SSE #273

Sports Science Exchange (2026) Vol. 41, No. 273, 1- 6

GATORADE SPORTS SCIENCE INSTITUTE

BEYOND ALGORITHMS: THE DATA ECOSYSTEM BEHIND AI-DRIVEN PERSONALIZED SPORTS PERFORMANCE

Eric C. Freese, MS, PhD[1], Corey T. Ungaro, PhD[2], Liam R. Brown, MSc[3]

[1] Gatorade Sports Science Institute, PepsiCo Life Sciences, Global R&D, Frisco, TX, USA
[2] Gatorade Sports Science Institute, PepsiCo Life Sciences, Global R&D, Chicago, IL, USA
[3] Gatorade Sports Science Institute, PepsiCo Life Sciences, Global R&D, UK

KEY POINTS

- AI can support individualized fueling, hydration, and recovery guidance, but personalization is limited by the completeness, quality, and integration of athlete data.
- Inter-individual variability (e.g., sweat losses, carbohydrate oxidation, gastrointestinal tolerance, and recovery responses) makes one-size-fits-all recommendations insufficient for many athletes.
- Most near-term value comes from AI that integrates context (training load, environment, behavior) and assists practitioner workflows, rather than fully automated decision-making.
- Key risks include hallucinations and overconfidence, non-representative training data, noisy or missing inputs, and automation bias; guardrails and validation are required.
- Organizations should prioritize data standards, interoperability, governance, privacy/consent, and human-in-the-loop oversight to deploy AI personalization safely and effectively.

INTRODUCTION

Recent advances in artificial intelligence (AI) are rapidly changing how individuals' access and interact with health and performance information. Large language models (LLM) and AI-enabled digital platforms are increasingly used to answer questions related to training, nutrition, hydration, and recovery. At the same time, AI-driven coaching tools are emerging across healthcare and wellness settings, where early research suggests that these systems can influence behaviors such as physical activity, weight management, and adherence to lifestyle interventions. Together, these developments suggest that AI may play an expanding role in delivering health and performance guidance at scale.

Parallel to the growth of AI tools has been a dramatic expansion in the availability of personal health and performance data. Wearable technologies, athlete monitoring systems, and digital health platforms now generate continuous streams of physiological and behavioral information, including metrics related to heart rate (HR), sleep, activity, and recovery. These technologies are widely used by both elite athletes and recreationally active individuals, creating new opportunities to better understand individual responses to training, nutrition, and environmental stressors.

The growing availability of these data streams is particularly relevant in sports nutrition and exercise physiology, where substantial inter-individual variability exists in factors such as sweat rate and electrolyte losses, carbohydrate oxidation capacity, gastrointestinal tolerance to fueling strategies, and recovery responses to training. As a result, population-based guidelines may not always reflect the needs of individual athletes or physically active individuals.

AI offers the potential to integrate diverse data sources, including physiology, nutrition intake, training load, and environmental conditions, to support more personalized performance guidance. However, the effectiveness of these approaches will depend not only on advances in algorithm development but also on the availability and integration of high-quality physiological and behavioral data. In this context, the future of AI-enabled personalized sports performance may be determined less by the sophistication of algorithms and more by the development of robust athlete data ecosystems capable of supporting individualized decision-making.

This Sports Science Exchange (SSE) article examines the potential of AI to advance personalized sports nutrition and performance guidance, but meaningful progress will depend less on algorithm sophistication and more on the development of integrated data ecosystems capable of centralizing physiological, behavioral, and environmental determinants of performance.

WHY PERSONALIZATION MATTERS: INTER-INDIVIDUAL VARIABILITY

Even when individuals are matched for sex, age, body mass, or training status, their responses to the same training load, nutritional strategy, or recovery modalities can differ meaningfully. Substantial inter-individual variability in physiology and behavior means that population-based guidelines may not suffice. Personalization aims to refine guidance using individual-level data and observed responses over time (Figure 1).

