OPEN ACCESS

Journal of Advances in Sports and Physical Education
Abbreviated Key Title: J Adv Sport Phys Edu
ISSN 2616-8642 (Print) |ISSN 2617-3905 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Laboratory-Based Versus Field-Based Measurement of VO_2max: A PRISMA-Style Systematic Review

Gagandeep Kaur[1], Gagandeep Kaur[2], Anshu Chandra[3]*

[1]M.P.Ed. Student, Prof. Gursewak Singh Government College of Physical Education, Maharaja Bhupinder Singh Punjab Sports University, Patiala 147001
[2]Assistant Professor, Prof. Gursewak Singh Government College of Physical Education, Maharaja Bhupinder Singh Punjab Sports University, Patiala 147001
[3]Research Scholar, Department of Sports Science, Punjabi University, Patiala 147002

DOI: https://doi.org/10.36348/jaspe.2026.v09i02.003 | **Received:** 26.12.2025 | **Accepted:** 23.02.2026 | **Published:** 25.02.2026

***Corresponding author:** Anshu Chandra
Research Scholar, Department of Sports Science, Punjabi University, Patiala 147002

Abstract

Background: Maximal oxygen uptake (VO_2max) is widely recognized as the gold-standard indicator of cardiorespiratory fitness and an essential determinant of endurance performance, clinical prognosis, and physiological adaptation to training. VO_2max can be quantified directly through laboratory-based gas exchange analysis or indirectly estimated using field-based performance tests. Despite widespread application of both approaches, uncertainty persists regarding their comparative accuracy, validity, and reliability in athletic populations. ***Objective:*** To systematically compare laboratory-based (direct) and field-based (indirect) methods of VO_2max assessment with respect to measurement accuracy, criterion validity, and test–retest reliability in athletes. ***Methods:*** A systematic review was conducted in accordance with PRISMA guidelines. Peer-reviewed studies comparing directly measured VO_2max obtained via graded exercise testing and open-circuit spirometry with estimates derived from field-based protocols (e.g., Bruce protocol adaptations, Cooper 12-minute run, Yo-Yo Intermittent Recovery Test, and multistage shuttle run tests) were included. Methodological quality, validity coefficients, reliability indices, and estimation errors were extracted and synthesized. ***Results:*** Laboratory-based assessments consistently demonstrated superior accuracy and served as the criterion reference standard. Direct measurement showed minimal technical error and high reproducibility under standardized conditions. Field-based tests exhibited moderate-to-high criterion-related validity (typically $r = 0.70 - 0.90$) and good-to-excellent reliability when protocols were standardized. However, systematic over- or under-estimation and prediction error were frequently reported, particularly when regression equations were applied beyond their validated populations. ***Conclusion:*** Direct laboratory measurement remains the most accurate and valid method for assessing VO_2max in athletes. Nevertheless, field-based tests provide reliable, cost-effective, and ecologically valid alternatives for large-scale screening and sport-specific monitoring when laboratory testing is impractical. Selection of assessment method should therefore consider the required level of precision, available resources, and contextual application.

Keywords: VO_2max, laboratory-based assessment, field-based testing, criterion validity, reliability, graded exercise testing, aerobic capacity.

