
BEEP TEST: ORIGIN AND APPLICATION IN CONTEMPORARY SPORTS PRACTICE

Yuliyana Zlatkova

South-West University “Neofit Rilski”, Bulgaria, zlatkov_foza@swu.bg

Krasimira Zlatkova

South-West University “Neofit Rilski”, Bulgaria, hristovakrassi@swu.bg

Abstract: The assessment of aerobic capacity is a fundamental element of functional diagnostics in sports science and practice. The maximal oxygen uptake (VO_{2max}) is considered the “gold standard” for evaluating cardiorespiratory fitness, but laboratory-based measurements are costly, require specialized equipment, and are not always feasible in large groups of athletes. To address this limitation, field-based protocols such as the 20-meter multistage shuttle run test (commonly known as the Beep test) have been developed and widely applied. The aim of the present review is to summarize the historical development, predictive equations, validity, reliability, and adaptations of the Beep test, and to highlight its application in contemporary sports practice. A literature review was conducted including both classical publications and recent validation studies. The analysis shows that the Beep test is valid and reliable for large-scale group screening and monitoring, particularly in student and endurance athlete populations, but results may vary depending on the sport and the predictive formula applied. In basketball and rowing, systematic under- or overestimation of VO_{2max} has been reported, while in female team sports only selected predictive equations provide accurate results. Modified protocols, such as 10-meter versions and child-specific adaptations, expand the usability of the test, and recent work demonstrates that even different sound signals can influence performance and enjoyment. The review concludes that the Beep test remains a practical and cost-effective tool for assessing aerobic fitness; however, practitioners should interpret results with caution and in relation to sport specificity. For precise individual diagnostics and training prescription, it is recommended to combine the Beep test with portable gas analysis systems, which provide direct measurement of gas exchange. This combined approach is particularly relevant for intermittent sports and combat disciplines such as taekwondo, where the accuracy of physiological assessment is essential. Furthermore, the review emphasizes that the Beep test holds value not only in elite sport, but also in physical education, clinical exercise settings, and population health monitoring, where quick and inexpensive assessment of aerobic fitness is required. By consolidating evidence from different populations and protocols, the present analysis provides guidance for coaches, physiotherapists, and sports physicians on how to apply and interpret the Beep test effectively in both research and practice. Future research should focus on refining predictive equations for specific sports and age groups, as well as exploring digital and sensor-based integrations that may further enhance the precision and usability of the Beep test in modern sports diagnostics.

Keywords: Beep test; shuttle run; aerobic capacity; VO_{2max} ; validity; sports practice

1. INTRODUCTION

The assessment of aerobic endurance is an essential component of functional diagnostics in athletes and physically active individuals. Maximal oxygen uptake (VO_{2max}) is considered the “gold standard” for determining cardiorespiratory capacity; however, its laboratory measurement is expensive, requires specialized equipment, and is difficult to apply to large groups. This has led to the development of field tests that are simpler, less costly, and more practical in sports settings.

In 1980, Léger and Boucher introduced the concept of a continuous, progressive, multistage running test (Université de Montréal Track Test), demonstrating its validity and reliability. Earlier physiological studies had indicated that the energy cost of running could vary depending on the surface (Margaria et al., 1963), as well as age, sex, and training status of the individuals (Bransford & Howley, 1977). Nevertheless, when validating the 20-meter multistage shuttle run test, Léger and Lambert (1982) showed that results were reliable regardless of sex and running surface. The concept was further refined and standardized into the 20-meter shuttle run test (MSFT, commonly known as the Beep test) (Léger, Mercier, Gadoury, & Lambert, 1988), which quickly gained popularity in sports, schools, and military practice as an indirect predictor of VO_{2max} . Contemporary studies and meta-analyses continue to confirm its wide application and scientific relevance (Mayorga-Vega, Aguilar-Soto, & Viciano, 2015).

The Beep test is characterized by its simplicity, the possibility of testing large groups simultaneously, and minimal financial costs. It is widely applied for assessing physical fitness in sports clubs, national teams, and educational practice.

The aim of the present review is to present the origin of the Beep test, to systematize available evidence regarding its validity and reliability, and to analyze the possibilities and limitations of its application in contemporary sports practice./

2. MATERIALS AND METHODS

The present article is a narrative literature review focusing on the origin, validity, and application of the 20-meter multistage shuttle run test (Beep test). A targeted approach was used to identify and select publications related to the development, validation, and practical use of the test.