For years, population-level recommendations have provided valuable starting points, but they fail to capture how individuals actually respond over time. Traditional sports science studies have relied heavily on group averages and between-person comparisons. However, with the

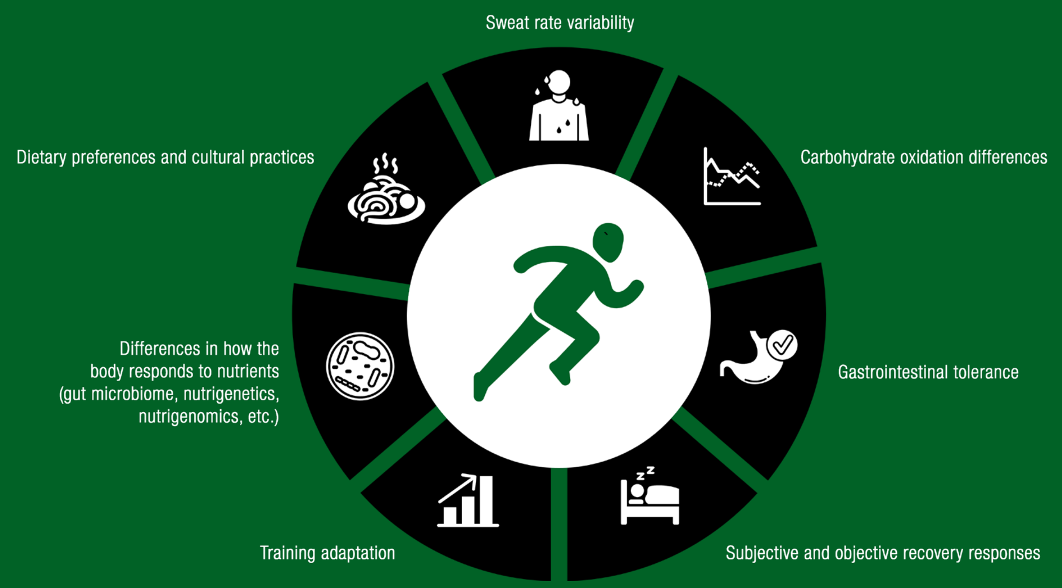

Figure 1. Common sources of variability in physiology and behaviors that can be used to help personalize sports nutrition and performance support.

recent democratization of validated digital capabilities, researchers and individuals can understand and react to their own within-person patterns. For example, individuals exhibit wide interindividual variability in sweat rate (0.5–2.0 $L{\cdot}h^{-1}$) and sweat sodium concentration (10–90 $mmol{\cdot}L^{-1}$) (Baker, 2017), and protein requirements scale with body mass, such that a 45 kg female will have different needs than a 75 kg female. Moreover, individuals exhibit substantial variability in carbohydrate oxidation rates, including the capacity to modify these rates through diet and exercise intensity (Burke, 2021).

THE ENABLER: EXPLOSION OF ATHLETE DATA

Continuous athlete monitoring is making personalization more feasible by generating data streams that can be linked to training, nutrition, recovery, and performance outcomes, from a range of different sources (Figure 2). A recent publication focused on AI and wearable biosensors concluded that machine learning (ML) models trained on wearable data can support training adaptation and recovery estimation, with improved performance over traditional workload metrics in endurance, strength, and team-sport contexts when evaluated using athlete-wise or longitudinal validation schemes (Madrigal-Cerezo et al., 2026).

Furthermore, a scoping review of AI in biomechanics demonstrated that learning management systems enhanced knowledge transfer, raising coaches' understanding by 45%, and athlete adherence by 3.4 times (Souaifi et al., 2025). Additionally, implementing integrated AI systems resulted in a 23% reduction in reinjury rates (Souaifi et al., 2025).

