



QUALITY IN SPORT

eISSN 2450-3118 · Open Access · Peer-reviewed

apcz.umk.pl/QS Nicolaus Copernicus University in Toruń



STANEK, Justyna Anna, PERŁOWSKA, Aleksandra Anna, SUKAR, Magdalena Maria, MARSZALEK, Katarzyna, PÓLCHŁOPEK, Justyna Agnieszka, WEREMCZUK, Karolina, WOJCIECHOWSKA, Klaudia, RADŁOWSKA, Natalia Barbara, STYŚ, Katharina-Sophia and OSENKOWSKI, Tomasz Jakub. Signals Are Not Symptoms: Can Wearable Analytics Validly Detect Burnout in University Student-Athletes? A Critical Review of Construct Validity, Longitudinal Evidence, and Data Governance. *Quality in Sport*. 2026;78:76448. <https://doi.org/10.12775/QS.2026.78.76448>

ARTICLE TIMELINE

Received: 04.10.2026. Accepted: 15.10.2026. Published: 25.10.2026.

The journal has been awarded 20 points in the parametric evaluation by the Polish Ministry of Higher Education and Science (Annex to the announcement of 05.01.2024, No. 32553). Unique Journal Identifier: 201398. Scientific disciplines assigned: Economics and Finance (Field of Social Sciences); Management and Quality Sciences (Field of Social Sciences).

Punkty Ministerialne z 2019 – aktualny rok 20 punktów. Załącznik do komunikatu Ministra Szkolnictwa Wyższego i Nauki z dnia 05.01.2024 Lp. 32553. Posiada Unikatowy Identyfikator Czasopisma: 201398. Przypisane dyscypliny naukowe: Ekonomia i finanse (Dziedzina nauk społecznych); Nauki o zarządzaniu i jakości (Dziedzina nauk społecznych). © The Authors 2026.

OPEN ACCESS · CC BY-NC-SA 4.0 This article is published with open access under the License Open Journal Systems of Nicolaus Copernicus University in Toruń, Poland, and is distributed under the terms of the Creative Commons Attribution Non-commercial Share Alike License (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the work is properly cited. The authors declare no conflict of interest regarding the publication of this paper.

Signals Are Not Symptoms: Can Wearable Analytics Validly Detect Burnout in University Student-Athletes?

A Critical Review of Construct Validity, Longitudinal Evidence, and Data Governance

AUTHORS & AFFILIATIONS

Justyna Anna Stanek

ORCID: <https://orcid.org/0009-0009-7256-1695>

E-mail: justynaannastanek@wp.pl

University of Rzeszów: Rzeszów, Subcarpathia, PL

Aleksandra Anna Perlowska

ORCID: <https://orcid.org/0009-0002-9479-7665>

E-mail: olaperlowska@gmail.com

University of Rzeszów: Rzeszów, Subcarpathia, PL

Magdalena Maria Sukar

ORCID: <https://orcid.org/0009-0008-0904-6810>

E-mail: magda.sukar@gmail.com

University of Rzeszów: Rzeszów, Subcarpathia, PL

Katarzyna Marszałek

ORCID: <https://orcid.org/0009-0005-1313-3739>

E-mail: kmarszalek29@gmail.com

University of Rzeszów: Rzeszów, Subcarpathia, PL

Justyna Agnieszka Pólchlopek

ORCID: <https://orcid.org/0009-0006-4827-8193>

E-mail: justyna.polchlopek1205@gmail.com

University of Rzeszów: Rzeszów, Subcarpathia, PL

Karolina Weremczuk

ORCID: <https://orcid.org/0009-0008-4292-1285>

E-mail: kalawer@gmail.com

University of Rzeszów: Rzeszów, Subcarpathia, PL

Klaudia Wojciechowska

ORCID: <https://orcid.org/0009-0000-7978-3033>

E-mail: Wojciechowska301@gmail.com

University of Rzeszów: Rzeszów, Subcarpathia, PL

Natalia Barbara Radłowska

ORCID: <https://orcid.org/0009-0004-3919-7696>

E-mail: nm098421@stud.ur.edu.pl

University of Rzeszów: Rzeszów, Subcarpathia, PL

Katharina-Sophia Styś

ORCID: <https://orcid.org/0009-0002-7757-2034>

E-mail: ks113954@stud.ur.edu.pl

University of Rzeszów: Rzeszów, Subcarpathia, PL

Tomasz Jakub Osenkowski

ORCID: <https://orcid.org/0009-0009-3923-0034>

E-mail: t.osenkowski1@gmail.com

University of Rzeszów: Rzeszów, Subcarpathia, PL

Corresponding Author

Justyna Anna Stanek, E-mail: justynaannastanek@wp.pl

Abstract

Background: Wearables estimate physiological and behavioral signals, but burnout is a chronic, multidimensional psychological construct.

Aim: This review assesses whether wearable analytics can detect burnout in university student-athletes.

Approach: A critical narrative synthesis examines athlete-burnout research, psychophysiology, student wearable sensing, and data governance; it is not a systematic review or meta-analysis.

Synthesis: Direct evidence is sparse. Burnout studies mainly rely on self-report, while wearable-AI studies mostly target acute stress, fatigue, sleep, or affect in other populations. Physiological markers remain exploratory, with no accepted diagnostic thresholds. Collegiate studies also raise concerns about data sharing and athlete control.

Conclusion: Wearables may inform voluntary recovery conversations but should not diagnose burnout or determine selection, eligibility, or care. Multisite, season-long validation using repeated burnout measures and athlete-governed data practices is needed.

Keywords: athlete burnout; student-athletes; wearable technology; heart-rate variability; digital monitoring; construct validity; privacy

1. Introduction

University student-athletes manage two linked but distinct sets of demands. Training, competition, injury, selection, and recovery occur alongside coursework, examinations, finances, relationships, and decisions about future careers. The combination can create sustained stress, but stress is not itself burnout. Athlete burnout is typically described as a cognitive-affective syndrome involving emotional and physical exhaustion, reduced sense of accomplishment, and sport devaluation [1], [2], [3]. It is therefore possible to be acutely tired after a demanding week without being burned out, and possible to experience sport devaluation or reduced accomplishment without an unusually high heart rate on a given day.

Wearable systems promise continuous information about heart rate, heart-rate variability (HRV), sleep, activity, and training exposure. Machine-learning models can combine such data and produce estimates of stress, fatigue, readiness, or risk. In a university sport setting, those estimates may appear useful: a coach or practitioner might hope to notice accumulating strain before an athlete withdraws, becomes injured, or needs clinical support. Yet the claim that a

device can detect burnout is much stronger than the claim that it can measure a physiological response associated with exertion or acute stress.

Recent research reviews cover adjacent parts of this problem, but not the whole question. A scoping review identified 29 studies of burnout among 8,022 intercollegiate athletes; most were cross-sectional, and the review did not examine wearables or physiological biomarkers [4]. A systematic review of smartwatch-based burnout detection in students included school and higher-education populations but did not analyze student-athletes as a distinct group; stress was the most common burnout-related target [5]. Reviews of AI-enabled wearables in athletes focus on transient psychological states, such as stress, emotion, attention, and readiness, rather than the chronic syndrome of burnout [6], [7]. A broader review of AI and wearables in sport addresses athlete well-being and governance, but does not establish burnout detection in university student-athletes [8].

The ethical literature forms another adjacent strand. Reviews and empirical studies describe athlete data collection, privacy, institutional monitoring, and data access, while often treating well-being as a broad concern rather than measuring burnout as an outcome [9], [10], [11]. The overlap among these areas makes it tempting to join them into a single story: sensors detect strain, algorithms predict burnout, and monitoring either supports or surveils athletes. The evidence does not yet support that complete causal chain.

This review takes a narrower position. It asks whether wearable-derived signals have been validated against an appropriate measure of athlete burnout in university student-athletes, whether they distinguish burnout from nearby states, and whether the surrounding data practices allow monitoring to be used in a way that protects athlete agency. The central argument is that wearable data can add context to conversations about training and recovery, but they do not currently provide a stand-alone measure or diagnosis of burnout. A signal is not a symptom, and a prediction is not a clinical or educational judgment.

This distinction matters for practice. If a system labels an athlete “high risk” based on poor sleep, an HRV change, or reduced training output, the label may affect coach behavior, team status, or the athlete’s own understanding of their health. If the model is wrong, the consequences are not limited to a technical error. Conversely, if a potentially useful pattern is hidden because an athlete fears consequences from disclosure, the monitoring system may fail at its stated purpose. The value of an analytic tool therefore depends on construct validity, longitudinal evidence, interpretation, and governance—not on sensor availability alone.

2. Review approach and evidence boundaries

This article is a critical narrative review. Scholarly-paper searches and citation chaining were used to identify work on university student-athlete burnout, athlete burnout measurement, wearable and physiological monitoring, student stress detection, AI or machine learning, and athlete data governance. The evidence base was updated through October 2, 2026. Priority was given to systematic and scoping reviews, full-text studies directly involving university or collegiate athletes, and studies that clarify measurement or data-sharing practices.

The review distinguishes five evidence groups. Direct evidence includes research on university student-athletes that measures burnout and uses wearable data in the same participants. Athlete-burnout evidence includes studies of student-athletes or athletes using validated burnout questionnaires without wearable monitoring. Wearable-psychophysiology evidence includes athlete monitoring of stress, fatigue, recovery, sleep, or affect without a validated burnout outcome. Student-sensing evidence includes wearable-AI research on stress or mental-health indicators in general student populations. Governance evidence includes research on collegiate or elite-athlete data practices, privacy, and athlete experience.

These groups are not interchangeable. A machine-learning model that classifies laboratory-induced stress in general college students does not establish that a model can detect chronic burnout in student-athletes. A study of athlete sleep and readiness does not establish that its outcome is burnout. A review of athlete burnout correlates does not test sensors. This review therefore treats population, construct, measurement tool, duration, and device role as separate dimensions.

No protocol was registered. Searches were iterative rather than exhaustive across a fixed set of named databases, screening was not conducted independently by two reviewers, and no formal risk-of-bias tool was applied across the heterogeneous evidence. No meta-analysis was attempted. These limits mean that this article cannot be presented as a systematic census of every study. Its purpose is to identify what the available evidence can and cannot support and to set out a defensible standard for future claims.

2.1 Evidence tiers used in the synthesis

The phrase “wearable burnout detection” can refer to several different research designs. Some studies use sensors to detect acute stress and then discuss possible relevance to burnout. Others study wearables and athlete well-being without assessing burnout. A third group measures

burnout but has no sensor data. Only a study that links wearable-derived features to repeated, validated burnout measures in the intended athlete population can directly evaluate the proposed detection claim.

Evidence tier	Population and outcome	What it can establish	What it cannot establish
Direct validation	University student-athletes, wearable data, repeated validated burnout assessment	Whether sensor features track or predict burnout dimensions in this population	Clinical diagnosis unless diagnostic criteria and validation are also established
Athlete burnout research	Athletes assessed with burnout scales, without wearable data	Burnout dimensions, correlates, and possible time course	Sensor validity or algorithm performance
Athlete monitoring research	Athletes monitored for load, stress, fatigue, sleep, or readiness	Feasibility and measurement of selected physiological or performance states	Burnout detection unless burnout is explicitly measured
Student wearable-AI research	General student populations, often acute stress	Technical feasibility for the measured task and sample	Transfer to student-athletes or chronic burnout
Governance research	Athlete data practices and experiences	How access, power, consent, and interpretation shape monitoring	The physiological validity of a burnout predictor

The available literature is substantial in the second through fifth tiers. The first tier remains the decisive missing test. The review therefore does not interpret technical model accuracy in adjacent samples as evidence of burnout prediction.

