foundational wearable-validation literature, the training-load monitoring consensus and framework literature, and the behaviour-change trial literature for consumer wearables, restricted to peer-reviewed studies reported in English.