Investigators evaluating AI in endurance sports showed that AI systems effectively integrated multimodal physiological, environmental, and behavioral data to enhance metabolic health monitoring, predict recovery states, and personalize nutrition (Grivas & Safari, 2025). Continuous glucose monitoring combined with AI algorithms allowed precise carbohydrate management during prolonged events, improving performance outcomes. AI-driven supplementation strategies, informed by genetic polymorphisms and individual metabolic responses, demonstrated enhanced ergogenic effectiveness (Grivas & Safari, 2025).

Additionally, blood biomarkers and AI integration are rapidly evolving, providing objective measures for interdisciplinary teams to support athletes' health, nutrition and performance across a broad spectrum of physiological systems. The success of a blood-biomarker monitoring program is dependent not only on the selection of appropriate biomarkers, but also upon the timing of the testing (Pedlar et al., 2019).

Researchers developed a model for predicting over training syndrome while examining 120 youth (12-18 y) male soccer players from six elite South Korean soccer academies (Qian & Lee, 2025). Biweekly blood sampling of testosterone, cortisol, creatine kinase, interleukin-6, and tumor necrosis factor-α levels, as well as weekly psychological assessments (Recovery-Stress Questionnaire for Athletes (RESTQ-Sport), Profile of Mood States), continuous GPS training load monitoring, and monthly performance tests were conducted. Machine learning modeling identified testosterone to cortisol ratio (0.89), RESTQ Sport balance (0.83), and acute to chronic workload ratio (0.78) as the strongest predictors, enabling a three tier risk stratification system that correctly identified 85% of high-risk overtraining syndrome cases one week before performance decline (Qian & Lee, 2025).

All authors cite limitations and future need to be focused on external validations, diverse populations, sport types and various athletic levels.

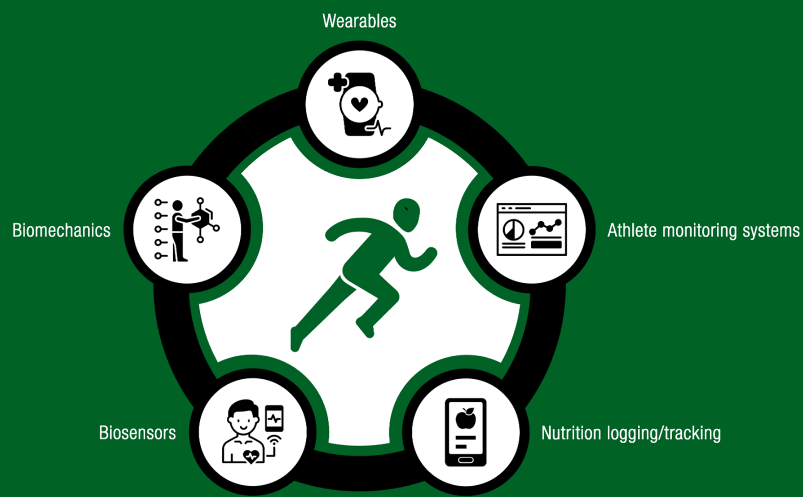

Figure 2. Examples of key data sources to support AI-driven personalized sport nutrition and performance.

WHERE AI FITS IN SPORTS SCIENCE (AND WHAT IT CAN AND CANNOT DO)

In applied sport settings, AI is most useful as a decision-support and integration tool by combining multiple inputs, amplifying signals, summarizing patterns, and helping practitioners and athletes act on information more efficiently. However, outputs are only as reliable as the data provided and the validation of the underlying models and tools.

AI best fits with pattern recognition for training load management, menstrual cycle tracking, performance analysis, sleep tracking, off-training behavior, injury prevention, and return-to-play. AI can integrate multimodal datasets (e.g., physiology, nutrition, sleep) and it is able to identify nonlinear relationships (e.g., heart rate variability (HRV) x sleep x training). AI can be an excellent fit for within-person patterns which are critical for n-of-1 personalization.