3. What athlete burnout means—and what it does not mean

3.1 A multidimensional construct

The Athlete Burnout Questionnaire (ABQ) was developed to assess three dimensions: physical and emotional exhaustion, reduced sense of accomplishment, and sport devaluation [1]. These dimensions capture different experiences. Exhaustion concerns depletion; reduced accomplishment concerns the athlete’s appraisal of competence or progress; sport devaluation

concerns a more detached or cynical relationship with participation. The dimensions can co-occur, but they need not rise and fall together.

The three-part account has shaped much of the sport literature, while other traditions emphasize stress appraisal, sport commitment, motivation, identity, or the organization of sport [2], [3], [12]. A broad review published in *Pedagogy and Psychology of Sport* describes sport psychology as spanning research, assessment, intervention, counseling, and mental training across physical-culture settings [13]. It provides disciplinary context for this review, not evidence about wearable validity or burnout detection. The cognitive-affective stress model describes a process in which demands, appraisal, coping, and response interact over time [14]. Sociological accounts emphasize restricted autonomy and a narrow athletic identity; commitment-based accounts focus on feeling trapped in sport rather than freely choosing it [2], [3]. Self-determination theory links burnout to the frustration of autonomy, competence, and relatedness needs [15]. These perspectives differ, but they share a point important for sensor research: burnout is not reducible to a single physiological variable.

The ABQ remains the most common measure in athlete-burnout studies. In the recent scoping review of university student-athletes, it was used in 27 of 29 included studies [4]. A study of 267 NCAA Division I athletes also used its three subscales to examine differences by sex, eligibility year, early specialization, and injury history [16]. Such instruments capture athletes' reported experience over a reference period. They do not turn a high score into a clinical diagnosis by themselves.

The absence of accepted diagnostic cutoffs is a major constraint. Reviews of athlete burnout have noted that studies often report questionnaire scores or prevalence estimates without a universally agreed threshold that separates a clinical case from a high symptom score [3], [17]. This affects any proposed wearable model: a classifier needs a target label. If the label is a questionnaire score, the model predicts that score or category, not an independently established disease state. If the label is a researcher-created cutoff, its validity must be justified rather than assumed.

3.2 Burnout is not acute stress, fatigue, or overtraining

Burnout overlaps with fatigue, stress, sleep difficulty, depressed mood, and training maladaptation, but these terms are not synonyms. Acute stress is a time-bounded response to a demand. Fatigue can follow a single training session, travel, illness, sleep loss, or academic

workload. Overtraining syndrome concerns maladaptation to training and is usually evaluated through sustained performance impairment and a broader clinical picture. Burnout includes attitudinal and motivational features, especially sport devaluation and reduced accomplishment, that are not directly observable from a heart-rate trace [3], [17], [18].

A physiological marker may be relevant to exhaustion without measuring the other dimensions. For example, low HRV or altered cortisol might reflect autonomic or endocrine regulation under some conditions. Even if a relationship with exhaustion were replicated, it would not automatically indicate sport devaluation or perceived accomplishment. A model trained on a general “stress” label could therefore detect transient arousal while being described incorrectly as a burnout detector.

Burnout also unfolds over time. A high load on Tuesday, a poor night of sleep, or an unusual HRV value may prompt a short-term recovery conversation. Burnout, by contrast, implies a more sustained pattern of strain and altered engagement. The gap between the time scale of sensor data and the time scale of the target construct is a core measurement problem. Continuous data do not guarantee that a chronic construct is being measured continuously.

3.3 Burnout versus neighboring clinical and sport constructs

The boundaries among athlete burnout, depressive symptoms, anxiety, sleep problems, overreaching, and overtraining are not always clean in applied settings. The constructs can co-occur and can share exhaustion, negative affect, or impaired functioning. Athlete burnout research has examined links to mental and physical health outcomes, while work on overtraining highlights overlapping fatigue and performance change [17], [19]. A wearable model that flags “high risk” without clarifying its target may blur these distinctions and send athletes toward the wrong support pathway.

The practical solution is not to demand a perfect one-sensor definition. It is to define the question. Is the system intended to identify short-term physiological strain? To prompt a check-in? To forecast questionnaire-measured burnout symptoms over a season? To make a clinical referral? Each purpose requires a different target, validation standard, and response to errors. If the intended result is only a contextual prompt, it should be called a prompt rather than a diagnosis.

4. What is known about burnout in university student-athletes

4.1 Prevalence and design of the evidence base

The most direct overview is the scoping review of intercollegiate athlete burnout. It included 29 studies and 8,022 athletes; 22 studies were cross-sectional and only five were longitudinal [4]. Most studies used the ABQ. The review organized associated factors into demographic characteristics, personal characteristics, psychosocial states, and interpersonal factors. It also noted that the largely cross-sectional literature cannot determine whether correlates precede burnout, follow it, or reinforce it.

A separate study of NCAA Division I athletes illustrates the limits of a typical cross-sectional snapshot. Among 267 respondents at two institutions, women reported higher exhaustion scores than men, upperclassmen reported greater sport devaluation than underclassmen, and athletes with an injury history reported lower accomplishment; early sport specialization did not distinguish burnout scores in that sample [16]. These are group associations, not individual predictions. They do not show that sex, eligibility year, or injury history alone can identify which athlete will become burned out.

Burnout research in team sports is larger but not fully specific to university sport. A systematic review and meta-analysis of 59 papers, representing 55 samples and more than 11,000 athletes, linked burnout dimensions to motivation, affect, perfectionism, coaching climate, autonomy, and social support [15]. A meta-analysis of social constructs found that support and relatedness were inversely associated with burnout, while negative social interactions showed positive associations [20]. These findings support the relevance of the athlete's social environment, but most are observational and cannot be read as simple causal effects.

4.2 Predictors and protective factors

Across student-athlete and team-sport studies, perfectionistic concerns, negative affect, amotivation, controlling coaching, sport-specific stress, and low social support recur as risk-related correlates. Positive associations have been reported between perfectionistic concerns and subsequent burnout dimensions in college athletes across a three-month competitive season [21]. Earlier work distinguishes perfectionistic concerns from high personal standards: the former are more consistently linked with maladaptive outcomes, while striving for demanding goals can have a different association [22], [23]. Studies of collegiate athletes have also examined grit, gratitude, social support, and well-being as potential protective factors [24], [25].

The distinction between sport demands and other student demands matters. Student-athletes may face academic deadlines, financial pressure, social expectations, and sport obligations at the same time. Research on dual careers and student-athlete burnout describes the possibility that pressure can arise from the interaction of educational and sport systems, rather than from training load alone [26]. University samples and populations differ by country, sport structure, scholarship status, and institutional support; findings from one context should not be treated as a universal student-athlete profile [27], [28].

Interpersonal conditions are also important. Coaching behavior and motivational climate have been examined in collegiate athletes, including relationships with burnout and competitive anxiety [29], [30]. Across broader team-sport evidence, autonomy-supportive coaching and social support tend to be protective correlates, whereas controlling practices and negative social climates tend to align with higher burnout scores [15], [20]. These patterns argue against a purely individual or biometric explanation. A wearable score cannot show whether an athlete feels trapped, unheard, or unable to ask for reduced load.

4.3 The limited longitudinal evidence

Burnout is a process, so repeated measurement is essential. A scoping review of 32 longitudinal athlete-burnout studies found that the ABQ was used in most studies and that measurement points often followed the competitive season or occurred at intervals of several months [31]. The reviewed work suggests that dimensions can move differently over a season; exhaustion does not necessarily change in parallel with devaluation or accomplishment. The review also found that most studies relied on self-report and that more qualitative and culturally diverse work is needed.

The three-month study of college athletes found that perfectionistic concern cognitions predicted increases in reduced accomplishment and devaluation, but not exhaustion [21]. That dimension-specific result is useful for wearable research: a model trained only against “total burnout” might hide different time courses. A signal associated with exhaustion should not be assumed to predict devaluation. Likewise, a weekly readiness metric cannot stand in for a season-long change in the athlete’s relationship to sport.

Sleep research provides a caution about interpreting physiological data. In a sample of 257 young elite athletes, higher self-reported burnout symptoms were associated with subjective sleep complaints and less time in bed on weekdays. Yet the polysomnographic subsample

showed no significant difference in objective sleep parameters between athletes with higher and lower burnout symptoms. In longitudinal analyses, burnout predicted later insomnia symptoms, while baseline insomnia did not predict later burnout in the tested model [32]. This was not a wearable-validation study and the athletes were younger than university populations, but it demonstrates that subjective distress and objective sleep physiology can diverge.

4.4 Consequences and the role of screening

Burnout can be associated with disengagement, reduced performance, and sport withdrawal; reviews also connect it with broader mental-health concerns [2], [3], [19]. But the evidence on consequences, trajectories, and treatment is not equivalent to a validated screening system. A screening tool must show what it detects, how often it misses cases or raises false alarms, and what happens after a positive result. Athlete burnout questionnaires can structure assessment, but absent agreed diagnostic thresholds, a score should be interpreted as symptom information within an appropriate clinical or sport-psychology process [1], [3].

Wearable monitoring could eventually contribute to early support if it identifies sustained changes that precede worsening self-reported symptoms. That is a hypothesis to test, not a result established by the current literature. The distinction between screening and surveillance is partly determined by what follows the signal: a confidential, voluntary check-in differs from an automated label visible to coaches or selection staff.

5. What wearable analytics measure

5.1 Signals, estimates, and inferred states

Wearables record signals, not psychological constructs. Depending on the device, these may include photoplethysmography, electrocardiography, accelerometry, temperature, electrodermal activity, GPS-derived movement, or sleep-related estimates. Software then transforms the raw stream into measures such as heart rate, HRV, steps, sleep duration, training load, or readiness. A machine-learning model can combine features and output a classification or probability.

Each step introduces choices. Sampling frequency, artifact removal, missing-data rules, body placement, device firmware, and algorithm version can alter the estimate. A proprietary “recovery score” is not a direct observation of recovery, and an “energy” or “readiness” score is not a validated measure of burnout. Reviews of wearable monitoring in sport describe diverse

sensor systems and applications for internal and external workload, while stressing that interpretation depends on the activity and protocol [33], [34], [35].

In athletic settings, HRV is often used to characterize autonomic regulation and recovery. It is sensitive to posture, time of day, breathing, recent exercise, illness, sleep, hydration, medication, and measurement artifacts. Resting HRV may contribute useful within-person context, but group-level comparisons and one-off readings are less informative when individual baselines differ. Athlete-monitoring reviews increasingly recommend pairing physiological information with subjective reports rather than allowing either channel to substitute for the other [36], [37].

5.2 The proposed physiology of burnout

The hypothalamic-pituitary-adrenal axis and autonomic nervous system are plausible areas of study because chronic demands can affect stress regulation. Candidate indicators include salivary cortisol patterns and vagally mediated HRV, particularly when measured under standardized conditions [18]. Yet the supporting evidence is mixed, and the markers are not specific to burnout. They may reflect sleep, training, acute stress, illness, nutrition, menstrual-cycle effects, or other conditions.