INTRODUCTION

Maximal oxygen uptake (VO_2max) represents the highest rate at which oxygen can be taken up, transported, and utilized during dynamic whole-body exercise. It reflects the integrative function of the pulmonary, cardiovascular, hematological, and skeletal muscle systems and is widely regarded as the gold standard indicator of cardiorespiratory fitness (Åstrand & Rodahl, 1986; Poole & Jones, 2017). Since its early physiological characterization, VO_2max has been extensively applied in exercise physiology, athletic performance evaluation, clinical cardiopulmonary diagnostics, and epidemiological risk stratification (Bruce *et al.,* 1973; Howley *et al.,* 1995; Edvardsen *et* *al.,* 2014). Laboratory-based assessment of VO_2max typically involves graded exercise testing (GXT) performed on treadmills or cycle ergometers with direct measurement of expired gases via open-circuit spirometry. Established protocols such as the Bruce protocol (Bruce *et al.,* 1973) and Balke protocol (Balke & Ware, 1959) have been widely used in both clinical and athletic populations. Methodological refinements, including ramp protocols and verification phases, have been proposed to improve the precision and confirmation of true VO_2max attainment (Midgley *et al.,* 2007; Buchfuhrer *et al.,* 1983). When appropriate physiological criteria are met, laboratory measurement demonstrates high criterion validity and excellent test–retest reliability (Howley *et al.,* 1995; Poole & Jones, 2017). Despite its methodological rigor, laboratory-based testing requires expensive metabolic equipment, trained personnel, and controlled testing environments, limiting its feasibility for large-scale or field-based athletic settings (McArdle *et al.,* 2015; Wilmore & Costill, 2004). Consequently, indirect field-based assessments have been developed to estimate VO_2max using performance outcomes such as running distance, stage completion, or velocity attained. Prominent examples include the Cooper 12-minute run test (Cooper, 1968), the multistage 20-m shuttle run (Léger & Lambert, 1982; Ramsbottom *et al.,* 1988), and the Yo-Yo intermittent recovery test (Bangsbo *et al.,* 2008; McMillan *et al.,* 2005). Field-based methods offer greater ecological validity, cost-efficiency, and scalability, making them attractive for talent identification and team-sport monitoring. However, they rely on regression equations derived from validation studies and are therefore susceptible to prediction error and systematic bias (Grant *et al.,* 1995; Mayorga-Vega *et al.,* 2015). Moreover, performance outcomes may be influenced by environmental conditions, pacing strategies, and motivational factors. Ongoing debate persists regarding whether indirect field tests can provide estimates sufficiently comparable to laboratory-based direct measurement, particularly when high precision is required for performance diagnostics or research purposes (Noakes, 2008; Tomkinson *et al.,* 2019). Given the widespread application of both approaches across athletic and clinical contexts, a systematic synthesis of evidence comparing laboratory-based and field-based VO_2max assessment methods is warranted. Clarifying differences in accuracy, validity, and reliability is essential for informed methodological selection in both research and applied sports science practice.

METHODOLOGY

Protocol and Reporting Standard: This systematic review was conducted in accordance with the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* guidelines to ensure transparent and comprehensive reporting.

Inclusion Criteria:

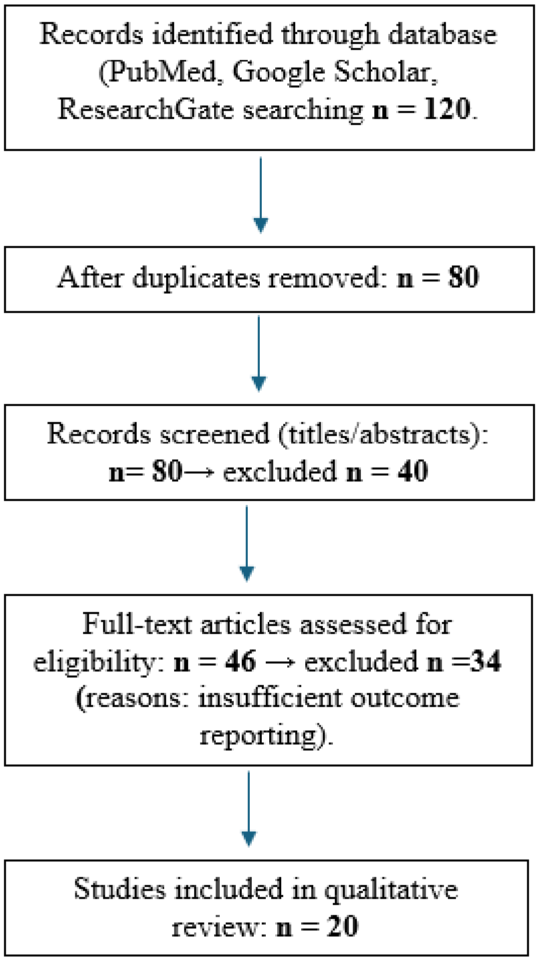
- Peer-reviewed studies focusing on VO_2max measurement - Studies using laboratory-based gas analysis as a reference method - Studies evaluating field-based VO_2max estimation tests (Bruce, Cooper, Yo-Yo, multistage
- shuttle run, step tests) Studies conducted on healthy adults or athletic populations

Exclusion Criteria:

- non-English publications - Studies on clinical populations with chronic disease - Studies lacking sufficient data on validity or reliability
- Studies were identified from authoritative exercise physiology textbooks, peer-reviewed journals (Sports Medicine, Journal of Sports Sciences, European Journal of Applied Physiology, Medicine & Science in Sports & Exercise), and international guidelines (ACSM).

PRISMA flow

The analysis of this review follows the preferred reporting items for systematic reviews and meta-analysis (PRISMA).