AI must be context-specific and need the interpretation of coaches, trainers, or practitioners. AI models are limited by the unpredictability of sport (Mateus et al., 2024) and the lack of scale across sports, real-world scenarios, and populations. Moreover, there is a need to provide interpretable outputs for coaches, trainers, and practitioners: AI-powered, human-led (Kumar et al., 2025).

AI-enabled tools may integrate signals across:

- Physiology
- Nutrition
- Behavior
- Environment

AI Use Cases:

- **Generative AI (LLMs)**: For synthesis/communication/coaching interactions.
- **Predictive/ML models**: For forecasting responses/risk/readiness.

CORE ARGUMENT: PERSONALIZED PERFORMANCE IS A DATA ECOSYSTEM PROBLEM

Even highly capable algorithms cannot deliver meaningful personalization if key inputs are missing, inaccurate, or disconnected across systems. Robust data ecosystems that link physiology, training load, nutrition, environment, behavior, context, and outcomes are necessary to make AI outputs interpretable and actionable.

Common ecosystem gaps include:

- Physiological measurement.
- Nutrition and dietary intake and preference data.
- Environmental context.
- Logistical context (Stage of macro-, meso- and microcycles, phase of competition, travel requirements).
- Individual context/subjective responses (Injury, illness, psychological feelings).
- Fragmented pairable data (Wearables, training and nutrition logs, surveys).
- Individual behavior (Capability, opportunity, and motivation).

EMERGING APPLICATIONS FOR AI-ENABLED PERSONALIZATION

AI is emerging as a key enabler of personalized sports nutrition, hydration, recovery and performance support. When supported by a high-quality data ecosystem, AI may help democratize the translation of complex data and identification of patterns, into practical insights that can improve provision of precision-based interventions (Table 1) (Grivas & Safari, 2025; Richter & King, 2025).

POTENTIAL BENEFITS

When supported by robust and integrated data ecosystems, paired with sufficient human interaction, AI-enabled systems may provide several practical benefits for athletes and practitioners (Table 2) (Hliš et al., 2024; Mancin, et al., 2020; Naughton et al., 2024; Papastratis et al., 2024; Richter & King, 2025).

Potential Applications	Examples of Data Enablers
Hydration personalization	• Sweat rate and composition • Environmental conditions • Fluid intake habits
Fueling strategies	• Exercise demands • Nutritional intake • Carbohydrate oxidation rates • Gastrointestinal tolerance
Recovery strategies	• Training load • Sleep history • Diet history • Stress measures
Practical food-based advice (meal plans, recipes)	• Training load • Recovery demands • Physiological measurements • Dietary preferences and behaviors
Practitioner-facing dashboards	• Training schedules and load • Nutrition and hydration plans and intake • Recovery schedules and scores
Athlete feedback and nudges	• Training and recovery schedules • Nutrition and hydration plans • Behavioral interventions
Training optimization	• Readiness and recovery scores • Adaptation goals and measures
Injury prediction and management	• Longitudinal patterns in injury risk • Diet history • Training load • Recovery practices
Digital twins (virtual representation of individual to stimulate, monitor and predict performance and health outcomes)	• Physiological data • Behavioral data • Contextual data

Table 1. Potential applications of AI in sport nutrition and performance.

PITFALLS, LIMITATIONS, AND IMPLEMENTATION BARRIERS

Each of these applications will be most impactful where AI helps augment practitioner workflows, rather than replacing decision making entirely. Fully autonomous workflows may provide risk as a result of limitations in data quality and completeness, AI processes, validation, contextual understanding, and interpretability. Despite advances in validated algorithms, such data limitations may result in the provision of inadequate, misleading, or potentially dangerous personalized recommendations. To deploy AI responsibly in sport settings, practitioners should assume that errors will occur and build systems that can limit, detect, and learn from them, whilst prioritizing data standards.