The physiological perspective on athlete burnout has long warned against treating a biological response as a settled diagnostic criterion. Related work on overtraining has examined endocrine, immune, autonomic, and performance changes, but findings have not produced a universally accepted predictive marker [17]. The recent critical review of psychophysiological markers similarly describes promising hypotheses alongside small samples, inconsistent timing, heterogeneous burnout definitions, and limited longitudinal evidence [18].

A useful model is therefore not “burnout equals low HRV” or “burnout equals poor sleep.” It is “some physiological measures may covary with some dimensions of burnout under certain conditions, and the relationship may depend on personal baseline and context.” That statement is less dramatic, but closer to the evidence.

5.3 Technical prediction is not construct validity

A model can achieve high accuracy on a dataset and still fail to measure the construct that matters. Performance depends on how labels are assigned, how participants are sampled, how data are split between training and test sets, and whether validation includes new people, teams,

seasons, and devices. If a model distinguishes an exam-induced stress condition from rest, it may be learning the experimental task, not a chronic burnout process.

A meta-analysis of wearable AI for student stress detection reported a pooled accuracy estimate around 0.86 across selected model results, but performance varied with the number of stress classes, device placement, sample size, and ground-truth method. The authors concluded that performance remained suboptimal and recommended using wearable AI alongside established assessments rather than as a replacement [38]. This is evidence about student stress classification, not about athlete burnout.

A 2026 scoping review of wearable stress measurement in college students found heavy reliance on existing datasets, with repeated use of small benchmark samples, inconsistent demographic reporting, and limited sample-size justification [39]. This further weakens simple transfer from technical demonstration to campus practice. Stress detection and burnout prediction require different reference outcomes and time horizons.

Reviews of athlete AI and wearable sensing report a similar translation gap. Athlete studies often focus on short sessions, laboratory settings, or transient states such as emotion and stress; small samples and limited external validation constrain claims about individualized prediction [6], [7]. A broad sport review likewise emphasizes the need for model validation, contextual interpretation, practitioner expertise, and responsible governance [8]. High within-study accuracy should not be interpreted as evidence that a model generalizes to student-athletes in ordinary university sport.

5.4 Heart-rate variability and autonomic features

HRV is attractive for field monitoring because it can be sampled repeatedly with relatively low burden. In sport practice, it is often interpreted as an indicator of autonomic regulation or readiness. That interpretation is conditional: HRV is affected by the measurement protocol and by recent physiological demands, and a change does not identify its cause. The same athlete may show a different value after a hard session, a short night, illness, travel, or a stressful examination week. Those events may all matter to well-being, but they are not interchangeable with burnout.

The critical review of psychophysiological burnout markers discusses resting vagally mediated HRV as a candidate signal, while emphasizing heterogeneous results and the lack of adequate longitudinal athlete studies [18]. A review of athlete monitoring finds HRV widely used across

competitive sports but notes that objective and subjective channels capture partly overlapping processes rather than one unified state [36]. This supports using HRV as contextual information, not as a direct index of exhaustion, devaluation, or reduced accomplishment.

In a validation study, the model should establish a standardized recording context, report signal quality and artifact handling, and distinguish an acute response from a persistent deviation. The analysis should test whether HRV adds information beyond an athlete's own symptom report and workload history. A group threshold should not be substituted for a personal reference range, and a personal baseline should not be interpreted without context. If the same metric changes with training adaptation as well as distress, the model must test which interpretation is supported by the outcome data.

5.5 Sleep, activity, and workload features

Sleep estimates are appealing because sleep disturbance may accompany stress and impaired recovery. Consumer wearables, however, infer sleep from movement and physiological signals rather than directly measuring all stages with laboratory polysomnography. A reported sleep duration or recovery score can be useful for tracking routines, but its precision and meaning depend on device and algorithm. The study of young elite athletes found that subjective insomnia and burnout symptoms were associated, while objective polysomnographic parameters did not significantly distinguish the higher-symptom group [32]. This finding does not show that wearable sleep data are useless; it shows that objective sleep features cannot be presumed to stand in for subjective burnout experience.

Activity and workload measures have a different role. GPS, accelerometry, session duration, and training-load summaries describe exposure or movement. They may help identify an unusually dense training period or a mismatch between load and recovery. They do not measure how the athlete interprets the demands, whether participation feels voluntary, or whether sport has lost meaning. Reviews of athlete wearable systems describe broad use for internal and external workload monitoring, but the value of each metric depends on sport, device, and decision context [33], [34].

The most plausible future approach is not to merge all streams into an opaque wellness number. It is to retain a clear distinction between workload, sleep, autonomic features, self-reported recovery, and burnout dimensions. Researchers can then test which signals are related to which

outcomes and whether the association is useful in a defined time window. A composite score that masks its inputs may be easier to display but harder to validate and challenge.

6. Evidence at the intersection of student-athletes, wearables, and burnout

6.1 What is directly available

The direct intersection remains thin. A study proposing passive mobile and wearable sensing for student-athlete condition monitoring collected four weeks of behavior and stress-recovery scale data from 19 student-athletes. It examined training intensity and lifestyle patterns and described a platform as a basis for later machine-learning prediction; it did not validate a wearable model against a burnout measure [40]. This is relevant feasibility work, but not evidence that wearables detect burnout.

A related intervention study in Pedagogy of Physical Culture and Sports tested 12 sessions of VR-based Tabata training with 40 male university swimmers in Indonesia. The experimental group showed greater improvements in physical fitness and self-reported psychological well-being than the control group, while interviews noted both enjoyment and practical limits such as equipment access and the need for a safe space [41]. This paper broadens the review's educational-sport context: it concerns a digital training intervention and well-being, but it did not use wearable sensing, AI, or a burnout measure, so it is not direct evidence for wearable burnout detection.

A review of wearable technology in youth athletes describes physiological, biomechanical, performance, and safety applications, including discussion of internal workload and potential burnout concerns [42]. Its broad scope is useful for locating wearable use in youth sport, but it does not establish a student-athlete burnout biomarker or an algorithm with validated diagnostic performance.

The student smartwatch review is the closest evidence on wearable-based student burnout symptoms. It included students from childhood through young adulthood and reported that stress was the most frequently studied burnout-related signal, with HR and HRV common among the physiological measures [5]. The population was not a distinct student-athlete cohort, and the review itself notes the gap between stress detection and evidence that burnout symptoms can be detected during academic study. The word “burnout” in the review objective should therefore not be read as proof that the included models diagnose the full syndrome.

6.2 Adjacent athlete-wearable evidence

Athlete monitoring studies commonly assess training load, sleep, recovery, readiness, injury risk, and acute stress. A recent review of psychophysiological monitoring in competitive athletes synthesized 22 studies and found HRV to be the dominant physiological measure; physiological and subjective indicators were directionally consistent in some settings but differed in magnitude and timing [36]. Burnout was not the primary construct across that evidence base. The review supports multimodal, dual-channel monitoring for athlete state, not wearable diagnosis of burnout.

AI-enabled wearable reviews of athletes examine cognitive state, affect, stress, or readiness. The systematic review by Mia and Soniawan found small, often laboratory-based athlete studies and emphasized the difficulty of distinguishing psychological stress from the physiological noise of exercise [6]. A related review of machine-learning and multimodal AI similarly highlights construct heterogeneity and external-validation gaps [7]. Neither review establishes a model trained and validated on longitudinal athlete-burnout outcomes in student-athletes.

6.3 Adjacent student-wearable evidence

Student sensing research provides methodological lessons but not direct population transfer. The systematic review of wearable AI for student stress detection found variable performance and problems in the ground truth used to label stress [38]. College-student reviews document the range of devices, sensors, and machine-learning methods in stress measurement, while also noting dataset and demographic limitations [39]. Research on wearable and smartphone sensing has also explored stress and mental-health indicators outside sport, but such results cannot by themselves establish validity for athletes balancing sport and study [43], [44], [45].

Studies of prolonged stress using wearable biosensors and ecological momentary assessments are potentially useful because they link repeated physiological measurements to repeated self-report in everyday settings [46], [47]. Their design logic can inform future athlete research. However, chronic stress remains distinct from athlete burnout, and a model must still be validated against the specific burnout dimensions it claims to predict.

6.4 Evidence map

Research strand	Typical population	Common target	Relevance to this review	Main limitation for burnout claims
University athlete burnout	Intercollegiate athletes	ABQ dimensions and psychosocial correlates	Defines the population and construct [4], [16]	Mostly questionnaire-based; little sensor data
Athlete psychophysiology	Competitive athletes	HRV, stress, sleep, recovery, affect	Identifies candidate signals and confounders [18], [36]	Acute or adjacent outcomes; no validated burnout model
Student wearable AI	General student samples	Acute or perceived stress	Offers model-validation lessons [38], [39]	Not athlete-specific; stress is not burnout
Collegiate athlete sensing	Student-athlete samples	Training, lifestyle, stress-recovery state	Shows feasibility of passive sensing [40]	Small preliminary sample; no burnout endpoint
Athlete data governance	Collegiate and elite sport	Access, consent, sharing, privacy	Shows how context shapes use [11], [48]	Usually not linked to measured burnout outcomes

6.5 The central evidence gap

No available evidence in this synthesis demonstrates that wearable analytics can diagnose or reliably predict athlete burnout among university student-athletes. This is not a claim that no such work exists anywhere; it is a statement about the evidence found in the current workspace research and the closely related indexed literature examined for this review. The available work either studies burnout without sensors, sensors without burnout, or stress-related indicators in populations that differ from student-athletes.

This separation prevents a common inferential leap. If a wearable estimates stress, and stress is associated with burnout, it does not follow that the wearable measures burnout. A plausible causal pathway is not a validated measurement model. A strong article or intervention must

report the target outcome, the time window, the population, the comparator, and what the model's prediction is intended to change.

6.6 Why adjacent evidence cannot be combined as if it were direct evidence

The evidence streams are complementary, but they answer different questions. The university athlete burnout literature describes self-reported symptoms and correlates; the wearable literature estimates physiological or behavioral signals; the student-AI literature tests classification of stress; and the governance literature describes how data are handled. Joining these findings into a single claim that a sensor can detect student-athlete burnout would treat several untested links as if they had already been demonstrated.

Source	Population and measured outcome	Contribution	Boundary on interpretation
[4]	Intercollegiate athletes; burnout questionnaires	Maps burnout correlates and study designs	Does not test sensors or biomarkers
[40]	19 student-athletes; passive sensing and stress-recovery reports	Demonstrates a small-scale sensing platform	Does not validate burnout prediction
[5]	General student population; smartwatch and burnout-related symptoms	Maps student wearable work and common signals	Student-athletes are not a distinct subgroup; stress dominates
[38]	Students; AI classification of stress	Reports model-performance and ground-truth concerns	Stress classification is not chronic athlete burnout
[6]	Athletes; AI/wearable inference of psychological states	Maps technical feasibility and validation gaps	Burnout and student-athlete status are not the review's target
[11]	Collegiate athletes; wearable-data information flows	Shows how data access is experienced	Does not test burnout prediction or health effects

The table also clarifies how the review uses evidence. Direct claims about the construct and its correlates rely on athlete-burnout studies. Claims about sensor interpretation rely on

psychophysiology and device-validation research. Claims about athlete experience rely on qualitative and ethical studies. Where the evidence does not cross these domains, the article identifies the gap rather than inferring a result.

