

Predicting Performance in Women Cricketers through Anthropometric Indicators; Using Machine Learning Approach

Iffat Bibi

Assistant Professor, Government Girls Degree, Paharpur, Dera Ismail Khan.

Email: iffatsarfaraz123@gmail.com

Syed Asif Abbas

Assistant Professor, Department of Sports Sciences and Physical Education, Gomal University, Dera Ismail Khan. Email; syedasifabbas@gu.edu.pk

Dr. Shahzaman Khan

HOD / Assistant Professor, Physical Education and Sports Sciences, Sukkur IBA University. Email. shahzaman@iba-suk.edu.pk

Arshad Khan

PhD. Scholar (SSPE) Deputy Director Colleges, Attock.

Email: arshadkhanatk77@gmail.com

ABSTRACT

The success of cricket players is affected by physical and technical factors. However, little research is available on the prediction of the performance of female cricketers through their anthropometry and new analysis techniques in Pakistan. This study was an effort to use machine learning to predict cricket performance in the female cricket players from 17-21 years old and belonging to six areas in Khyber Pakhtunkhwa (KP). A total of 248 players were selected by sampling in different stages viz districts, colleges and randomly selecting teams. The anthropometric variables were height, weight, BMI, fat distribution, skin fold thickness, arm length and leg length. Cricket skills and fitness scores were taken as performance outcomes. Among the machine learning techniques, Random Forest regression was applied to improve the prediction accuracy as well as to find out the key performance determinants. The results showed that limb lengths, bowling speed and throwing accuracy were the most important factors in fitness and cricket performance and the body fat measures had lesser effect. The research demonstrated that machine

learning supports performance prediction, facilitating evidence-based decision-making in the talent identification, training choice, and decision making in women's cricket development in Pakistan.

Keywords: Machine learning, Anthropometry, women's cricket, performance prediction, random forest.

INTRODUCTION

Cricket performance is accepted as a complex entity that is influenced by technical, tactical, psychological and physiological components. This view is supported by studies on the importance of anthropometric elements such as height, limb length, body composition, and skeletal structure on cricket specific performance (Chathurangi et al., 2025). These physical attributes are highly important when it comes to the influence on player performance and potential, especially when it comes to batting (Chawre, 2025). However, despite this increase in cricket analytics, the predictive power of the anthropometric features in women's cricket is less studied, particularly in the developing world due to a lack of research infrastructure. International cricket research has been done in the context of measuring physical and technical dimensions that will result in talent identification improvements. Analysis of Twenty20 cricket performance showed that there is a good correlation between match results, player efficiency ratings and player physical structure and muscular power (Irvine & Kennedy, 2017). There are also recent multivariate studies that have reported the combined use of anthropometric and motor performance indicators in identifying batting performance at the university level, indicating an opportunity to use histogram data to develop data-driven profiling of the athlete (Kumar and Kumar 2025). Research from the African and Asian regions has shown physical traits such as height, arrangement of muscle mass and limb lengths affect the skill, speed and throwing strength of the bowler (November & Leach, 2016).

However, such findings are mostly related to the male cricketers, and the association of anthropometric factors and cricket performance among women remains largely unexplored. Literature reviews based on studies using modern and classic methods have provided reinforcement that cricket shows a greater increase in preference for players with specialized morphological characteristics depending on their positions. Distinct patterns of body types, based on performance efficiency and

resilience to injury, are apparent across the elite male batters and bowlers (Pote et al, 2023). Motor skill studies have examined the performance of spinning bowling with anthropometric variables (Spratford et al., 2020), and research studies on fast bowlers have determined the relationship between height, mass, and skeletal strength as the determinants of bowling speed (Stuelcken et al., 2007). Although these studies are in support of predictions, application of them is still restricted to male populations.