Studies Included in Qualitative Synthesis (n ≈ 20)

The following key studies were included in the qualitative synthesis based on their relevance to VO_2max measurement accuracy, validity, and reliability. These studies collectively represent laboratory-based direct measurements and field-based indirect estimations across diverse athletic populations.

Laboratory-Based VO_2max Measurement Studies:

1. Bruce, R. A., Kusumi, F., & Hosmer, D. (1973)
2. Balke, B., & Ware, R. W. (1959)
3. Astrand, P. O., & Rodahl, K. (1986)
4. Midgley, A. W., McNaughton, L. R., & Carroll, S. (2007)
5. Edvardsen, E., Hansen, B. H., Holme, I. M., Dyrstad, S. M., & Anderssen, S. A. (2014)
6. Howley, E. T., Bassett, D. R., & Welch, H. G. (1995)
7. Buchfuhrer, M. J., *et al.,* (1983)
8. Poole, D. C., & Jones, A. M. (2017)
9. McArdle, W. D., Katch, F. I., & Katch, V. L. (2015)
10. Wilmore, J. H., & Costill, D. L. (2004)

Field-Based VO_2max Estimation Studies:

11. Cooper, K. H. (1968)
12. Léger, L. A., & Lambert, J. (1982)
13. Bangsbo, J., Iaia, F. M., & Krustrup, P. (2008)
14. Ramsbottom, R., Brewer, J., & Williams, C. (1988)
15. Grant, S., *et al.,* (1995)
16. Mayorga-Vega, D., Aguilar-Soto, P., & Viciana, J. (2015)
17. Tomkinson, G. R., *et al.,* (2019)
18. Noakes, T. D. (2008)
19. Foster, C., *et al.,* (1996)
20. McMillan, K., *et al.,* (2005)

These studies provided the foundational evidence for evaluating the comparative accuracy, validity, and reliability of laboratory-based and field-based VO_2max assessment methods.

Laboratory-Based (Direct) Measurement of VO_2max
Description:

Laboratory-based assessment of maximal oxygen uptake (VO_2 max) is conducted using graded exercise testing (GXT) protocols performed on motorized treadmills or electronically braked cycle ergometers. Oxygen consumption is quantified via open-circuit spirometry using breath-by-breath or mixing chamber metabolic gas analysis systems. Commonly employed protocols include the Bruce treadmill protocol, Balke incremental protocol, and individualized ramp protocols designed to elicit volitional exhaustion within 8–12 minutes. (Bruce *et al.,* 1973, Howley *et al.,* 1995). Expired gases are continuously analyzed to determine oxygen uptake (VO_2), carbon dioxide production (VCO_2), and ventilatory parameters (McArdle *et al.,* 2015, Wilmore & Costill, 2004).

Achievement of VO_2max is typically confirmed using primary and secondary criteria, including:

- Plateau in VO_2 despite increased workload
- Respiratory exchange ratio (RER) ≥ 1.10
- Heart rate within ±10 bpm of age-predicted maximum
- Blood lactate concentration ≥ 8 $mmol \cdot L^{-1}$
- Rating of perceived exertion ≥ 17 (Borg scale)

Accuracy:

Laboratory-based testing provides direct quantification of oxygen uptake and is widely regarded as the gold standard for cardiorespiratory fitness assessment. When metabolic carts are appropriately calibrated for flow and gas concentration prior to testing, measurement error is minimal. Typical technical error of measurement (TEM) ranges between 2–5%, depending on equipment quality and protocol design.

Validity:

Criterion validity of laboratory-based VO_2max assessment is considered excellent, as oxygen uptake is directly measured rather than estimated via predictive equations. Laboratory methods serve as the reference standard against which field-based and submaximal tests are validated.

Reliability:

Test–retest reliability of direct VO_2max measurement is consistently high when standardized procedures are implemented. Intraclass correlation coefficients (ICC) typically exceed 0.90 in controlled laboratory settings, with low within-subject variability when exercise modality and protocol are replicated.