Potential data limitations include:

- **Incomplete data**: Missing nutritional, physiological, behavioral, environmental, and/or contextual data.
- **Poor data quality**: Inaccurate self-reported data, incorrect/inconsistent wearable use, poor controls.
- **Non-representative data**: Data that doesn't reflect target population, sport, sex, or environment.
- **Noisy inputs**: Measurement error, device variability, inconsistent data sets, and signal artifacts that may obscure true physiological patterns.
- **Limited longitudinal monitoring**: Short or fragmented datasets may not provide sufficient representations of an individual's norm.

Potential Benefits	Implications
Democratization of personalized advice	Integration of large, complex data sets, and automation of data processing may allow recommendations to be tailored to individual, context-specific needs, which otherwise would require technical knowledge and significant time.
Early identification of meaningful changes	AI systems can monitor and evaluate athlete's data over time, highlighting changes to their norm. In turn, informing the practitioner of 'abnormal' responses outside of that norm, when intervention may be required.
Scalable support and improved efficiency	Ability to reduce manual work (e.g. data entry) and scale personalization may enable practitioners to support a larger number of athletes without significantly increasing workload. This may allow practitioners to focus expertise where it adds greatest value (e.g. building rapport or resolving barriers to implementation).
Scenario modelling	Emerging tools (e.g. digital twins, or tools that can generate synthetic data) may allow practitioners to model different training, nutrition or recovery scenarios and predict the results they will have on health and performance.
Translation into practice	Converting technical, scientific recommendations (e.g. nutrient targets) into practical outputs such as food/exercise/recovery-based guidelines, plans, or schedules enables democratization of personalized support.
Pattern recognition and problem solving	The ability to analyze complex, longitudinal datasets and recognize patterns may be useful in solving particular problems, e.g., why an athlete may be more susceptible to injury at a specific time point in the macrocycle.

Table 2. Potential benefits of linking AI with high-quality data ecosystems.

Even when the data ecosystem contains a relatively complete, quality dataset, AI systems themselves can have limitations, and come with several implementation challenges (Table 3) (Grivas & Safari, 2025; Naughton et al., 2024; Richter & King, 2025). Without acknowledging these limitations and thus incorporating human-in-the-loop oversight within the system to help limit and detect, and correct potential issues, there is a risk of generating incorrect, overconfident, or poorly contextualized personalized recommendations which may impair both health and performance.

Potential AI limitations include:

- **Hallucinations**: Incorrect outputs in generative systems.
- **Overconfidence**: Models/tools that are not valid for target populations (sport, sex, environment and use case) may still express high level of certainty that is not justified.
- **Bias against non-representative athletic populations**: Under-represented groups, para-sports, female athletes, youth, regions for heat-acclimation may be poorly served by existing models.
- **Inconsistency in responses**: Similar data input may not always yield identical recommendations from Generative AI due to inconsistency of AI systems.

Challenge	Potential Solution(s)
Lack of validation and transparency of predictive models	• Models should be validated against gold-standard, real-world applications. • An overview (minimum) of predictive model algorithms should be provided.
Adapting to changes in patterns over time	• AI models will learn new patterns over time and should be adapted respectively to remain effective and reliable.
Limited practitioner oversight	• Ensure systems provide relevant context alongside outputs. • Incorporate stages of expert oversight to reduce risk of acting on inappropriate or unsafe outputs.
Risks of applying systems outside of intended context and unpredictability of sport	• Ensure systems provide relevant context alongside outputs, only applying models to validated sports, sexes, climates, or performance levels. • Incorporate stages of expert oversight to reduce risk of acting on inappropriate or unsafe outputs. • Ensure systems can provide audit trails.
Transparency, explainability and interpretability	• Each AI model should be transparent to its functioning and development, with humans able to understand and interpret how decisions are being made (including specific context), and subsequent results/outputs.
Regulatory landscape	• Athlete data should be managed in line with local privacy and security regulations, which should be monitored and adapted as required regularly.

Table 3. Potential implementation challenges and solutions for the use of AI in sport nutrition and performance.