6.7 AI outputs, thresholds, and the meaning of a risk label

Machine-learning systems can produce a probability, score, or category. That output is only as meaningful as its target and calibration. A model that predicts a questionnaire score may be useful for research even if it has no diagnostic role. A model that assigns “low,” “moderate,” or “high” risk needs evidence that those categories correspond to different future outcomes and that the chosen threshold supports a worthwhile action.

A high overall accuracy can be misleading when the target outcome is uncommon. If most participants are classified as not having elevated burnout symptoms, a model can achieve a high accuracy by predicting the majority class while missing athletes who need support. Conversely, a sensitive threshold may produce many false alerts. Authors should report sensitivity, specificity, predictive values, calibration, confidence intervals, and the consequences of false positives and negatives at realistic prevalence levels. The student wearable-AI review illustrates that accuracy changes with label definition and ground-truth method [38].

Risk categories should not be created merely because a dashboard requires colors. A red status can carry a social meaning even if its statistical basis is weak. If the purpose is to invite a conversation, neutral wording such as “a recent change worth reviewing” may be more accurate than “burnout risk.” The language should be co-designed with athletes, tested for comprehension, and separated from decisions about selection or discipline.

7. Construct validity: the decisive test

7.1 Define the target before selecting the sensor

A validation study should begin by specifying whether the intended target is (1) same-day physiological strain, (2) perceived stress, (3) physical and emotional exhaustion, (4) total ABQ score, (5) clinically relevant burnout symptoms, or (6) future change in burnout over a season. These are not equivalent endpoints. The time window also needs to be explicit: a five-minute HRV sample cannot be compared meaningfully with a questionnaire asking about experiences across a different period unless the temporal relation is theorized and tested.

The ABQ is a reasonable athlete-specific criterion for symptom dimensions, but there is no universally accepted clinical cutoff [1], [3]. A model trained to classify the upper quartile of ABQ scores would predict membership in that sample-defined quartile. It would not automatically detect a clinical syndrome. Reports should avoid words such as diagnosis, disease detection, or prevention unless the relevant thresholds and clinical follow-up are established.

7.2 Match each sensor to its plausible construct

Heart-rate and HRV features might be tested as correlates of autonomic state. Sleep duration or timing might be tested as a behavioral pattern associated with recovery. GPS or accelerometry may quantify external load. None of these directly records perceived accomplishment or sport devaluation. Those psychological dimensions require direct measures, interviews, or other validated assessment.

This is why composite “wellness” or “readiness” scores need transparency. If the output combines sleep, HRV, and training load, the model may be useful for a narrowly defined decision such as prompting a review of training. But the score should not be relabeled as burnout simply because burnout is one possible consequence of chronic strain. The components and their uncertainty should remain visible to the people affected.

7.3 Distinguish within-person change from between-person ranking

A person’s HRV baseline can differ from another person’s baseline for reasons unrelated to health. For athlete monitoring, within-person departures from a stable personal baseline may be more interpretable than ranking athletes against one another, but within-person change still does not specify cause. A change can reflect a hard training block, illness, sleep loss, travel, academic exams, medication, or sensor error.

A valid model should therefore test whether repeated features add predictive information beyond baseline self-report, training load, sleep, injury, season timing, and relevant contextual factors. It should use person-level or team-level holdout validation rather than randomly splitting records from the same athlete across training and test data. Otherwise, the model may learn each participant’s stable signature and appear to generalize when it does not.

7.4 Calibrate labels and include athlete interpretation

Questionnaire scores, diagnostic interviews, clinician judgments, and momentary self-reports answer different questions. Labels should be pre-specified and their reliability reported.

Repeated ABQ measurement can track symptom dimensions; interviews can clarify whether a score reflects sport devaluation, a temporary conflict, injury, or loss of choice. Athlete commentary can identify contextual events that sensor data do not capture.

A useful design would compare the model with a base case using repeated ABQ scores and established psychosocial indicators. If wearable data do not improve calibration, sensitivity, specificity, or decision usefulness over the base case, continuous collection may not be justified. If they do improve prediction, the next question is whether an intervention based on the model improves outcomes without causing new harms.

7.5 Confounding, subgroup validity, and transportability

Physiological features are shaped by more than psychological state. Recent exercise, illness, sleep opportunity, travel, nutrition, hydration, caffeine, medication, ambient temperature, and measurement posture can alter the same signals proposed for burnout research. For athletes, training phase and sport type further affect the relationship between workload, perceived fatigue, and autonomic recovery [33], [34], [36]. A model that does not measure or account for these conditions may attribute a normal training response to psychological risk.

Confounding is not solved by adding every available variable. More data can increase privacy costs, overfitting, and the chance that a model uses a proxy for a protected or sensitive characteristic. The relevant question is whether each additional feature has a defined role, adequate quality, and a proportionate justification. The physiological-marker literature specifically calls for standardized measurement and control of factors such as sleep, age, fitness, and season timing [18].

Transportability is also uncertain. A model trained on one team may reflect its sport, coaching style, facilities, device type, or reporting norms. It may not perform similarly for athletes in individual sports, athletes with disabilities, athletes recovering from injury, or students whose training and academic schedules differ. Athlete-burnout studies themselves are geographically concentrated and vary in sex representation and measurement practices [4], [31]. Wearable stress reviews report gaps in demographic reporting and small or reused datasets, which further limit claims about fairness across groups [39].

The issue is not only whether average accuracy is lower in one group. A false positive may prompt unwanted scrutiny, while a false negative may reassure staff despite a student's distress. Performance should therefore be reported separately across relevant groups, but small subgroup

samples should not be used to make confident claims. Qualitative work can identify how the system is experienced, while quantitative validation estimates whether its errors differ.

Source of variation	Possible effect on sensor features	Why it matters for burnout inference	Study safeguard
Training phase and sport	Alters load, HRV, sleep, and recovery	May mimic or mask a sustained strain pattern	Record sport-specific phase and test across seasons
Injury, illness, or medication	Changes activity and physiological response	Can create apparent risk unrelated to burnout	Record relevant context with consent; do not treat missingness as normal
Sleep opportunity and travel	Changes sleep timing and autonomic measures	A poor night may be transient rather than chronic burnout	Use repeated measures and contextual diaries
Sex-related and individual physiology	Changes baseline and response patterns	Group thresholds can misclassify individuals	Use person-specific baselines and subgroup validation
Device, firmware, and placement	Changes data quality and derived features	A model may fail after a device or software update	Report model/device version and validate externally
Institutional culture	Shapes self-report, compliance, and disclosure	Athletes may conceal symptoms where data are high stakes	Keep assessment confidential and document power context

These limitations make a one-size-fits-all burnout threshold particularly difficult to defend. Personalization can improve sensitivity to individual change, but it does not eliminate the need for an appropriate outcome label or athlete consent. Individualized models also need enough observations per athlete and safeguards against becoming an opaque personal profile.

7.6 Screening, prevalence, and the cost of errors

A screening system is useful only if it improves on existing pathways. The absence of a diagnostic cutoff for athlete burnout means a wearable model cannot inherit a clinical threshold from the ABQ without validation [1], [3]. If a study selects a threshold for convenience, the resulting categories should be described as analytic groups, not as confirmed cases. If the model is meant to detect change, it should be evaluated against repeated symptom measurements and functioning over time.

The base rate of the target matters. Even a model with apparently strong sensitivity and specificity may produce more false positives than true positives when the outcome is uncommon. This is especially important where a label can lead to unwanted attention or affect team status. Evaluation should therefore include positive and negative predictive values at the prevalence expected in the intended setting, as well as decision-curve or utility analyses where appropriate. Overall accuracy alone does not show that alerts help athletes.

The response to an alert also changes the acceptable error profile. A confidential, optional check-in has lower potential harm than a training restriction or a referral made without the athlete's agreement. The system should define the action before selecting a threshold. A model should not create an alert for which the institution has no appropriate support pathway. Conversely, a false reassurance can delay a conversation if staff treat a "low-risk" status as evidence that the athlete is well.

Research should report alert frequency per athlete-season, not only per observation. Repeated false alarms can create fatigue and undermine trust even when individual alerts appear statistically modest. The athlete should know that the output is uncertain and should be able to correct the context, decline discussion, or request help independent of the score.

8. Time, season, and dual-career context

8.1 Burnout unfolds across several calendars

University athletes move through academic semesters, examinations, training phases, competition schedules, travel, injury rehabilitation, and off-season periods. These schedules do not line up neatly. A wearable may register reduced sleep during exam week or higher load during competition, but the meaning depends on what else is happening. A model trained on one sport or season may not transfer to a different calendar or institution.

Longitudinal athlete-burnout research indicates that dimensions and correlates can change over time, and that measurement intervals matter [31]. Most available university student-athlete burnout evidence, however, remains cross-sectional [4]. A cross-sectional association between poor sleep and burnout cannot show whether sleep disruption preceded burnout, followed it, or arose from a common cause.

8.2 Repeated measurement does not automatically mean prediction

A dataset can contain many sensor records from a small number of people. Thousands of time points do not equal thousands of independent athletes. When observations are clustered within athlete, team, and institution, analysis must account for that structure. If a model predicts an outcome from the same athlete's past records, its performance should be evaluated on held-out future time periods, not only on randomly selected records.

The strongest prospective test would enroll athletes before a season, establish baseline symptom and sensor profiles, collect repeated measures through competition and exams, and assess burnout at planned intervals and after the season. The model should be evaluated at the athlete level and again in an independent institution. This design could test whether signals lead changes in symptoms, merely accompany them, or add no useful information.

8.3 Contextual information must be protected

Academic workload, injury, financial stress, scholarship status, coach communication, and social support can influence both physiological signals and burnout risk. Capturing more context may improve interpretation, but it also expands data sensitivity. A model that requires an athlete to disclose menstrual information, mental-health symptoms, location, or sleep habits must justify each data field and restrict use to the agreed purpose.

No single indicator should decide that an athlete is “resilient,” “unmotivated,” or “at risk.” Perfectionistic concerns, autonomy, social support, coaching climate, and injury history are established correlates in athlete burnout research, but those constructs need direct assessment and contextual interpretation [15], [20], [21]. Data should support care conversations rather than substitute for them.

8.4 Academic stress is not a nuisance variable

The student-athlete's academic calendar is part of the target population's lived context, not merely a confounder to remove. Examination weeks, laboratory placements, travel, and missed classes may affect sleep and perceived stress at the same time as training demands. If a wearable signal changes during exams, that pattern could reflect study pressure, sport pressure, or their interaction. A model that uses only training load may misattribute the source of strain.

At the same time, collecting academic activity data can deepen surveillance. A system might combine training records with attendance, location, screen use, or study schedules in the name

of personalization. Each additional data stream requires a new purpose justification and a new consent discussion. Research on student stress sensing does not automatically justify broad collection of educational data, and a student-athlete should not have to disclose every aspect of daily life to receive support.