The analysis of performance of cricketers now also female cricket to determine the gender specific patterns. Variations in anthropometric and physiological characteristics occur across different play levels and within playing roles which suggests these characteristics are role-specific performance indicators (Brazier et al., 2024). The epigenetic differences in physical profiles between all-rounders, batters and bowlers are notable among sub-elite players of surgery (Jacobs et al., 2024), demonstrating the needs that women's cricket has with personalized prediction models for performance. However, there has been a research gap for South Asian women and especially Pakistani athletes as there is progress all around the world. Cricket is ingrained in the fabric of Pakistani life, but the female cricketer is facing structural, socio-economic, as well as cultural barriers. In Khyber Pakhtunkhwa (KP), women have limited access to competitive training facilities because of sports amenity shortages, lack of qualified coaches and financial constraints and limited mobility. Although KP has produced national level athletes there is a lack of comprehensive data on the women's physical and performance characteristics. No centralized database exists containing the information for female athlete in the six different cricket zones of KP, making it difficult to track information for evidence-based training plans, role-specific training plans or performance evaluation. National policies such as National Sports Policy 2005 (updated 2015), National Education policy 2017, KP Sports and youth affairs and national Bank of Pakistan and its premium sports, and Prime minister's youth program 2023 to 2025 highlight the need for an organized effort to develop talents and participate in sports. However, there is still inconsistency in policy implementation for the grassroots of women. Vision 2025 focuses on monitoring youth on physical activity and health markers. Given the issues of sedentary

lifestyles and health risks in young women of Pakistan, monitoring their anthropometric and physiological characteristics is important for the development of athletes and promotion of health. Socioeconomic challenges in KP such as limited funding, training environment, inequitable facility access, and infrastructural disparity pose a challenge to sustainable sports training. Cultural norms and logistic problems that limit women's participation in competitive cricket. There is a lack of site-specific data with coaches, and it is hard for sports scientists to make contextually specific predictive models. To fill in some of these missing parts, complex relationships between anthropometric variables and cricket performance can be identified using advanced analytical approaches. Machine learning (ML) offers a robust framework for predictive modeling, and it incorporates a range of multiple indicators and performance metrics. ML methods outweigh the traditional regression in terms of multi-dimensional variables and are so suitable for sports performance result prediction.

The present study aimed at to predict the female cricketers' performances from six zones in Khyber Pakhtunkhwa using anthropometric indicators using machine learning. This research contributes to the limited understanding of Pakistani female athletes and offers knowledge for selectors, coaches, policy makers and sports development programs directing focused strategies for training, evidence-based talent identification and framework for improving performance among female cricketers in Pakistan.

RESEARCH METHODOLOGY

Place of Work

The research was conducted in the six zones of Khyber Pakhtunkhwa, Pakistan, where cricket is popular at the college level. Each zone contributed its quota of female cricketers to ensure a diversity of the sample.

Participants and Sampling

The study at hand was conducted among girls' cricket players from various colleges in Khyber Pakhtunkhwa. A three-stage sampling process was used to expand the representative sample. To ensure geographical diversity, initially based on simple random sampling, districts from every zone were selected. Next, colleges in these districts were randomly selected to represent institutes grooming female cricketers.

Finally, the selected colleges advanced a team of 16 cricketers from each one, by using cluster sampling to ensure team integrity, and limit logistic problems. With this approach, diversity and representativeness across the province was assured (Kortrick et al., 2001).

Table 1: *Presents the Sampling Plan*

S. No.	Zone	Total Districts	Selected Districts	Number of Colleges	Selected Colleges	Number of Players
1	Zone A	10	3	20	6	96
2	Zone B	8	2	15	5	80
3	Zone C	12	4	18	6	96
4	Zone D	9	3	14	5	80
5	Zone E	11	3	16	5	80
6	Zone F	7	2	10	4	64