The search was conducted in electronic databases including PubMed, Scopus, and Web of Science. The following keywords were applied: “20-m shuttle run test,” “beep test,” “multistage fitness test,” “VO₂max prediction,” and “validity reliability.” In addition, references cited in classical publications were screened to capture relevant studies not retrieved in the initial search.

Inclusion criteria were defined as follows:

- Classical publications describing the creation and standardization of the test.
- Validation studies comparing the Beep test with direct laboratory measurements of VO₂max.
- Contemporary research and meta-analyses (last 5–10 years) reflecting the application and limitations of the test in different populations (children, athletes, females, and males).
- Peer-reviewed articles published in English.

A total of 19 primary sources were analyzed, of which seven were published after 2015 to ensure contemporary relevance.

3. RESULTS

Predictive equations

Over the years, several equations have been developed to predict VO₂max from Beep test performance, based on the last completed stage or the maximal running speed. The original formula by Léger et al. (1988) was derived from data on children and adolescents. Later, Chatterjee et al. (2008) proposed an equation for female university students in India, while Ramsbottom et al. (1988) validated their formula in British military personnel. Flouris et al. (2005) suggested a modification aiming to improve accuracy in adult men. More recently, Magee et al. (2021) compared several of these predictive models in NCAA Division I female field hockey players and reported differences in predictive value.

Validity and reliability

A meta-analysis by Mayorga-Vega, Aguilar-Soto, and Viciano (2015), including more than 100 studies, demonstrated a high correlation between the Beep test and directly measured VO₂max ($r = 0.66\text{--}0.95$) with a standard error of $3\text{--}6 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. In NCAA female field hockey players ($n = 65$), Magee et al. (2021) found a directly measured VO₂max of $46.4 \pm 4.6 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. Among the predictive equations tested, the closest values were obtained with Ramsbottom et al. ($46.5 \pm 4.2 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) and Flouris et al. ($46.3 \pm 3.8 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$). In students and runners, Paradisis et al. (2014) reported a strong correlation ($r = 0.93$) between Beep test-derived VO₂max and direct measurements, with a mean value of $52.3 \pm 6.2 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. In young basketball players, Stojanovic et al. (2016) compared Beep test results with a portable metabolic analyzer (COSMED K4b2) and observed lower values, with a mean difference of $-6.54 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. In rowers, Senanayake et al. (2024) reported overestimation of VO₂max, with a mean difference of $+3.2 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. In soccer players, Aziz et al. (2005) compared the Beep test with the Yo-Yo intermittent recovery test and found lower accuracy for the Beep test ($\text{SEE} \approx 5 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$). In squash athletes and runners, St. Clair-Gibson et al. (1998) observed differences between measured and predicted VO₂max values ($63.4 \pm 6.1 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ in squash vs. $69.6 \pm 4.2 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ in runners).

Protocol adaptations

To facilitate broader use in various contexts, several modifications of the Beep test have been introduced. As early as 1982, Léger and Lambert proposed a version with one-minute stages for easier application in children and schools. In situations with limited space, 15-meter or 10-meter variants have been used. A study by Cho, Park, and Nam (2022) validated a 10-meter version of the test, reporting strong correlations with VO₂max (men: $r = 0.816$; women: $r = 0.821$; $\text{SEE} \approx 4\text{--}4.5 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$). More recently, the Modified Shuttle Test – Paeds, using a 10-meter track, was developed for children with reduced exercise capacity or health problems, while maintaining the principle of progressive stages guided by sound signals (Aertssen et al., 2024). This version enables reliable assessment of cardiorespiratory fitness in clinical and school settings where the standard 20-meter protocol is not feasible.

Contemporary research has also demonstrated that not only track length but also the type of sound stimulus can affect test outcomes. Ketelhut (2025) reported that using customized sound signals during the 20-meter shuttle run

test increased total distance covered, peak heart rate, and perceived enjoyment compared to the traditional “beep” sound.

Collectively, these findings confirmed that the Beep test is a valuable tool for assessing cardiorespiratory fitness, but its accuracy is highly dependent on the tested population, the sport, and the predictive formula applied. This emphasizes the need for careful interpretation and, when possible, combination with direct methods such as portable gas analysis systems.