A strong study can start with a minimal set of contextual variables chosen in advance: academic assessment periods, training phase, travel, injury, and athlete-reported stress. The goal is not to build a complete digital twin of the student. It is to identify the smallest amount of context needed to interpret a particular signal and test a defined hypothesis. Student-athlete burnout reviews identify academic and sport-specific stressors as relevant, while the wearable evidence remains focused mainly on physiological signals [4], [40].

Researchers should also consider whether the timing of assessment itself changes disclosure. Athletes may be more candid when surveys are administered by an independent research team than when a coach is present. They may underreport symptoms if they believe responses are visible to selection staff. Confidential survey administration, explicit separation from coaching decisions, and a clear statement of who receives results are therefore part of measurement validity, not only ethics.

9. From monitoring to surveillance

9.1 The same device can serve different power relationships

A wearable can be used as a private self-monitoring aid, a clinician's contextual tool, a coach-facing dashboard, or an institutional compliance system. The sensor does not decide which role it plays. Who owns the device, who receives the data, whether sharing is automatic, whether the athlete can refuse, and what consequences follow all shape the meaning of monitoring.

Research with collegiate athletes shows that data can move among athletes, coaches, teams, vendors, and social networks. In interviews with 15 elite U.S. collegiate athletes, participants described differences between data they chose to share and information automatically shared through platforms; some worried that refusing access could affect their standing with coaches [11]. A qualitative study with collegiate coaching staff found that staff collect and interpret many data types, while balancing performance goals, athlete care, privacy, and the risk of making athletes feel watched [48].

An ethnographic study of women's professional football monitoring found that team-wide sharing and decontextualized metrics could contribute to anxiety, demotivation, and body-

image concerns for some participants, while others found the data useful [49]. Those findings should not be generalized directly to university athletes because the league, sample, gendered context, and employment setting differ. They do show why outcomes and harms must be assessed rather than assumed.

9.2 Consent under institutional dependence

Informed consent is difficult when participation occurs inside an institution that controls training opportunities, playing time, scholarships, selection, or access to sport. The fact that an athlete accepts a device or checks a consent box does not establish that refusal is safe or practically available. Ethical discussion of collegiate biometric collection has long highlighted the power imbalance between student-athletes, universities, and commercial technology providers [10]. Recent work on athlete autonomy and proprietary data use raises related questions about access, ownership, and commercialization [50].

Consent should therefore be treated as an ongoing process. Athletes need to know what is collected, which staff can see it, how long it is retained, whether vendors receive it, and whether it can affect selection or eligibility. They should be able to correct inaccurate records, withdraw from optional monitoring, and receive an equivalent route to care or training support. Data collected for health support should not silently be repurposed for disciplinary or performance decisions.

9.3 Monitoring can affect the state it claims to measure

Monitoring is not always a neutral observation. An athlete who knows their sleep, location, or recovery score is visible may change behavior, worry about the number, or avoid reporting a poor state. A coach may interpret missing data as noncompliance. A low readiness score may itself increase anxiety or lead to reduced participation. In this way, measurement can become part of the environment that influences the outcome.

Work on invisible athlete monitoring argues that unobtrusive measurement has practical benefits but raises challenges around validity, reliability, burden, data governance, and communication [51]. Historical and contemporary sport-surveillance scholarship likewise treats monitoring as a social practice tied to discipline, commercial value, and institutional power—not merely a collection of sensors [52], [53], [54]. In university sport, this matters because monitoring designed as wellness support can be experienced as surveillance when athletes cannot control visibility or consequences.

9.4 The Quality in Sport context

A Quality in Sport article describes smart sports services in school physical education as a cycle of data collection, analysis, personalized plans, tracking, feedback, and evaluation [55]. It is useful as an adjacent account of how educational sport systems frame data-enabled personalization. Its subject is school PE and general student physical health, not intercollegiate athletes or burnout. It does not validate any wearable burnout metric, and it is used here only as contextual literature on educational sport technology.

9.5 Minimum governance conditions for a student-athlete setting

Data governance should be designed before a team selects a product or recruits athletes. A practical policy needs to identify the purpose of collection, the minimum data required, who can access raw and derived data, whether information is sent automatically to coaches, how long it is retained, and whether a vendor can reuse it. These conditions should be understandable to athletes, not only documented in procurement or research paperwork [9], [10].

Access should be role-based. A team coach may need a training-load summary for a specific practice decision, while a clinician may need sensitive health information. Those are not identical permissions. Automatic access to continuous sleep, location, or biometric data can exceed what an athlete expects from a training tool, as collegiate athletes' accounts of data flows illustrate [11]. Coaches also describe having to decide which data to share, how to contextualize them, and when information risks becoming intrusive or distressing [48].

A monitoring agreement should state that individual wearable scores will not independently determine selection, scholarship, discipline, or playing time. If a system is used for well-being support, a separate confidential channel should exist for raising concerns. Athletes should be able to view, correct, and, where appropriate, delete their data; to pause optional collection; and to ask what a model output means. The research literature on collegiate biometric data has repeatedly raised questions about consent, commercial interests, and institutional power [10], [56], [57], [58].

Governance should also cover group displays and data sharing among teammates. Even when no name is shown, team members may infer who has a low score from rank, schedule, injury status, or context. Public dashboards can turn a private health estimate into a social comparison. Athlete-centered monitoring studies in elite sport document the emotional effects of team-wide

metrics and incomplete interpretation [49]. Privacy-preserving design is therefore not merely encryption; it includes limits on visibility, purpose, and consequences.

Finally, data governance needs periodic review. Teams change staff, software vendors update models, and the original purpose of a device can expand. Re-consent should be required when recipients, data types, or uses change. Independent review is warranted when monitoring affects health decisions or team opportunities. The wider surveillance literature frames these systems as part of an expanding data economy, which makes purpose limitation and athlete representation important safeguards [9], [51], [54].

9.6 Potential benefits and potential harms are both empirical questions

Wearable monitoring may support self-awareness, help athletes notice patterns across training and sleep, and provide a prompt for a timely conversation. Those are plausible benefits, but they need measurement. A study should ask whether athletes understand the metric, whether they find it useful, whether they can choose what to share, and whether the information changes a decision in a way they regard as supportive. A device used privately for reflection may have a different effect from the same device connected to a coach dashboard.

Potential harms also need direct assessment. These can include worry about low scores, compulsive checking, social comparison, concealment, unwanted disclosure, loss of trust, or a sense that everyday behavior is being judged. Qualitative work with monitored athletes has documented both practical value and negative emotional responses, including anxiety and demotivation in contexts where data were shared or insufficiently contextualized [49]. Collegiate athlete interviews likewise show that data flows and automatic access may not match athlete expectations [11]. Neither finding proves that monitoring causes burnout, but both make it inappropriate to assume neutral effects.

The benefit-harm balance may differ across groups. Athletes returning from injury may value recovery feedback but also feel pressure to demonstrate readiness. Athletes in lower-status roles may be less able to refuse data sharing. Athletes with a history of disordered eating or anxiety may experience weight, calorie, or sleep metrics differently. Those possibilities should guide co-design and pre-specified harm measures, while avoiding unsupported claims that any subgroup is uniformly vulnerable.

A proportionate system collects the least data needed for the stated purpose, keeps individual information private by default, and gives athletes a clear way to pause or question monitoring.

If the benefit is only speculative and the data are sensitive, the burden of justification is higher. Monitoring should be discontinued or redesigned if its costs exceed the benefit or if the institution cannot protect the data as promised.

10. A defensible model of wearable-supported burnout care

The evidence supports a layered model rather than an automated diagnosis. First, the device estimates a signal. Second, a validated algorithm summarizes that signal with known uncertainty. Third, the estimate is compared with the athlete's own baseline and relevant context. Fourth, a voluntary check-in allows the athlete to explain the pattern. Fifth, a qualified practitioner may use a validated burnout measure or clinical assessment if concerns remain. Finally, any support decision is made with the athlete rather than imposed by a model.

Stage	Appropriate interpretation	Safeguard	Inappropriate leap
Sensor recording	Heart rate, movement, or sleep-related signal	State device and data limitations	"The device measures burnout"
Feature estimate	Change from personal baseline under specified conditions	Validate reliability and account for missing data	Treat one abnormal value as a symptom
Model output	Probability of a defined near-term state	Report calibration and external validation	Convert a stress score into a diagnosis
Athlete check-in	Context about training, study, sleep, health, and choice	Make disclosure confidential and voluntary	Treat silence as refusal or noncompliance
Professional assessment	Validated symptom measure and clinical judgment	Use trained staff and clear referral routes	Let model output decide playing time
Follow-up	Change in symptoms and functioning	Review benefit, burden, and harms	Continue collection without evidence of value

The model also clarifies the limits of "early detection." An algorithm might flag an unusual within-person pattern before an athlete reports elevated burnout. That could be useful if it leads to a supportive, confidential conversation. But a predictive alert should not be called early

detection unless the study shows that it identifies a defined outcome earlier than existing practice and improves the response. It must also show that false alerts do not create avoidable stigma or surveillance.

11. Recommendations for research and practice

11.1 Specify the decision the system is meant to support

Researchers should state whether the system aims to support self-reflection, guide training load, prompt a private well-being conversation, or screen for burnout symptoms. These purposes require different data and different thresholds. A system for training adjustment should not quietly become a mental-health screening tool. A system for research should not automatically feed individual data to coaches.

11.2 Measure burnout with validated athlete-specific tools

Studies should use a clearly described burnout measure, report each dimension separately, and state whether scores are continuous symptoms or categories. If an algorithm uses a threshold, the threshold should be justified, and sensitivity analyses should examine alternative choices. Because no universal diagnostic cutoff is established, authors should not imply that questionnaire-based categories equal clinical diagnosis [1], [3].

11.3 Collect repeated data over a full season

A cross-sectional sample cannot test early prediction. Future studies should span preparation, competition, examination periods, injury or return-to-play where relevant, and post-season recovery. They should include enough repeated burnout assessments to estimate change, not only a baseline and final questionnaire. Longitudinal studies should distinguish within-athlete change from differences between athletes [31].

11.4 Test incremental and external validity

Wearable features should be compared with a base model containing established predictors, such as prior burnout score, sport-specific stress, injury, and social support. Models should be validated on athletes not used in training and, where possible, on a different team, season, institution, and device. Researchers should report calibration, sensitivity, specificity, predictive values, uncertainty, and missingness, rather than accuracy alone. Participant-level and time-aware data splitting are essential when repeated records are available.

11.5 Treat athlete-reported experience as a co-equal channel

The data should not be described as more objective simply because they came from a sensor. Self-report captures aspects—such as meaning, autonomy, and sport devaluation—that a wearable cannot directly observe. Physiological and subjective channels can diverge, and that divergence may itself be informative [32], [36]. A defensible system pairs sensor features with athlete report and practitioner interpretation.

11.6 Predefine governance before recruitment

Protocols should specify who sees raw signals, derived scores, and burnout responses; where data are stored; how vendors can use them; how long they are retained; how athletes can withdraw; and whether data can influence selection. Governance should include an independent route for complaints and correction. The information-flow problems described by collegiate athletes and staff show that “the team has access” is not a sufficient data policy [11], [48].