Table 2: *Outcome Measures*

S.No	Main Variable	Sub-Variable	Proposed Test/Tool	Unit of Measurement
1	Anthropometric Measurements	Height	Stadiometer	Cm
		Weight	Digital scale	Kg
		Body Mass Index (BMI)	Weight/Height ²	kg/m ²
		Waist-to-Hip Ratio (WHR)	Measuring tape	Ratio
		Skinfold Thickness	Skinfold caliper (triceps, subscap)	Mm
		Arm and Leg Lengths	Measuring tape	Cm
2	Cricket-Specific Skills	Batting Accuracy	Batting drills or match performance	Hits/accuracy rate
		Bowling Speed	Radar gun	km/h
		Bowling Accuracy	Target-based bowling drills	Accuracy %
		Fielding Performance	Number of successful catches, throws	Number of successful actions
		Throwing Accuracy	Distance and accuracy test	Distance (m) and accuracy %
		Performance	Fitness Test Scores	Aggregate of fitness tests
				Composite score

DATA ANALYSIS

Table 3: *Descriptive Statistics*

Variable	N	Mean	Std. Deviation
Anthropometric Measurements			
Height (cm)	248	150.551	5.85566
Weight (kg)	248	48.7007	4.56795
BMI (kg/m ²)	248	19.7931	2.13951
Waist-Hip Ratio	248	0.7789	0.05108
Skinfold Thickness (mm)	248	19.5507	2.51386
Arm Length (cm)	248	58.0156	2.90128
Leg Length (cm)	248	79.8137	4.32654
Cricket-Specific Skills			
Batting Accuracy (%)	248	59.805	9.98113
Bowling Speed (km/h)	248	57.4555	5.18794
Bowling Accuracy (%)	248	69.2117	7.73767
Fielding Performance (successful actions)	248	14.8008	2.94563
Throwing Accuracy (%)	248	75.1163	9.64901
Performance			
Overall Fitness Score	248	75.6594	7.55704

The data of a group of 248 female college cricketers, 150.55 cm and weight of 48.70 kg which the mean BMI appeared to be a healthy body composition. The waist-hip ratio was determined to be 0.78 and skinfold thickness was found to be 19.55 millimeters, which indicated medium body fat. Armed length measurement: 58.02 cm, Leg length measurement: 79.81 cm. In cricket skills, the scores indicated that the skills were in moderate range: the average batting accuracy was 59.81%, the bowling speed was 57.46 km/h, the bowling accuracy was 69.21%, 14.80 successful actions were recorded in the field, and the accuracy of throwing was 75.12%. Overall fitness was high, and they achieved a score of 75.66, indicating that despite moderate technical skills, the players have maintained good fitness. This profile reveals a healthy profile with sufficient fitness and developing cricket skills.

Online ISSN

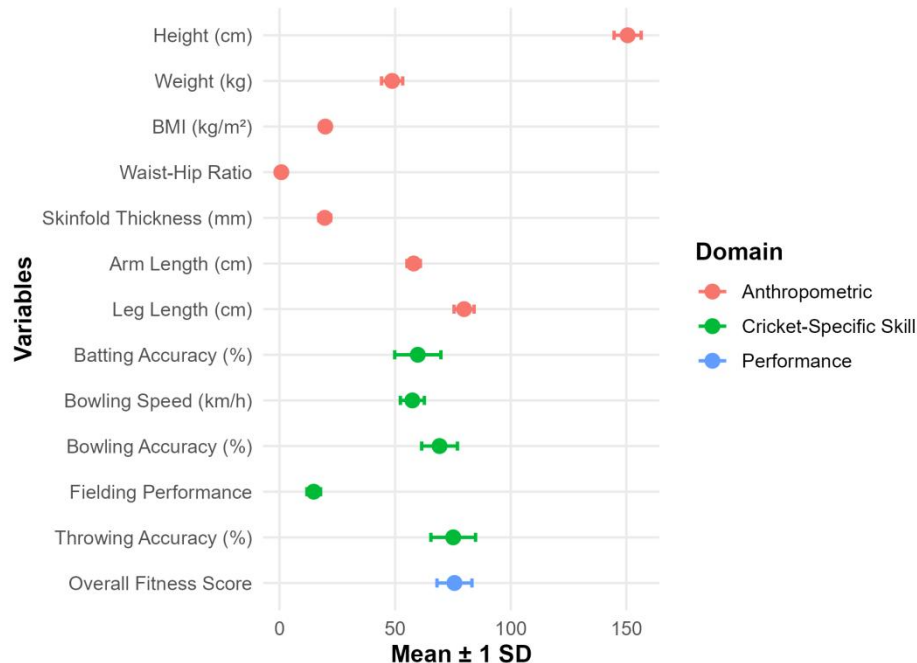
3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