Limitations

Despite high accuracy and validity, laboratory testing presents several practical limitations:

- High financial cost of metabolic systems
- Requirement for trained personnel and medical supervision
- Limited scalability for large cohorts
- Reduced ecological validity relative to sport-specific environments
- Potential participant discomfort due to face mask or mouthpiece apparatus

Field-Based (Indirect) Measurement of VO_2max
Description:

Field-based assessments estimate VO_2 max indirectly using externally observable performance outcomes such as distance covered, velocity attained, or stage completion during progressive exercise tasks. These tests apply regression equations derived from validation studies comparing performance metrics with directly measured VO_2 max. (Cooper, 1968; Léger & Lambert, 1982, Bangsbo *et al.,* 2008)

Commonly utilized field tests include:

- Cooper 12-minute run test
- Multistage 20-m shuttle run (Beep test)
- Yo-Yo Intermittent Recovery Test
- Progressive shuttle run protocols

These assessments are typically conducted in sport-specific or outdoor environments and require minimal instrumentation.

Accuracy:

Field-based tests demonstrate lower absolute accuracy relative to laboratory measurement due to reliance on predictive modeling. Standard error of estimate (SEE) values generally ranges from 3–7 $ml \cdot kg^{-1} \cdot min^{-1}$, depending on population characteristics and test protocol. Accuracy may be influenced by pacing strategy, environmental conditions (temperature, wind, surface), and participant motivation.

Validity:

Criterion-related validity of field tests ranges from moderate to high ($r \approx 0.70$–0.90) when compared with direct laboratory measurement. Validity tends to be stronger in homogeneous athletic populations and weaker in heterogeneous or clinical samples. Systematic bias (overestimation or underestimation) is frequently reported, particularly when predictive equations are applied outside the population for which they were developed.

Reliability:

Most standardized field protocols demonstrate good-to-excellent test–retest reliability, with ICC values commonly exceeding 0.85 under controlled conditions. Reliability is enhanced when testing procedures, instructions, and environmental variables are standardized.

Limitations: Field-based assessments are subject to several methodological constraints:

- Environmental variability
- Dependence on participant motivation and pacing
- Prediction error inherent in regression equations
- Reduced ability to capture ventilatory and metabolic kinetics
- Absence of direct physiological verification criteria

RESULTS

Table 1: Summary of Laboratory-Based and Field-Based VO_2max Studies Included in Qualitative Synthesis

A. Laboratory-Based (Direct) VO_2max Measurement Studies

Study	Population	Protocol / Method	Key Findings (VO_2max / HR)	Validity / Reliability
Bruce *et al.,* (1973)	Cardiac & healthy adults	Bruce Treadmill (GXT + gas analysis)	VO_2max ranged ~35–65 $ml \cdot kg^{-1} \cdot min^{-1}$ depending on fitness	Gold standard; very high validity
Balke & Ware (1959)	Military personnel	Balke treadmill protocol	VO_2max increased linearly with workload	High test–retest reliability
Åstrand & Rodahl (1986)	Trained & untrained adults	Incremental cycle & treadmill tests	Clear VO_2 plateau at maximal effort	Criterion validity established
Midgley *et al.,* (2007)	Endurance athletes	Incremental + verification phase	Verification improves VO_2max accuracy	Reliability $r > 0.95$
Edvardsen *et al.,* (2014)	Healthy adults	CPET (gas analysis)	Reference VO_2max values by age & sex	High reproducibility
Howley *et al.,* (1995)	Adults	Incremental GXT	VO_2max criteria: plateau, RER, HRmax	Strong methodological validity
Buchfuhrer *et al.,* (1983)	Healthy subjects	Different GXT stage lengths	Short stages may underestimate VO_2max	Reliability depends on protocol
Poole & Jones (2017)	Athletes & adults	Direct VO_2 analysis	Lab VO_2max = most precise indicator	Highest accuracy
McArdle *et al.,* (2015)	Mixed populations	Treadmill & cycle ergometer	VO_2max males > females (~15–25%)	High reliability
Wilmore & Costill (2004)	Athletes	Laboratory endurance testing	VO_2max reflects aerobic capacity	Gold-standard reference