11.7 Report harms and implementation outcomes

Studies should report perceived pressure, mistrust, unwanted disclosure, sleep or exercise preoccupation, data avoidance, and inequitable consequences. They should report who opted out or lacked a usable device, who received alerts, and what staff did in response. A model that improves prediction but damages trust may have poor real-world value. Monitoring interventions should be judged on athlete welfare and meaningful support, not just sensor completion or model performance.

11.8 Distinguish generative AI from predictive analytics

Generative AI may eventually translate sensor summaries into conversational feedback or help practitioners organize information. That use is different from a predictive model classifying physiological time series. Sport-psychology discussions of generative AI emphasize both potential applications and the need for AI literacy and ethical judgment [59]. At present, generative AI should not be presented as an established method for wearable-based burnout detection in student-athletes. If used, it should not invent clinical interpretations, create opaque athlete profiles, or issue high-stakes decisions. The sensor pipeline and the language-generation layer require separate validation and governance.

11.9 A proposed validation study

A feasible first study would recruit multiple university teams and follow athletes across one competitive season and an examination period. The primary outcome would be repeated ABQ dimensions, supplemented by a confidential clinical or sport-psychology assessment for athletes who report elevated symptoms. The study would collect only the sensor features needed for pre-specified hypotheses—such as resting HRV and sleep timing—along with training load and brief weekly reports of perceived stress, fatigue, and recovery.

The study should compare three models: a self-report and context model, a wearable-only model, and a combined model. A wearable-only model should not be treated as the default. The comparison tests whether sensor data add predictive value beyond information already available through athlete-centered assessment. Model training should use participant-aware splits, with a final test set from a separate team or institution. An independent replication should precede routine implementation.

The primary analysis should report whether wearable features predict later changes in each burnout dimension, with uncertainty and calibration. A separate feasibility analysis should measure device adherence, missingness, burden, acceptability, and whether athletes altered behavior because data were visible. The protocol should predefine how an alert triggers a confidential conversation and forbid using the output for playing-time, scholarship, or selection decisions. Athletes should help govern the data and review the communication language used for any alert.

Study component	Minimum design choice	Reason
Population	Multi-team university athletes with varied sports and genders	Reduces dependence on one team or sport
Duration	Full season plus academic examination period	Captures sport and study demands over time
Burnout outcome	Repeated ABQ dimensions, not only a composite score	Preserves multidimensionality
Sensor exposure	Pre-specified, minimally necessary signals	Limits exploratory overfitting and overcollection

Comparator	Context and self-report model without wearable features	Tests incremental value
Validation	Participant- and institution-aware held-out testing	Assesses generalization to new athletes
Governance	Athlete-controlled visibility and no punitive use	Preserves voluntariness and psychological safety
Outcomes	Calibration, utility, acceptability, and harms	Evaluates real-world value, not accuracy alone

11.10 Test usefulness, not prediction alone

A model can predict a score and still fail to improve athlete care. Before implementation, researchers should test whether an alert changes a meaningful decision, shortens the time to an appropriate support conversation, or improves an athlete-centered outcome. Prediction studies should therefore be followed by impact studies comparing usual confidential support with support informed by the model. The comparison should measure both benefit and burden: earlier help, distress from alerts, unwanted disclosure, time spent responding, and whether athletes trust the process.

Decision thresholds should reflect the consequences of errors. If an alert only prompts a private, optional check-in, a lower threshold may be acceptable. If it changes training, triggers referral, or becomes visible to selection staff, the potential harm of false positives and false negatives is much greater. Thresholds should be chosen with athletes and clinicians and tested prospectively. They should not be selected solely to maximize overall accuracy, especially where burnout symptoms are uncommon and a high accuracy score can coexist with many false alerts.

Implementation should also assess whether the system is sustainable outside a research team. Device replacement, data support, staff training, and confidential referral capacity all carry costs. If the team cannot respond safely to an alert, collecting more sensitive data may create an expectation of care without providing care. Feasibility and data governance are part of the intervention, not ancillary administrative details [48], [51].

A phased implementation is more defensible than immediate team-wide deployment. First, conduct co-design and usability work with athletes and staff. Second, validate the measure and model without allowing outputs to influence team decisions. Third, test a confidential support

pathway with independent oversight. Fourth, evaluate longer-term outcomes and adverse effects. Only after these stages should an institution consider routine use, and then only for the specific purpose supported by evidence.

11.11 A staged research and deployment agenda

A useful research program can move through four stages. The first is construct and co-design work: student-athletes, clinicians, coaches, and data specialists specify what “burnout detection” is supposed to mean and which decisions are in scope. The second is observational validation: sensor features are compared with repeated athlete-specific burnout measures across the season, with no high-stakes action based on the output. The third is external replication and subgroup evaluation in new sports and institutions. The fourth is an impact study that tests whether a defined support pathway improves outcomes compared with usual care.

This sequence avoids a common deployment problem: building a technically impressive model before deciding what the model should do. It also separates measurement validity from intervention efficacy. A model may estimate an outcome accurately yet fail to improve support; an intervention may help athletes even if the sensor adds no predictive value. Those are distinct results and should be reported separately.

11.12 Athlete participation in model design

Athletes should be involved before feature selection, not only asked to rate a finished dashboard. Co-design can identify which data feel useful, which outputs are stigmatizing, what contexts are missing, and which sharing relationships are acceptable. It can also surface whether “voluntary” monitoring is experienced as optional in a team environment. Research on collegiate athlete data practices shows that athletes and staff may understand information flows differently [11], [48].

Participation should be meaningful and compensated where appropriate. An athlete advisory group should include varied sports, gender identities, scholarship and non-scholarship athletes, injured and returning athletes, and students with different levels of technology access. It should have a route to challenge the purpose of the system and recommend discontinuation, rather than serving only as a user-testing panel.

11.13 Practical implications for athletic departments

Before purchasing a wearable system for mental-health or burnout purposes, an athletic department should ask whether the intended construct is measurable by that system, whether direct confidential assessment is more appropriate, and what staff will do when a concerning pattern appears. Procurement reviews should examine data export, vendor retention, model updates, user permissions, data deletion, and whether individualized data can be monetized or reused. A product's "wellness" label does not resolve these questions.

Departments should assign responsibility for interpretation to qualified personnel, provide staff training in measurement uncertainty, and separate health support from team selection. They should publish a short athlete-facing data policy and revisit it each season. If the department cannot offer a genuinely private and voluntary pathway, it should not promise that continuous monitoring is a mental-health service. These recommendations are consistent with athlete-centered approaches to invisible monitoring and data governance [10], [50], [51].

12. Implications for practitioners

For sport psychologists, athletic trainers, clinicians, and coaches, the current evidence supports a cautious operational rule: wearable data may help identify a reason to ask a question, but should not answer the question by itself. An unexpected HRV trend, sleep change, or sustained load pattern can be discussed alongside training records and the athlete's own account. It should not be treated as proof of burnout, lack of motivation, or unwillingness to compete.

Where a team uses wearables, athletes should receive a plain-language account of what is measured and what is not. They should know whether information is private to them, visible to medical staff, or shared with coaching personnel. The same device can support personal learning when under athlete control and create pressure when linked to selection or public comparison. Qualitative work in collegiate and elite sport documents these differences in experience [11], [49].

A practitioner should also resist treating a missing sensor record as a behavioral judgment. Non-wear may reflect discomfort, battery failure, privacy preference, skin irritation, schedule, or technical difficulty. A fair system offers an equivalent route to support that does not require continuous tracking. This is especially important when monitoring is introduced in a setting where athletes may feel that saying no is not a real option [10], [50].

Finally, labels should be used carefully. “Atypical sleep estimate,” “change from personal HRV baseline,” or “self-reported exhaustion increased” describes an observation. “Burnout risk” is an interpretation requiring validation. “Burned out” is a consequential label and should not be assigned by a consumer wearable or a black-box dashboard.

13. Limitations of this review

This critical narrative review is not a systematic review. Searches were iterative, not protocol-registered, and did not use independent duplicate screening or formal risk-of-bias assessment. The evidence streams differ in population, device, outcome, time scale, and study design; no quantitative pooling was appropriate. Some recent records were available only as abstracts or metadata, so claims about them are limited to their accessible descriptions.

The review is also shaped by the available literature. University student-athlete burnout research is more developed than direct sensor validation, while wearable-AI research often focuses on general student stress or athlete performance. Qualitative governance evidence is concentrated in specific North American collegiate or elite-sport contexts. Findings may differ across countries, sports, institutional systems, and levels of athlete choice. The evidence map should therefore be read as a critical account of the current research boundary, not an exhaustive census.

The proposed validation framework is a research agenda, not a validated intervention protocol. It does not establish that any specific wearable feature is a useful burnout predictor. Its purpose is to state the minimum distinctions future studies should preserve before making such a claim.

14. Conclusion

Wearables can measure or estimate signals that may be relevant to athlete strain: heart rate, HRV, movement, sleep, and training exposure. AI can combine those data and classify defined outcomes under tested conditions. Neither fact demonstrates that wearable analytics can currently detect burnout in university student-athletes.

The direct evidence remains divided. University student-athlete burnout studies mostly use self-report and often cross-sectional designs; wearable-AI work usually examines acute stress, fatigue, affect, or recovery in different populations. Candidate physiological markers remain exploratory, and objective sleep data do not necessarily mirror subjective burnout-related sleep complaints. At the same time, collegiate monitoring research shows that data access, consent, and power relations can shape whether a tool is experienced as support or surveillance.

The strongest present use of wearable data is therefore as a limited, contextual input to an athlete-centered conversation—not as a diagnosis, selection score, or substitute for validated burnout assessment. A credible path forward requires season-long studies, explicit burnout constructs, person-level and external validation, comparison against self-report and context, and governance that gives athletes meaningful control. Until those conditions are met, the most accurate conclusion is also the simplest: signals may alert us to ask, but they are not symptoms by themselves.

Appendix A. Core evidence sources and their role in the synthesis

Source	Evidence type	Main contribution	Boundary retained in this review
[4]	Scoping review of intercollegiate burnout	Maps populations, factors, and design limitations	No wearable or biomarker synthesis
[16]	NCAA Division I cross-sectional study	Applies the ABQ to current student-athletes	Two institutions; no causal or predictive inference
[31]	Scoping review of longitudinal athlete-burnout studies	Describes repeated-measure designs and changing dimensions	Most evidence remains questionnaire-based
[18]	Critical review of psychophysiological markers	Assesses cortisol and HRV as candidate indicators	Markers remain exploratory and heterogeneous
[32]	Young elite athlete sleep study	Shows subjective sleep complaints can diverge from objective sleep indices	Younger elite sample; not a wearable burnout model
[40]	Student-athlete sensing feasibility study	Demonstrates passive sensing of activity and stress-recovery context	Small preliminary sample; no validated burnout endpoint
[5]	Student smartwatch systematic review	Maps wearable AI and burnout-related symptoms in students	Not a student-athlete-specific evidence base

[38]	Student stress AI systematic review and meta-analysis	Highlights accuracy, ground-truth, and validation issues	Stress detection does not equal burnout detection
[6]	Athlete wearable-AI systematic review	Summarizes sensing of psychological states in sport	Focuses on transient states, not burnout
[11]	Collegiate-athlete qualitative study	Maps information flows and athlete privacy concerns	Does not estimate health or burnout effects
[48]	Collegiate coaching staff qualitative study	Describes data practices and care-control tensions	Staff perspectives do not substitute for athlete outcomes
[55]	Quality in Sport conceptual article	Provides adjacent context on smart sport services in school PE	Not direct evidence on university athletes or burnout

Appendix B. Minimum reporting checklist for wearable-burnout studies

A study claiming to detect or predict burnout in student-athletes should report, at minimum:

- The operational definition of burnout and the theoretical model used.
- The burnout instrument, subscales, recall period, scoring procedure, and any threshold rationale.
- Whether the target is a symptom score, a sample-defined category, a clinical assessment, or future change over time.
- The athlete population, institution type, sports, season phase, eligibility status, and relevant inclusion or exclusion criteria.
- Device manufacturer and model, sensor type, placement, firmware or algorithm version, sampling schedule, and preprocessing steps.
- Whether each wearable feature is a raw signal, device-derived estimate, proprietary score, or model-generated probability.
- The role of the device: measurement only, athlete feedback, coach feedback, clinical support, or team decision-making.
- The data-sharing pathway, storage location, retention period, vendor access, and who can view raw versus summarized data.