SAFETY AND RISK MANAGEMENT

Given the potential for error, AI-enabled systems should be implemented with clear safeguards to protect athlete health and performance and practitioner integrity. These safeguards (Table 4) should be built into the system from the start, to enable safe and relevant personalization (Grivas & Safari, 2025; Naughton et al., 2024).

Key risks may include:

- Inappropriate nutrition or hydration guidance.
- Misinterpretation of readiness or recovery metrics.
- Over-reliance on automated recommendations (Automation bias).
- Overstepping between wellness tools and medical decision-making.
- Inappropriate handling of athlete data: Privacy, consent and security.
- Removal of human interaction: Risk of poorly contextualized recommendations, missed errors, and loss of jobs due to automation.
- Failure to detect red-flag conditions requiring escalation: Without sufficient context and human-in-the-loop oversight, AI systems may fail to recognize situations that require practitioner review or referral (e.g. hypo- or hypernatremia, rapid weight loss, disordered eating risk, abnormal HR/HRV patterns).
- Inequality (digital divide): Amplification of pre-existing issues and widening of equality between organizations or individuals with differing financial power.

FUTURE RESEARCH PRIORITIES

To ensure safe, practical and valid AI-enabled personalization in sport, several key areas require further research:

- Establishing safety, ownership, transparency, and validation standards.
- Developing integrated athlete data ecosystems.
- Optimizing prompting and human–AI interaction.
- Equity and representation of diverse populations.
- Evaluating the efficacy of AI-led diagnostics and coaching.
- Integrating behavioral change frameworks.

Risk Mitigation Strategies	
Data governance	Ensure that data collection, storage and retention and management processes are clear, appropriate, time-bound and supported by ongoing consent.
Human-in-the-loop oversight	Practitioners retain responsibility for decision making. Where appropriate, involve human review throughout the process.
Escalation triggers	Pre-defined thresholds for further human oversight.
Contraindications	Clear boundaries for when a particular recommendation should/should not be made.
Clear boundaries between wellness and medical decisions	Ensure any potential medical issues are flagged and recommendations are not provided without oversight of medical professional.
Diversity and equality	Include diverse user groups in design, testing and implementation. Pre-specifying model purpose and target population.
Audit trains	Allow tracking of data storage, processes and review of outputs, with reasoning behind each step.
Clear governance	Define roles for data ownership, model validation and data review.

Table 4. Potential risk mitigation strategies.

PRACTICAL APPLICATIONS / RECOMMENDATIONS

- **Start with the decision, then the data**: Define the applied question (e.g., hydration plan for heat, race-day fueling, recovery nutrition) and identify the minimum viable inputs and outcomes needed to evaluate success.
- **Set minimum data standards and quality control**: Establish acceptable ranges, missing-data rules, and device/logging protocols for key inputs (training load, sweat losses, nutrition intake, sleep/recovery, environment).
- **Prioritize interoperability and governance**: Standardize athlete identifiers, data schemas, and version control; define who owns data, who can access it, and who is accountable for model updates and outputs.
- **Validate on the target context**: Require evidence that tools perform appropriately for the specific sport, sex, environment, competitive level, and use case; re-validate when inputs, populations, or vendors change.
- **Keep a human in the loop**: Use AI for decision support, not autonomous prescription; define sign-off roles (e.g., sports RD/physiologist/medical) and document review workflows.
- **Build red-flag rules and escalation pathways**: Specify thresholds that trigger review or referral (e.g., rapid weight change, suspected hypo-/hypernatremia risk, disordered eating risk indicators, abnormal HR/HRV patterns).
- **Address representativeness and equity**: Assess model performance across under-represented groups (female athletes, youth, para-sport, varied climates/heat acclimation) and avoid extrapolating beyond available data.
- **Protect privacy and consent**: Implement informed consent, data minimization, retention limits, and clear policies for secondary use; ensure athletes can understand and access their own data.
- **Integrate behavior change**: Pair personalization with adherence supports (education, friction reduction, nudges, feedback cadence) so recommendations translate into consistent action.
- **Monitor for automation bias**: Train staff to challenge outputs, track errors/near-misses, and maintain audit trails so systems improve over time.