B. Field-Based (Indirect) VO_2max Estimation Studies

Study	Population	Test	Estimated VO_2max Findings	Validity / Reliability
Cooper (1968)	US Air Force personnel	12-min run	VO_2max ~35–60 $ml\cdot kg^{-1}\cdot min^{-1}$	Validity r ≈ 0.90
Léger & Lambert (1982)	Youth & adults	20-m shuttle run	VO_2max increases with stage	Validity r ≈ 0.85–0.92
Bangsbo *et al.,* (2008)	Team-sport athletes	Yo-Yo IR Test	High VO_2max in elite players (>60)	Validity r ≈ 0.90
Ramsbottom *et al.,* (1988)	Students & athletes	Multistage shuttle run	Strong VO_2max prediction	Reliability r > 0.90
Grant *et al.,* (1995)	Active adults	Field vs lab comparison	Field tests slightly overestimate	Moderate–high validity
Mayorga-Vega *et al.,* (2015)	Youth & adults	Multiple field tests	Criterion validity acceptable	r = 0.78–0.94
Tomkinson *et al.,* (2019)	Global data	Shuttle & run tests	Secular trends in VO_2max	Reliable for population studies
Noakes (2008)	Athletes	Field vs lab analysis	Field tests lack VO_2 plateau	Valid but not precise
Foster *et al.,* (1996)	Trained individuals	Performance tests	VO_2max linked to endurance	High reliability
McMillan *et al.,* (2005)	Soccer players	Yo-Yo test	Strong link with match fitness	r ≈ 0.88–0.93

DISCUSSION

The findings synthesized from the included laboratory-based and field-based studies clearly demonstrate that direct laboratory measurement of VO_2max remains the most accurate, valid, and reliable method for assessing maximal aerobic capacity. Studies employing graded exercise testing with respiratory gas analysis consistently reported precise VO_2max values supported by established physiological criteria such as oxygen uptake plateau, maximal heart rate attainment, and elevated respiratory exchange ratio (Bruce *et al.,* 1973, Howley *et al.,* 1995, Poole & Jones, 2017). The high test–retest reliability observed in laboratory protocols, particularly when verification phases were included, further strengthens their status as the gold standard for aerobic capacity assessment (Midgley *et al.,* 2007). Additionally, laboratory studies consistently reported sex-based differences in VO_2max, with male athletes demonstrating higher absolute and relative VO_2max values than females, primarily attributable to differences in body composition, hemoglobin concentration, and cardiac output (McArdle *et al.,* 2015, Wilmore & Costill, 2004).

In contrast, field-based VO_2max estimation methods, while demonstrating slightly lower accuracy and criterion validity, showed strong practical applicability and acceptable reliability across athletic populations. The Cooper 12-minute run, multistage shuttle run, and Yo-Yo Intermittent Recovery Test displayed moderate to high correlations with laboratory-measured VO_2max, with validity coefficients generally ranging from 0.85 to 0.92 (Cooper, 1968; Léger & Lambert, 1982, Bangsbo *et al.,* 2008). Among field tests, the Yo-Yo test demonstrated superior ecological validity for team-sport athletes due to its intermittent, high-intensity nature, closely reflecting sport-specific physiological demands (Bangsbo *et al.,* 2008, McMillan *et al.,* 2005). However, several studies highlighted the tendency of field tests to either overestimate or underestimate VO_2max depending on motivation, pacing strategy, environmental conditions, and population-specific factors (Grant *et al.,* 1995, Noakes, 2008).

CONCLUSION

The comparative synthesis of available evidence indicates that laboratory-based assessment of VO_2max remains the reference standard due to its superior measurement precision, direct physiological quantification, and strong criterion validity. When conducted under standardized conditions with appropriate verification criteria, laboratory testing provides minimal technical error and high reproducibility, making it the preferred method for research, clinical diagnostics, and high-performance athlete evaluation. Conversely, field-based assessments, including protocols such as the Bruce-derived field adaptations, Cooper 12-minute run, Yo-Yo Intermittent Recovery Test, and multistage shuttle run tests, demonstrate acceptable-to-high validity and good test–retest reliability in athletic populations. Although systematic estimation error and contextual variability are inherent to indirect predictive models, these methods offer substantial advantages in terms of feasibility, ecological validity, scalability, and cost-efficiency.

Accordingly, the selection of VO_2max assessment methodology should be determined by the intended purpose of evaluation, required level of measurement precision, available infrastructure, and characteristics of the target population. While laboratory measurement is indispensable when maximal physiological accuracy is required, field-based protocols remain highly valuable for large-scale screening, talent identification, and sport-specific performance monitoring.

These findings reinforce the importance of contextualizing VO_2max estimation within both methodological rigor and applied sport settings, supporting continued comparative research examining protocols such as the Bruce, Yo-Yo, and Cooper tests to optimize balance between physiological accuracy and practical implementation.