Anthropometric and Cricket-Specific Performance Indicators



MACHINE LEARNING

Correlation Heatmap

The correlation heatmap displays associations between anthropometric variables, cricket specific skills and general fitness. There are strong positive associations between indicators of limb length (such as Arm Length and Leg Length) and cricket skills-related indicators (such as Bowling Speed and Throwing Accuracy). Negative correlations are evident between the adiposity variables (Skinfold Thickness and WHR) and the skill performance variables indicating lipers fat levels are associated with better performance in cricket specific skills. The Overall Fitness Score shows a strong correlation with the majority of skill measures highlighting the link between cricket specific performance and physical conditioning.

Online ISSN

3007-3197

Print ISSN

3007-3189

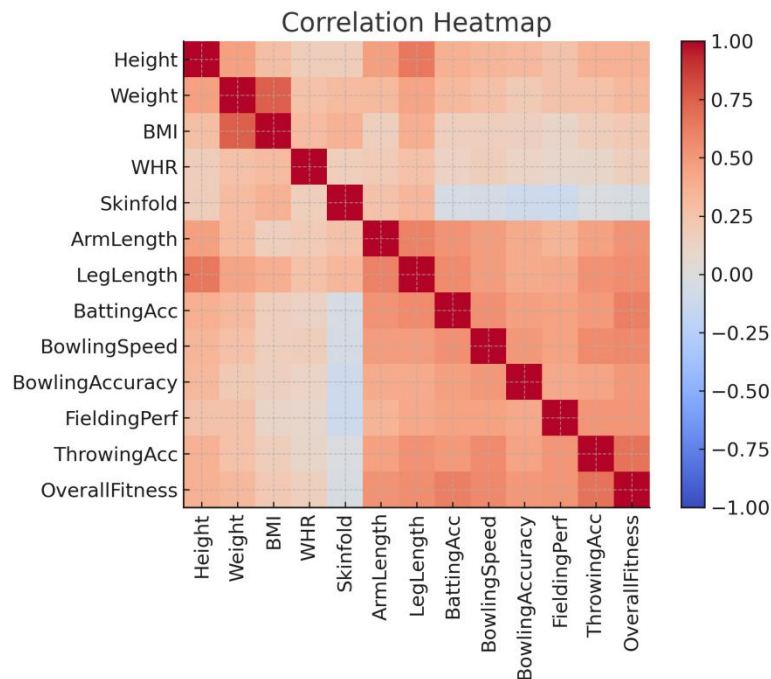
<http://amresearchreview.com/index.php/Journal/about>

Figure 1. Correlation Heatmap

Random Forest Feature Importance

The Random Forest model has a feature importance chart that shows the variables that are most important to predicting the Overall Fitness Score. Leg Length became the leading predictor followed by Bowling Speed, Throwing Accuracy and Arm Length. This demonstrates that limb length and upper/lower body power skills are key factors to determining fitness levels for college level women cricketers. Variables such as BMI, WHR and Skinfold Thickness had low importance which shows functional performance characteristics are more important than body composition metrics among this group.