4. DISCUSSIONS

The Beep test has been established as one of the most widely applied field methods for assessing cardiorespiratory endurance, owing to its simplicity, low cost, and suitability for testing large groups (Léger et al., 1988). Nevertheless, the present review indicates that its validity is not universal but depends strongly on the characteristics of the tested population and the predictive equations applied.

Studies in student populations and endurance runners have consistently confirmed high agreement with directly measured VO_2max , which supports the test’s usefulness in general fitness and endurance-oriented sports (Paradis et al., 2014). By contrast, in intermittent team sports such as field hockey, only a limited number of predictive equations have produced accurate estimates, highlighting the necessity of careful equation selection (Magee et al., 2021). Research in other disciplines further emphasizes the influence of sport-specific demands: in basketball, systematic underestimation has been reported (Stojanovic et al., 2016), whereas in rowing a tendency toward overestimation has been observed (Senanayake et al., 2024). Such findings confirm that the predictive accuracy of the Beep test is shaped by sport-specific physiological and mechanical factors.

Comparisons with other intermittent field protocols, such as the Yo-Yo test, suggest that the Beep test is less sport-specific for activities requiring frequent pace changes (Aziz et al., 2005). This is particularly relevant for combat sports, where aerobic and anaerobic demands are combined within an intermittent structure. Indeed, investigations in young taekwondo athletes have shown that the standard 20-meter protocol may not always correlate well with directly measured VO_2max (Chatterjee et al., 2006).

Protocol modifications—such as shortened 10-meter versions (Cho et al., 2022) or child-specific adaptations (Aertssen et al., 2024)—have expanded the applicability of the test in restricted spaces and clinical or school settings. However, these adaptations may compromise accuracy, reflecting the inherent trade-off between feasibility and validity. Recent developments even indicate that motivational and perceptual aspects can influence test outcomes: the use of customized sound signals has been shown to enhance performance and enjoyment compared with traditional beeps (Ketelhut et al., 2025).

From a practical perspective, the Beep test remains a valuable tool for group-level screening and for monitoring training progress over time. However, its individual accuracy is limited, and reliance on a single indirect method may lead to misinterpretation in sports where intermittent demands dominate. For this reason, combining the Beep test with direct gas analysis using portable systems such as COSMED K5 is recommended. This integrated approach provides more precise physiological profiling and is particularly valuable in combat sports like taekwondo, where accurate determination of aerobic capacity is essential for training prescription and therapeutic modeling.

5. CONCLUSIONS

The Beep test is one of the most widely applied field methods for evaluating aerobic capacity. It is reliable for group-level assessment and for monitoring training progress; however, its accuracy varies depending on the population, sport, and predictive formula applied. For individual diagnostics and precise training prescription, it is recommended that the Beep test be combined with direct gas analysis methods, such as the COSMED K5 system, particularly in intermittent sports and combat disciplines.

ACKNOWLEDGEMENTS

This review was carried out with the support of the internal funding project “Analysis of physical load through the application of a mobile gas analysis system in Taekwondo athletes” at South-West University “Neofit Rilski”, Blagoevgrad, Bulgaria (Project No. RP2/25).

REFERENCES

- Aertssen, W. F. M., van de Kamp, A., Jelsma, L. D., & Smits-Engelsman, B. C. M. (2024). Reliability and validity of the modified shuttle test-Paeds to measure cardiorespiratory fitness in children. *BMC Pediatrics*, 24(1), 343. <https://doi.org/10.1186/s12887-024-04812-0>