- Consent and refusal procedures, including whether athletes can opt out without losing training, selection, or support opportunities.
- Missingness, non-wear, device failure, and the handling of athletes who lack access or whose device is unsuitable.
- Confounders measured, including training phase, injury, illness, sleep opportunity, travel, medication where appropriate, and academic calendar.
- Model training and validation procedures, including participant-level separation, temporal validation, external validation, and calibration.
- Performance metrics beyond overall accuracy, with confidence intervals and subgroup results where sample sizes permit.
- The comparison model and whether wearable data add useful information beyond repeated self-report and contextual indicators.
- The action triggered by an alert, the person responsible, the athlete's right to contest it, and safeguards against punitive use.
- Benefits, harms, acceptability, trust, perceived pressure, and athlete-reported effects of being monitored.
- Whether the system was evaluated for decision usefulness and outcomes, not only for its ability to classify a dataset.

Without these details, readers cannot determine whether a model detects burnout, recognizes a correlated state, or learns an artifact of the sample or monitoring environment.

Appendix C. Questions for an athlete-centered monitoring review

Before a university team introduces wearable monitoring that is framed as a well-being or burnout tool, the following questions should be answered in writing:

1. What exact outcome is being monitored, and what evidence shows that the sensor or algorithm measures it?
2. Is the system intended for private self-reflection, clinical support, training adjustment, or performance management?
3. Can an athlete decline, pause, or remove the device without losing access to training, care, or team opportunities?
4. Who can see raw signals, derived scores, and questionnaire responses, and can access be changed by the athlete?

5. What happens after a high-risk alert, and who is responsible for a confidential follow-up?
6. Can the system distinguish a short-term response to training, exams, travel, illness, or injury from persistent burnout symptoms?
7. What evidence supports use in this sport, age group, institution, and device model?
8. How will the organization detect false alerts, missing data, bias, distress, and unwanted data sharing?
9. How can an athlete correct an inaccurate record or challenge an algorithmic interpretation?
10. When will collection stop if it does not improve athlete-centered outcomes?

A “no” or “not known” answer does not always prohibit research, but it should limit the claim and the deployment. The more consequential the decision, the stronger the evidence and governance requirements should be.

Supplementary Materials

Not applicable.

Author’s Contributions

Conceptualization, Justyna Anna Stanek; methodology, Justyna Anna Stanek; software, Tomasz Osenkowski and Klaudia Wojciechowska; check, Justyna Anna Stanek, Aleksandra Anna Perłowska, and Magdalena Maria Sukar; formal analysis, Justyna Anna Stanek; investigation, Katarzyna Marszałek, Katharina-Sophia Styś, and Justyna Agnieszka Półchłopek; resources, Justyna Anna Stanek; data curation, Natalia Barbara Radłowska and Karolina Weremczuk; writing - rough preparation, Justyna Anna Stanek, Aleksandra Anna Perłowska, and Magdalena Maria Sukar; writing - review and editing, Katarzyna Marszałek, Katharina-Sophia Styś, Justyna Agnieszka Półchłopek, Natalia Barbara Radłowska, and Karolina Weremczuk; visualization, Klaudia Wojciechowska and Tomasz Osenkowski; supervision, Justyna Anna Stanek; project administration, Justyna Anna Stanek. All authors have read and agreed with the published version of the manuscript.

Funding

The article did not receive any funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Acknowledgements

Not applicable.

Conflicts of Interest

The authors declare there are no conflicts of interest

References

- [1] T. Raedeke and A. L. Smith, “Development and Preliminary Validation of an Athlete Burnout Measure.” *Journal of sport & exercise psychology*, vol. 23 4, pp. 281–306, Dec. 2001, doi: [10.1123/jsep.23.4.281](https://doi.org/10.1123/jsep.23.4.281).
- [2] D. Gould and M. A. Whitley, “Sources and Consequences of Athletic Burnout Among College Athletes,” *Journal of Intercollegiate Sport*, vol. 2, pp. 16–30, Jun. 2009, doi: [10.1123/JIS.2.1.16](https://doi.org/10.1123/JIS.2.1.16).

- [3] H. Gustafsson, J. D. DeFreese, and D. Madigan, "Athlete burnout: review and recommendations." *Current opinion in psychology*, vol. 16, pp. 109–113, Aug. 2017, doi: [10.1016/j.copsyc.2017.05.002](https://doi.org/10.1016/j.copsyc.2017.05.002).
- [4] K. Waldhauser, S. Chen, S. Atkinson, and D. McEwan, "A Scoping Review of Athlete Burnout Among University Student-Athletes," *Journal of Intercollegiate Sport*, May 2026, doi: [10.17161/jis.v19i2.24228](https://doi.org/10.17161/jis.v19i2.24228).
- [5] P. Lialiou and I. Maglogiannis, "Students' Burnout Symptoms Detection Using Smartwatch Wearable Devices: A Systematic Literature Review," *AI Sensors*, May 2025, doi: [10.3390/aisens1010002](https://doi.org/10.3390/aisens1010002).
- [6] A. S. Mia and V. Soniawan, "AI-enabled wearable monitoring of athletes' psychological states: a systematic literature review," *Lentera Negeri*, Dec. 2025, doi: [10.29210/993220](https://doi.org/10.29210/993220).
- [7] A. S. Mia and B. Sujadesman, "Machine learning and multimodal AI for predicting athlete psychological states: a systematic review," *JRTI (Jurnal Riset Tindakan Indonesia)*, Dec. 2025, doi: [10.29210/30036927000](https://doi.org/10.29210/30036927000).
- [8] W. Alkasasbeh, A. Amawi, G. Grivas, B. E. Orhan, and T. Alawamleh, "Artificial intelligence and wearables in sport: performance, injury risk, and wellbeing," *Frontiers in Artificial Intelligence*, vol. 9, May 2026, doi: [10.3389/frai.2026.1838507](https://doi.org/10.3389/frai.2026.1838507).
- [9] A. S. Mia, I. Ifdil, and R. Rudyanto, "Digital surveillance and athlete privacy in elite sport: a prisma 2020-based systematic literature review," *Lentera Negeri*, Jul. 2026, doi: [10.29210/992690](https://doi.org/10.29210/992690).
- [10] J. F. Arnold and R. Sade, "Wearable Technologies in Collegiate Sports: The Ethics of Collecting Biometric Data from Student-Athletes," *The American journal of bioethics : AJOB*, vol. 17, pp. 67–70, Jan. 2017, doi: [10.1080/15265161.2016.1251648](https://doi.org/10.1080/15265161.2016.1251648).
- [11] K. Childs et al., Mapping the Flow of Information from Continuous Wearables: Insights from Elite Collegiate Athletes in the United States. 2025. doi: [10.1145/3749385.3749387](https://doi.org/10.1145/3749385.3749387).
- [12] T. Tanaka, "Trends and Prospects in Athlete Burnout Research : Aiming for Constructing a Theoretical Model," 2016. doi: [10.5997/SPOSUN.26.2_217](https://doi.org/10.5997/SPOSUN.26.2_217).
- [13] P. Piepiora, A. Rauk-Kubacka, and R. Kubacki, "Sport psychology in the physical culture sciences. A review," vol. 7, pp. 61–75, Feb. 2021, doi: [10.12775/PPS.2021.07.01.003](https://doi.org/10.12775/PPS.2021.07.01.003).
- [14] H. Gustafsson, G. Kenttä, and P. Hassmén, "Athlete burnout: an integrated model and future research directions," *International Review of Sport and Exercise Psychology*, vol. 4, pp. 24–3, Mar. 2011, doi: [10.1080/1750984X.2010.541927](https://doi.org/10.1080/1750984X.2010.541927).

- [15] S. Woods, S. Dunne, P. Gallagher, and A. McNicholl, "A systematic review of the factors associated with athlete burnout in team sports," *International Review of Sport and Exercise Psychology*, vol. 18, pp. 70–110, Nov. 2022, doi: [10.1080/1750984X.2022.2148225](https://doi.org/10.1080/1750984X.2022.2148225).
- [16] N. Giusti et al., "A Measure of Burnout in Current NCAA Student-Athletes," *Kansas Journal of Medicine*, vol. 15, pp. 325–330, Sep. 2022, doi: [10.17161/kjm.vol15.17784](https://doi.org/10.17161/kjm.vol15.17784).
- [17] K. Cureton, "Athlete Burnout: A Physiological Perspective," *Journal of Intercollegiate Sport*, vol. 2, pp. 31–34, Jun. 2009, doi: [10.1123/JIS.2.1.31](https://doi.org/10.1123/JIS.2.1.31).
- [18] L. J. Moore, S. Isoard-Gautheur, and H. Gustafsson, "Psychophysiological markers of athlete burnout: a call to arms," *International Journal of Sports Medicine*, Jan. 2024, doi: [10.1055/a-2433-3930](https://doi.org/10.1055/a-2433-3930).
- [19] H. L. Glandorf, D. Madigan, O. Kavanagh, and S. H. Mallinson-Howard, "Mental and physical health outcomes of burnout in athletes: a systematic review and meta-analysis," *International Review of Sport and Exercise Psychology*, vol. 18, pp. 372–416, Jun. 2023, doi: [10.1080/1750984X.2023.2225187](https://doi.org/10.1080/1750984X.2023.2225187).
- [20] C. E. Pacewicz, K. T. Mellano, and A. L. Smith, "A meta-analytic review of the relationship between social constructs and athlete burnout," *Psychology of Sport and Exercise*, Jul. 2019, doi: [10.1016/J.PSYCHSPORT.2019.02.002](https://doi.org/10.1016/J.PSYCHSPORT.2019.02.002).
- [21] D. Crowell and D. Madigan, "Perfectionistic concerns cognitions predict burnout in college athletes: a three-month longitudinal study," Jan. 07, 2021. doi: [10.1080/1612197X.2020.1869802](https://doi.org/10.1080/1612197X.2020.1869802).
- [22] J. Gotwals, "Perfectionism and Burnout Within Intercollegiate Sport: A Person-Oriented Approach," Dec. 01, 2011. doi: [10.1123/TSP.25.4.489](https://doi.org/10.1123/TSP.25.4.489).
- [23] L. F. Olsson, D. Madigan, A. Hill, and M. C. Grugan, "Do Athlete and Coach Performance Perfectionism Predict Athlete Burnout?" *European Journal of Sport Science*, vol. 22, pp. 1073–1084, Apr. 2021, doi: [10.1080/17461391.2021.1916080](https://doi.org/10.1080/17461391.2021.1916080).
- [24] H. Gray et al., "Examining the influence of grit on athlete burnout and well-being in female collegiate student-athletes," *Journal of athletic training*, Jun. 2023, doi: [10.4085/1062-6050-0121.23](https://doi.org/10.4085/1062-6050-0121.23).
- [25] H. Gray et al., "The Effects of Social Support on Athlete Burnout and Well-Being in Female Collegiate Athletes," *Journal of sport rehabilitation*, pp. 1–7, Aug. 2024, doi: [10.1123/jsr.2023-0421](https://doi.org/10.1123/jsr.2023-0421).