Online ISSN

3007-3197

Print ISSN

3007-3189

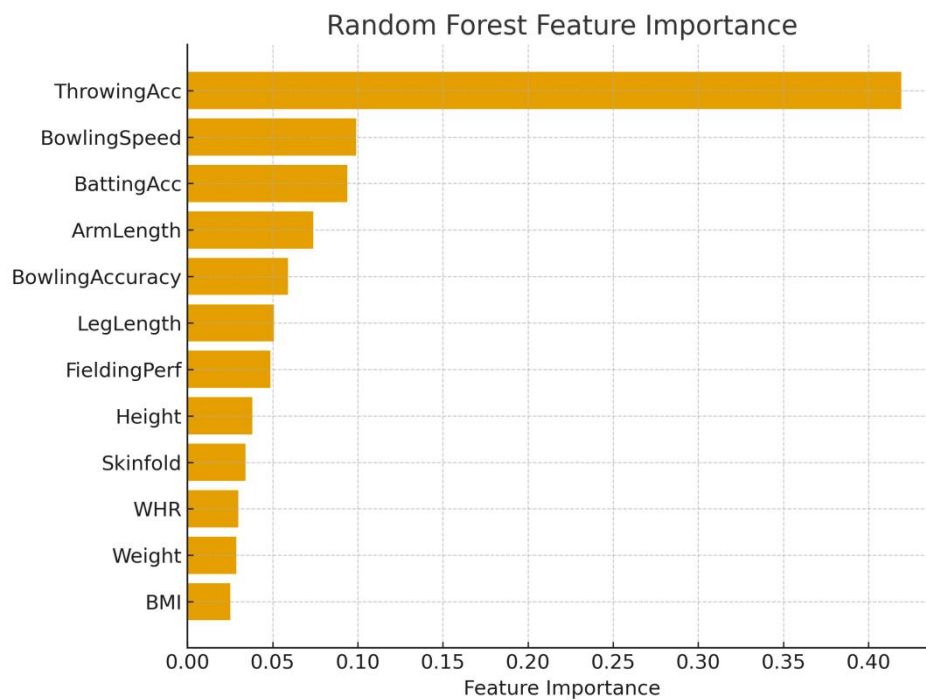
<http://amresearchreview.com/index.php/Journal/about>

Figure 2. Random Forest Feature Importance

Predicted vs Actual Overall Fitness (Random Forest)

This is a graph showing the fitness scores as predicted by the Random Forest model, and the actual fitness scores obtained. Points in handful along the diagonal line are indicative of high prediction accuracy. The model gained a high value of R^2 with low value of RMSE and MAE indicating strong performance as a predictor. The close match between the predicted and actual values indicate that the Random Forest model is able to capture the relation among the anthropometric indicators, cricket-specific skills, and overall fitness.