- Aziz, A. R., Tan, F. H., & Teh, K. C. (2005). A pilot study comparing multistage shuttle run test and Yo-Yo intermittent recovery test in soccer players. *Journal of Sports Science & Medicine*, 4(2), 105–112. <https://pubmed.ncbi.nlm.nih.gov/articles/PMC3880877/>
- Bransford, D. R., & Howley, E. T. (1977). Oxygen cost of running in trained and untrained men and women. *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, 43(6), 971–978. <https://doi.org/10.1152/jap.1977.43.6.971>
- Chatterjee, P., Banerjee, A. K., Das, P., Debnath, P., & Chatterjee, P. (2008). Validity of 20 meter multi stage shuttle run test for prediction of maximum oxygen uptake in Indian female university students. *Kathmandu University Medical Journal (KUMJ)*, 6(2), 176–180. <https://pubmed.ncbi.nlm.nih.gov/18769082/>
- Chatterjee, P., Banerjee, A. K., Majumdar, P., & Chatterjee, P. (2006). Validity of the 20-m multistage shuttle run test for the prediction of VO₂max in junior Taekwondo players of India. *International Journal of Applied Sports Sciences*, 18(1), 1–7. <https://www.researchgate.net/publication/261144318>
- Cho, H.-L., Park, H.-Y., & Nam, S.-S. (2022). Development of a multistage 10-m shuttle run test for VO₂max estimation in healthy adults. *Journal of Men's Health*, 18(1), 1–8. <https://doi.org/10.31083/jomh.2021.066>
- Flouris, A. D., Metsios, G. S., & Koutedakis, Y. (2005). Enhancing the efficacy of the 20 m multistage shuttle run test. *British Journal of Sports Medicine*, 39(3), 166–170. <https://doi.org/10.1136/bjism.2004.012658>
- Ketelhut, S., Benzling, V., & Hug, D. (2025). From simple beeps to complex sounds: Custom-designed sounds improve shuttle run test performance and perceived enjoyment. *Journal of Sports Sciences*, 43(15), 1462–1470. <https://doi.org/10.1080/02640414.2025.2502898>
- Léger, L. A., & Lambert, J. (1982). A maximal multistage 20-m shuttle run test to predict VO₂max. *European Journal of Applied Physiology and Occupational Physiology*, 49(1), 1–12. <https://doi.org/10.1007/BF00428958>
- Léger, L. A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sports Sciences*, 6(2), 93–101. <https://doi.org/10.1080/02640418808729800>
- Léger, L., & Boucher, R. (1980). An indirect continuous running multistage field test: The Université de Montréal Track Test. *Canadian Journal of Applied Sport Sciences*, 5(2), 77–84. <https://pubmed.ncbi.nlm.nih.gov/7389053/>
- Magee, M. K., White, J. B., Merrigan, J. J., & Jones, M. T. (2021). Does the multistage 20-m shuttle run test accurately predict VO₂max in NCAA Division I women collegiate field hockey athletes? *Sports*, 9(6), 79. <https://doi.org/10.3390/sports9060079>
- Margaria, R., Cerretelli, P., Aghemo, P., & Sassi, G. (1963). Energy cost of running. *Journal of Applied Physiology*, 18(2), 367–370. <https://doi.org/10.1152/jap.1963.18.2.367>
- Mayorga-Vega, D., Aguilar-Soto, P., & Viciano, J. (2015). Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *Journal of Sports Science & Medicine*, 14(3), 536–547. <https://pubmed.ncbi.nlm.nih.gov/articles/PMC4541117/>
- Paradis, G. P., Zacharogiannis, E., Mandila, D., Smirtiotou, A., Argeitaki, P., & Cooke, C. B. (2014). Multi-stage 20-m shuttle run fitness test, maximal oxygen uptake and velocity at maximal oxygen uptake. *Journal of Human Kinetics*, 41(1), 81–87. <https://doi.org/10.2478/hukin-2014-0036>
- 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. <https://doi.org/10.1136/bjism.22.4.141>
- Senanayake, S., Peiris, W., Rathnayake, R., et al. (2024). Validation study to assess the concurrent validity of the Beep test as a proxy for cardiopulmonary endurance using VO₂max as the criterion standard in rowers. *European Journal of Sport Science*, 24(2), 189–198. <https://doi.org/10.24018/ejsport.2024.3.1.131>
- St Clair Gibson, A., Broomhead, S., Lambert, M. I., & Hawley, J. A. (1998). Prediction of maximal oxygen uptake from a 20-m shuttle run as measured directly in runners and squash players. *Journal of Sports Sciences*, 16(4), 331–335. <https://doi.org/10.1080/02640419808559361>
- Stojanovic, M. D., Calleja-Gonzalez, J., Mikic, M., & Ostojic, S. M. (2016). Accuracy and criterion-related validity of the 20-m shuttle run test in well-trained young basketball players. *Montenegrin Journal of Sports Science and Medicine*, 5(2), 5–10. <https://doi.org/10.26773/mjssm.2016.08.00>