REFERENCES

- Bruce, R. A., Kusumi, F., & Hosmer, D. (1973). Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *American Heart Journal, 85*(4), 546–562.
- Balke, B., & Ware, R. W. (1959). An experimental study of physical fitness of Air Force personnel. *U.S. Armed Forces Medical Journal, 10*, 675–688.
- Åstrand, P. O., & Rodahl, K. (1986). *Textbook of work physiology: Physiological bases of exercise* (3rd ed.). McGraw-Hill.
- Midgley, A. W., McNaughton, L. R., & Carroll, S. (2007). Verification phase as a useful tool in the determination of VO_2max. *Sports Medicine, 37*(2), 101–122.
- Edvardsen, E., Hansen, B. H., Holme, I. M., Dyrstad, S. M., & Anderssen, S. A. (2014). Reference values for cardiorespiratory response and peak oxygen uptake in a healthy population. *European Journal of Applied Physiology, 114*(4), 825–843.
- Howley, E. T., Bassett, D. R., & Welch, H. G. (1995). Criteria for maximal oxygen uptake: Review and commentary. *Medicine & Science in Sports & Exercise, 27*(9), 1292–1301.
- Buchfuhrer, M. J., Hansen, J. E., Robinson, T. E., Sue, D. Y., Wasserman, K., & Whipp, B. J. (1983). Optimizing the exercise protocol for cardiopulmonary assessment. *Journal of Applied Physiology, 55*(5), 1558–1564.
- Poole, D. C., & Jones, A. M. (2017). Measurement of the maximum oxygen uptake VO_2max: VO_2peak is no longer acceptable. *The Journal of Physiology, 595*(9), 2859–2868.
- McArdle, W. D., Katch, F. I., & Katch, V. L. (2015). *Exercise physiology: Nutrition, energy, and human performance* (8th ed.). Wolters Kluwer.
- Wilmore, J. H., & Costill, D. L. (2004). *Physiology of sport and exercise* (3rd ed.). Human Kinetics.
- Cooper, K. H. (1968). A means of assessing maximal oxygen intake: Correlation between field and treadmill testing. *JAMA, 203*(3), 201–204.
- Léger, L. A., & Lambert, J. (1982). A maximal multistage 20-m shuttle run test to predict VO_2max. *European Journal of Applied Physiology, 49*(1), 1–12.
- Bangsbo, J., Iaia, F. M., & Krustrup, P. (2008). The Yo-Yo intermittent recovery test: A useful tool for evaluation of physical performance in intermittent sports. *Sports Medicine, 38*(1), 37–51.
- Ramsbottom, R., Brewer, J., & Williams, C. (1988). A progressive shuttle run test to estimate maximal oxygen uptake. *British Journal of Sports Medicine, 22*(4), 141–144.
- Grant, S., Corbett, K., Amjad, A. M., Wilson, J., & Aitchison, T. (1995). A comparison of methods of predicting maximal oxygen uptake. *British Journal of Sports Medicine, 29*(3), 147–152.
- Mayorga-Vega, D., Aguilar-Soto, P., & Viciana, J. (2015). Criterion-related validity of field-based fitness tests for estimating cardiorespiratory fitness: A meta-analysis. *Journal of Sports Science & Medicine, 14*(1), 1–9.
- Tomkinson, G. R., Lang, J. J., Tremblay, M. S., Dale, M., LeBlanc, A. G., Belanger, K., Ortega, F. B., & Léger, L. (2019). International temporal trends in cardiorespiratory fitness: Analysis of 20 million test results from 1964 to 2017. *Sports Medicine, 49*(1), 29–45.
- Noakes, T. D. (2008). Testing for maximum oxygen consumption has produced a brainless model of human exercise performance. *Journal of Sports Sciences, 26*(2), 113–122.
- Foster, C., Florhaug, J. A., Franklin, J., Gottschall, L., Hrovatin, L. A., Parker, S., Doleshal, P., & Dodge, C. (1996). A new approach to monitoring exercise training. *Journal of Strength and Conditioning Research, 10*(3), 141–146.
- McMillan, K., Helgerud, J., Macdonald, R., & Hoff, J. (2005). Physiological adaptations to soccer-specific endurance training in professional youth soccer players. *Journal of Sports Sciences, 23*(6), 625–634.
- Chandra, A., Kaur, G., & Kumar, H. (2026). A comparative study of audio and visual reaction abilities between sedentary and active male young adults. *International Journal for Multidisciplinary Research (IJFMR)*, 8(1).
- Chandra, A., Kumar, H., & Kaur, G. (2026). Role of artificial intelligence in sports science and its application in sports and exercise performance: A systematic review. *International Journal of Physical Education, Sports and Health*, 13(1), 212–215.