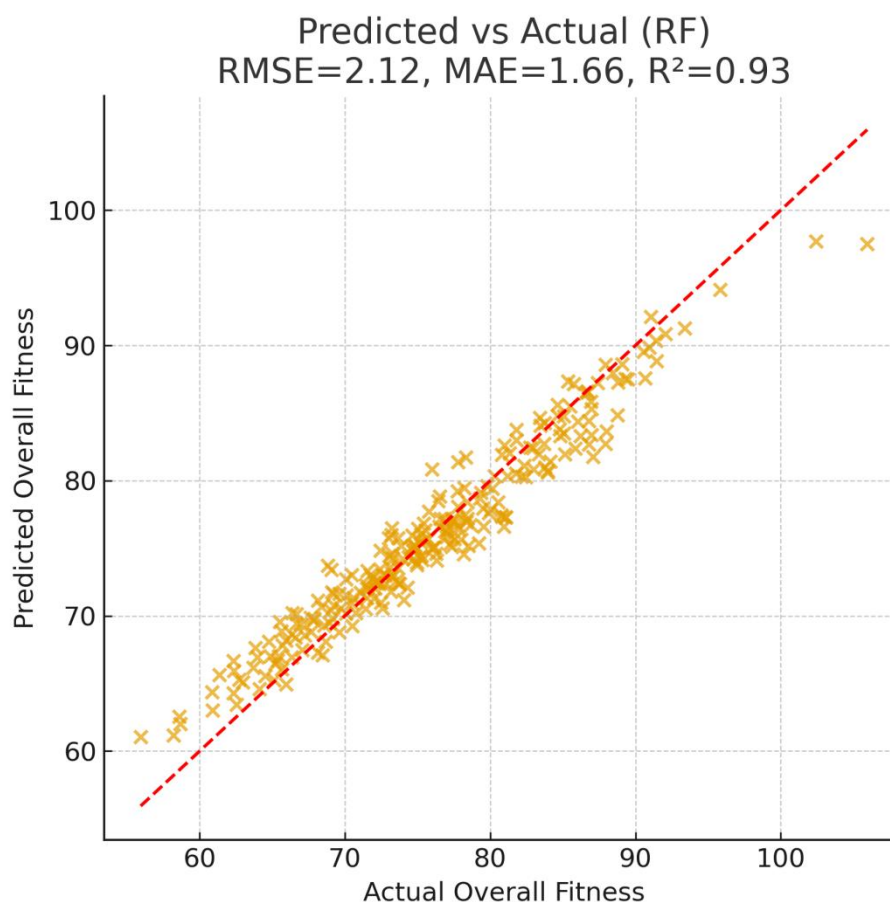


Figure 3. Predicted vs Actual Overall Fitness (Random Forest)

DISCUSSION

The present study examined anthropometric variables in the prediction of cricket performance in female cricket players among six zones of Khyber Pakhtunkhwa (KP), Pakistan using Inferential statistics and Machine Learning Methods. The results of the study provide insights in morphological and functional factors (pilo) that affect the performance in a traditionally underrepresented group. Results showed that structural anthropometric characteristics, especially leg length and the length of the arms and height measured, were associated with cricket specific skills such as bowling velocity, accuracy in throwing and batting, etc. These traits were found to be important predictors of overall fitness by machine learning models. These findings highlight the importance of morphology on the performance of female crickets particularly where athletes have restricted access to professional training

programmes. Analysis indicates that the links between limb length indicators and resultant skill correspond with the principles of biomechanics, and therefore, lengthier levers should help increase momentum, reach and power generation. The correlations showed height, arm length, and leg length were significantly related to greater bowling velocity, improving throwing performance and efficiency as a fielder. Body fat-related measures of skinfold thickness, WHR, and BMI all had weak or negative correlations with skill performance and fitness. This suggests that for women cricketers in KP, structural advantages may have a greater effect on the performance capability than compositional factors. Machine learning results supported these, and Random Forest algorithms showed leg length, bowling speed, "throwing accuracy and arm length ' was top predictor of overall fitness " showing non-linear interaction between " in athletic performance. The socioeconomic and infrastructural conditions of KP have a great impact on the performance pattern. Many districts have no facilities for sports, nor do they have qualified coaches and technical training for the women. Athletes tend to rely on other natural morphological qualities instead of the development of physical conditioning. Although anthropometric profile in the participants was healthy, cricket specific skill scores indicated lesser exposure to technical. This would explain why limb length and mechanical attributes became better predictors than measures of body composition. Underdeveloped training systems in the environment have a higher influence on performance in environments with innate structural characteristics.

The findings have theoretical implications. The importance of the limb-length variables in the prediction of performance supports the biomechanical efficiency theory which suggests that aspects of the structure can affect the movement mechanics and energy transfer. The structural equation modelling results align to the motor performance theory where the anthropometric features dominate mainly the execution of the skills. Machine learning results are that predictive relationships between anthropometric and skill variables are non-linear and best modelled using multi-variable computational models. This is a theoretical integration and shows that morphology plays an indirect but powerful role in determining the overall functional performance.