- [26] K. Kim, "Student-Athletes' dual career and sport burnout," Dec. 31, 2019. doi: [10.24985/kjss.2019.30.4.631](https://doi.org/10.24985/kjss.2019.30.4.631).
- [27] N. Dubuc-Charbonneau, N. Durand-Bush, and T. Forneris, "Exploring Levels of Student-Athlete Burnout at Two Canadian Universities." Aug. 31, 2014. doi: [10.47678/CJHE.V44I2.183864](https://doi.org/10.47678/CJHE.V44I2.183864).
- [28] J. W. P. Gaia, R. de C. Honorato, H. Coelho, J. F. Lima, and D. A. Pires, "SÍNDROME DE BURNOUT EM ATLETAS UNIVERSITÁRIOS," *Corpoconsciência*, Oct. 2023, doi: [10.51283/rc.27.e15063](https://doi.org/10.51283/rc.27.e15063).
- [29] R. S. Vealey, L. Armstrong, W. D. Comar, and C. Greenleaf, "Influence of perceived coaching behaviors on burnout and competitive anxiety in female college athletes," Sep. 01, 1998. doi: [10.1080/10413209808406395](https://doi.org/10.1080/10413209808406395).
- [30] B. Harris and M. L. Smith, "The influence of motivational climate and goal orientation on burnout: an exploratory analysis among Division I collegiate student-athletes." 2009.
- [31] B. E. Dišlere, K. Mārtinsone, and J. Kolesnikova, "A scoping review of longitudinal studies of athlete burnout," *Frontiers in Psychology*, vol. 16, Feb. 2025, doi: [10.3389/fpsyg.2025.1502174](https://doi.org/10.3389/fpsyg.2025.1502174).
- [32] M. Gerber et al., "Cross-Sectional and Longitudinal Associations Between Athlete Burnout, Insomnia, and Polysomnographic Indices in Young Elite Athletes." *Journal of sport & exercise psychology*, vol. 40 6, pp. 312–324, Dec. 2018, doi: [10.1123/jsep.2018-0083](https://doi.org/10.1123/jsep.2018-0083).
- [33] D. R. Seshadri et al., "Wearable sensors for monitoring the internal and external workload of the athlete," *NPJ Digital Medicine*, vol. 2, Jul. 2019, doi: [10.1038/s41746-019-0149-2](https://doi.org/10.1038/s41746-019-0149-2).
- [34] D. R. Seshadri et al., "Wearable Technology and Analytics as a Complementary Toolkit to Optimize Workload and to Reduce Injury Burden," *Frontiers in Sports and Active Living*, vol. 2, Jan. 2021, doi: [10.3389/fspor.2020.630576](https://doi.org/10.3389/fspor.2020.630576).
- [35] B. Sperlich, K. Aminian, P. Düking, and H. Holmberg, "Editorial: Wearable Sensor Technology for Monitoring Training Load and Health in the Athletic Population," *Frontiers in Physiology*, vol. 10, Jan. 2020, doi: [10.3389/fphys.2019.01520](https://doi.org/10.3389/fphys.2019.01520).
- [36] R. Rudyanto, M. Febrian, and K. Burhan, "Wearable and Portable Psychophysiological Monitoring of Stress, Recovery, and Affective States in Competitive Athletes: A Systematic Literature Review of Evidence Published from 2021 to 2026," *JRTI (Jurnal Riset Tindakan Indonesia)*, Aug. 2026, doi: [10.29210/30036922000](https://doi.org/10.29210/30036922000).

- [37] R. Cheng and J. H. M. Bergmann, "Impact and workload are dominating on-field data monitoring techniques to track health and well-being of team-sports athletes," *Physiological Measurement*, vol. 43, Mar. 2022, doi: [10.1088/1361-6579/ac59db](https://doi.org/10.1088/1361-6579/ac59db).
- [38] A. A. Abd-alrazaq et al., "The Performance of Wearable AI in Detecting Stress Among Students: Systematic Review and Meta-Analysis," *Journal of Medical Internet Research*, vol. 26, Sep. 2023, doi: [10.2196/52622](https://doi.org/10.2196/52622).
- [39] A. Sathyanarayana, O. B. Amin, J. An, and J.-P. Onnela, "Examining the Use of Consumer Wearable Devices and Digital Tools for Stress Measurement in College Students: Scoping Review of Methods," *JMIR mHealth and uHealth*, vol. 14, Mar. 2026, doi: [10.2196/64144](https://doi.org/10.2196/64144).
- [40] Z. Wang, A. Taya, T. Kato, K. Sezaki, and Y. Nishiyama, Toward Detecting Student-Athletes' Condition Using Passive Mobile and Wearable Sensing. 2024. doi: [10.1145/3675094.3677583](https://doi.org/10.1145/3675094.3677583).
- [41] R. A. Gani, E. Setiawan, I. Z. Achmad, R. Aminudin, T. Purbangkara, and M. Hofmeister, "Virtual reality-based tabata training: a professional method for changing levels physical fitness and psychological well-being on student-athletes," *Pedagogy of Physical Culture and Sports*, Apr. 2023, doi: [10.15561/26649837.2023.0201](https://doi.org/10.15561/26649837.2023.0201).
- [42] J.-H. Park et al., "Youth athletes and wearable technology," *F1000Research*, Nov. 2024, doi: [10.12688/f1000research.156207.1](https://doi.org/10.12688/f1000research.156207.1).
- [43] A. Sano et al., "Identifying Objective Physiological Markers and Modifiable Behaviors for Self-Reported Stress and Mental Health Status Using Wearable Sensors and Mobile Phones: Observational Study," *Journal of Medical Internet Research*, vol. 20, Jun. 2018, doi: [10.2196/jmir.9410](https://doi.org/10.2196/jmir.9410).
- [44] B. Hickey et al., "Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions and Stress: A Systematic Review," *Sensors (Basel, Switzerland)*, vol. 21, May 2021, doi: [10.3390/s21103461](https://doi.org/10.3390/s21103461).
- [45] G. Vos, K. Trinh, Z. Sarnyai, and M. Azghadi, "Generalizable machine learning for stress monitoring from wearable devices: A systematic literature review," *International journal of medical informatics*, vol. 173, pp. 105026, Sep. 2022, doi: [10.1016/j.ijmedinf.2023.105026](https://doi.org/10.1016/j.ijmedinf.2023.105026).

- [46] R. Tutunji et al., “Using wearable biosensors and ecological momentary assessments for the detection of prolonged stress in real life,” Jun. 30, 2021. doi: [10.1101/2021.06.29.450360](https://doi.org/10.1101/2021.06.29.450360).
- [47] R. Tutunji et al., “Detecting Prolonged Stress in Real Life Using Wearable Biosensors and Ecological Momentary Assessments: Naturalistic Experimental Study,” *Journal of Medical Internet Research*, vol. 25, Oct. 2023, doi: [10.2196/39995](https://doi.org/10.2196/39995).
- [48] M. Brewer et al., *Coach, Data Analyst, and Protector: Exploring Data Practices of Collegiate Coaching Staff*. 2025. doi: [10.1145/3706598.3714026](https://doi.org/10.1145/3706598.3714026).
- [49] P. Howell, E. Sherry, E. Pechenkina, and P. Scifleet, “‘It’s almost like a time trial day every day’: AFLW footballers’ experience of club-led digitised performance monitoring,” *Sport Management Review*, vol. 29, pp. 450–474, Feb. 2026, doi: [10.1080/14413523.2026.2630468](https://doi.org/10.1080/14413523.2026.2630468).
- [50] C. Corr, S. Stokowski, E. Staurowsky, A. E. Chisholm, and A. Paule-Koba, “Athlete Autonomy and Wearable Technology: Collegiate Athletes’ Understanding and Perceptions of the Proprietary Use of Their Physical Performance Data,” *Sociology of Sport Journal*, 2026, doi: [10.1123/ssj.2025-0026](https://doi.org/10.1123/ssj.2025-0026).
- [51] C. Leduc and D. Weaving, “Invisible Monitoring for Athlete Health and Performance: A Call for a Better Conceptualization and Practical Recommendations,” *International journal of sports physiology and performance*, pp. 1–5, Jan. 2025, doi: [10.1123/ijsp.2024-0292](https://doi.org/10.1123/ijsp.2024-0292).
- [52] I. Warren and N. Zurawski, “Editorial Surveillance and Sport,” 2014.
- [53] L. Jones, P. Marshall, and J. Denison, “Health and well-being implications surrounding the use of wearable GPS devices in professional rugby league: A Foucauldian disciplinary analysis of the normalised use of a common surveillance aid,” Dec. 01, 2016. doi: [10.1016/J.PEH.2016.09.001](https://doi.org/10.1016/J.PEH.2016.09.001).
- [54] H. Bowles and D. McGee, “Sport, surveillance and the data economy: an expanding horizon for research and governance,” *International Journal of Sport Policy and Politics*, vol. 16, pp. 339–347, Mar. 2024, doi: [10.1080/19406940.2024.2323015](https://doi.org/10.1080/19406940.2024.2323015).
- [55] Y. Liu, M. Zhang, and B. Ji, “Analysis of the Mechanism of Smart Sports Services in School Physical Education and Student Physical Health,” *Quality in Sport*, Apr. 2025, doi: [10.12775/qs.2025.40.59422](https://doi.org/10.12775/qs.2025.40.59422).

- [56] J. Reifenberg, “Ethical analysis of biometric data collection in collegiate athletics,” Aug. 15, 2020.
- [57] C. Rush and B. Osborne, “Benefits and Concerns Abound, Regulations Lack in Collegiate Athlete Biometric Data Collection,” *Journal of Legal Aspects of Sport*, Feb. 2022, doi: [10.18060/25479](https://doi.org/10.18060/25479).
- [58] C. Corr, S. Stokowski, E. Staurowsky, A. Paule-Koba, and R. Southall, “The Modernization of Monitorization: Technological Adaption and Innovation as Mechanisms for College Athlete Surveillance,” *Journal of Sport and Social Issues*, vol. 49, pp. 175–195, Apr. 2025, doi: [10.1177/01937235251330206](https://doi.org/10.1177/01937235251330206).
- [59] A. Oliver and J. Guiller, “Generative AI in sport and exercise psychology: Exploring opportunities and overcoming challenges,” *Sport & Exercise Psychology Review*, Dec. 2024, doi: [10.53841/bpssepr.2024.19.2.36](https://doi.org/10.53841/bpssepr.2024.19.2.36).