CONCLUSION

AI is poised to influence how athletes and practitioners access and apply sports nutrition and performance guidance, but the limiting factor is often not the sophistication of algorithms; it is the maturity of the underlying data ecosystem. Personalization requires high-quality, integrated, context-rich inputs and clear outcome measures, supported by governance, validation, and practitioner oversight. Organizations that invest in interoperability, data standards, and safety guardrails will be best positioned to use AI to scale individualized fueling, hydration, training, and recovery support while minimizing risk.

REFERENCES

Baker L. B. (2017). Sweating Rate and Sweat Sodium Concentration in Athletes: A Review of Methodology and Intra/Interindividual Variability. Sports medicine (Auckland, N.Z.), 47(Suppl 1), 111–128.

Burke L. M. (2021). Ketogenic low-CHO, high-fat diet: the future of elite endurance sport?. The Journal of physiology, 599(3), 819–843.

Grivas, G. V., & Safari, K. (2025). Artificial intelligence in endurance sports: Metabolic, recovery, and nutritional perspectives. Nutrients, 17, 3209.

Hliš, T., Fister, I., & Fister Jr, I. (2024). Digital twins in sport: Concepts, taxonomies, challenges and practical potentials. Expert Syst. Appl. 258.

Kumar, R., Bogia, P., Haq, A.U., Singh, V., Onima Reddy, T. (2025) Leveraging artificial intelligence and machine learning in sport sciences: a systematic literature review of applications, outcomes, and future directions. Sport Sci Health, 21, 2429–2443.

Madrigal-Cerezo, R., Domínguez-Sanz, N., & Martín-Rodríguez, A. (2026). Wearable biosensing and machine learning for data-driven training and coaching support. Biosensors, 16, 97.

Mancin, L., Rollo, I., Mota, J. F., Piccini, F., Carletti, M., Susto, G. A., Valle, G., & Paoli, A. (2021). Exerc Sport Sci Rev. 49(1), 42-49.

Mateus, N., Abade, E., Coutinho, D., Gómez, M.-Á., Lago Peñas, C., & Sampaio, J. (2025). Empowering the sports scientist with artificial intelligence in training, performance, and health management. Sensors, 25, 139.

Naughton, M., Salmon, P. M., Compton, H. R., & McLean, S. (2024). Challenges and opportunities of artificial intelligence implementation within sports science and sports medicine teams. Front. Sports. Act. Living. 6.

Papastratis, I., Konstantinidis, D., Daras, P., & Dimitropoulos, K. (2024). AI nutrition recommendation using a deep generative model and ChatGPT. Sci. Rep. 14(1).

Pedlar, C. R., Newell, J., & Lewis, N. A. (2019). Blood Biomarker Profiling and Monitoring for High-Performance Physiology and Nutrition: Current Perspectives, Limitations and Recommendations. Sports medicine (Auckland, N.Z.), 49(Suppl 2), 185–198.

Qian, H., & Lee, S. (2025). A multidimensional prediction model for overtraining risk in youth soccer players: Integrating physiological and psychological markers. Journal of Sports Sciences, 43(17), 1819–1834.

Richter, C., & King, E. (2025). Opportunities and Challenges of Artificial Intelligence in Sports Science. In D. Memmert, Artificial Intelligence and Machine Learning in Sports Science (pp. 307-324). Köln: Springer.

Souaifi, M., Dhahbi, W., Jebabli, N., Ceylan, H. I., Boujabli, M., Muntean, R. I., & Dergaa, I. (2025). Artificial intelligence in sports biomechanics: A scoping review on wearable technology, motion analysis, and injury prevention. Bioengineering, 12, 887.