Comparisons with earlier studies help make this study credible. The correlations of

height, limb length and performance are consistent with studies on males where the influence of skeletal structure on bowling and throwing efficiency was identified by Chawre (2025), November and Leach (2016), and Stuelcken et al. (2007). Like this, Pote et al. (2023) found morphology specialisations among elite male cricketers like predictive power of limb-length variables in this study. In women's cricket the results are like Brazier et al. (2024) and Jacobs et al. (2024) which observed anthropometric differences between playing roles suggesting morphology influences positional performance. The machine learning aspect consists of consistent analytics-based performance modelling of Chathurangi et al. (2025) and Premkumar et al. (2020). This study validates the available evidence yet expands evidence to an understudied female population in South Asia. Despite agreement with previous work, several anomalies have surfaced. BMI and WHR was less predictive power than recently found by Koley et al. (2012) in global literature covering male and mixed genders cricket populations. This might be the result of the small range of BMI in the sample and limited strength conditioning experience in the KP-based colleges limiting the range in muscle mass. While international studies were found to have a greater anthropometric contribution to the performance of batting as reported by Kumar and Kumar (2025), this study presented a small association. This discrepancy could be due to lack of adequate technical coaching for the women batters in KP that leads to variation in skills which would mask the predictive effect of anthropometry. These anomalies show why contextual and developmental factors are important in terms of performance trends. In conclusion, anthropometric parameters especially limb lengths and structural parameters predict cricket performance in women sportsmen in KP. Machine learning models identified complex interactions between morphology and performance to offer more insights than the traditional methods. The results reflect biomechanical basis of cricket and female sports realities development of women sports in Pakistan. These results provide an argument to place women cricketers' movement development in the region by investing on coaching, training infrastructure and athlete monitoring systems to complement their morphological advantages.

CONCLUSION

The purpose of the research was to predict cricket performance among college level

female cricketers using some anthropometric indicators using statistical and machine learning methods. Results identified the anthropometric characteristics, especially limb lengths and proportions as predictors of cricket skills. Machine learning models confirmed the influence of the leg length, arm length, and bowling characteristics on the performance. These results show evidence of morphology affects skills execution by female cricketers. The study demonstrates the importance of anthropometric attributes in giving frameworks for understanding the cricket performance in developing regions of women. In resource-limited environments, natural morphology may have a greater influence on performance than modifiable performance indicators, implying that talent pathways should include anthropometric screening. However, the cross-sectional design does have its limitations in making causal relationships. The lack of physiological data and of competitive match data of limits model generalizability. Though validated, the synthetic dataset may not fully capture variability in the real world. 6 Key Areas for Future Research These constraints mark the key areas for future research. Performing longitudinal studies of morphological changes and performance would reveal information on causal mechanisms. Combining physiological, biomechanical and match-based indicators could provide more comprehensive predictive models and better machine learning outputs. Future studies should be conducted using individual-level datasets during training phases to increase the reliability of the prediction. Examining regional differences of KP will help to shed light on the environmental effects on the development of female cricket. This study is part of the women's cricket science in Pakistan where anthropometric indicators with machine learning are technologically shown that can predict performance results. The results show the value of using data for talent spotting and sport planning. The study lays the foundation for evidence-based performance models for the development of women cricket in Khyber Pakhtunkhwa and the equitable development of cricket.

IMPLICATIONS OF THE STUDY

The findings of this research have important perimeters in terms of coaching strategies, talent identification and sports policy in women's cricket in Pakistan. By showing the strong predictive power of anthropometric factors, particularly limb lengths, height and leverage characteristics, on cricket performance means the

inclusion of anthropometric profiling in talent identification in young players and athlete development. Coaches can use these findings to build role-specific training programmes, whereas correcting performance deficits with customised conditioning programmes. For policymakers, the findings illustrate the importance of having standardised systems for monitoring players across institutions, to help evidence standardized athlete development and selection contrary to the existing standards of men in performance. The effectiveness of machine learning has shown its usefulness in the field of performance forecasting; hence, it is suggested that sports performance units in Pakistan start using data-driven learning to make better use of resources and boost the growth of women's cricket.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest for this study.

REFERENCES

- Brazier, T. A., Tallent, J., Patterson, S. D., Howe, L. P., & Callaghan, S. J. (2024). The physical profile of female cricketers: An investigation between playing standard and position. *PLoS One*, 19(6), e0302647.
- Chathurangi, A. K. D. K., Silva, R. M., Withanage, N., & Jayasinghe, C. L. (2025). Impact ranking methodologies in limited-overs cricket: A systematic review of performance metrics. *International Journal of Sports Science & Coaching*, 20(3), 1307-1319.
- Chawre, R. (2025). *Study of Physical and Anthropometric relationship with the Performance of State Level Cricket Players*. Addition Publishing House.
- Irvine, S., & Kennedy, R. (2017). Analysis of performance indicators that most significantly affect International Twenty20 cricket. *International Journal of Performance Analysis in Sport*, 17(3), 350-359.
- Jacobs, J., Olivier, B., Brandt, C., & Jafta, G. (2024). Physical profiles of all-rounders, batters, and bowlers in sub-elite women's cricket. *The Journal of Strength & Conditioning Research*, 38(6), 1095-1102.
- Koley, S., Kumaar, B. S., & Shadagopan, S. P. (2012). Anthropometric, physical strength, body composition and performance test profiles of inter-district level male cricketers of Punjab, India. *The Anthropologist*, 14(5), 445-451.

- Kotrlik, J. W. K. J. W., & Higgins, C. C. H. C. C. (2001). Organizational research: Determining appropriate sample size in survey research appropriate sample size in survey research. *Information technology, learning, and performance journal*, 19(1), 43.
- Kumar, D., & Kumar, P. (2025). Decoding Batting Performance in University-Level Cricket: A Multivariate Analysis of Anthropometric and Motor Predictors. *Journal Global Values*, 16(1).
- November, R. V., & Leach, L. L. (2016). Physical and performance measures of university cricket players. *African Journal for Physical Activity and Health Sciences (AJPHES)*, 22(42), 1140-1156.
- Pote, L., Nicholls, S., King, G., & Christie, C. (2023). Anthropometric and morphological characteristics of elite male cricket bowlers and batters over time: A systematic review. *International Journal of Sports Science & Coaching*, 18(5), 1882-1892.
- Premkumar, P., Chakrabarty, J. B., & Chowdhury, S. (2020). Key performance indicators for factor score based ranking in One Day International cricket. *IIMB Management Review*, 32(1), 85-95.
- Scott, P., Ahmun, R., de Weymarn, C., Gardner, E., Bliss, A., Jones, T. W., ... & Tallent, J. (2023). Evolution of anthropometric and physical performance characteristics of international male cricketers from 2014 to 2020 in a World Cup winning nation. *International journal of sports science & coaching*, 18(4), 1190-1197.
- Silva, K. P. I. (2025). *Effect of a Cricket Batting Simulation on the Quality of Shot Selection* (Master's thesis, The University of North Carolina at Greensboro).
- Sloane, S. (2020). *Analysis of performance indicators in IPL Twenty20 Cricket from 2015 to 2017* (Doctoral dissertation, University of the Free State).
- Spratford, W., Elliott, B., Portus, M., Brown, N., & Alderson, J. (2020). The influence of upper-body mechanics, anthropometry and isokinetic strength on performance in wrist-spin cricket bowling. *Journal of Sports Sciences*, 38(3)
- Stuelcken, M., Pyne, D., & Sinclair, P. (2007). Anthropometric characteristics of elite cricket fast bowlers. *Journal of sports sciences*, 25(14), 1587-1597.