Title: Predictors of match performance in the South African Rugby Union Women's Premier Division
Running head: Predictors of performance in women's rugby

Authors:

Nicola Sewry[1,2], Candice MacMillan[1,2], Willie Maree[3], Clint Readhead[3,4], Lynne Cantwell[3], Esme Jordaan[5,6], Sonja Swanevelder[5]

Affiliations

1. Sport, Exercise Medicine and Lifestyle Institute (SEMLI), Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
2. IOC Research Centre, Pretoria, South Africa
3. South African Rugby Union, Tygerberg, South Africa
4. Health, Physical Activity, Lifestyle, and Sport (HPALS) Research Centre, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa
5. Biostatistics Research Unit, South African Medical Research Council (SAMRC), South Africa
6. Statistics and Population Studies, University of the Western Cape, Cape Town, South Africa

Address for correspondence:
Dr Nicola Sewry, Sport, Exercise Medicine and Lifestyle Institute (SEMLI), Faculty of Health Sciences, University of Pretoria, South Africa, Sports Campus, Burnett Street, Hatfield, Pretoria 0020, South Africa
Telephone: -27-12-420 6133
Email: nicolasewry@hotmail.com

Word Count:
Abstract: 192; Text body: 3569

Data sharing statement:
No additional data are available.

Ethical Clearance:
Ethics not required as all data are freely accessible on the internet.

Acknowledgements:
We want to thank all the players and SARU for their continued support of the project and contributions. A specific thank you to Mobii for the coded data.

Funding:
The South African Medical Research Council (SAMRC) provided partial funding for the statistical analysis.

ORCID ID:
Nicola Sewry: https://orcid.org/0000-0003-1022-4780
Candice MacMillan: https://orcid.org/0000-0002-5284-3208
Esme Jordaan: https://orcid.org/0000-0002-0361-3473
Sonja Swanevelder: https://orcid.org/0000-0003-0885-3190

Abstract

Purpose: Performance indicators in women's rugby union are under-researched. The aim of this study was to describe the match activities and determine the predictors for winning/losing and points difference in the South African Rugby Union (SARU) Women's Premier Division competition. Methods: During the 2021 and 2022 seasons, SARU-coded matches included 33 indicators (including attacking, defensive, set plays, and error categories) which were analysed for two outcome variables: winning/losing and points difference (points for MINUS points against). Results: The results of the study showed that for winning compared to losing, tries (OR=6.48; p=0.0001), goal kicking success % (OR=1.03; p=0.0245), and dominant tackles % (OR=1.45; p=0.0333) increased the odds of winning, whilst yellow cards decreased the odds (OR=0.27; p=0.0199). For points difference, Factor 1 (OR=1.016) including variables related to clean breaks, tries and conversions thereof, were predictive of a larger points difference, whilst Factor 6 (positive tackles, lineout wins; OR=0.958) favoured a close match/smaller points difference. Conclusion: The predictive indicators differed from those found in previous literature. However, there was little to no data available on women's rugby union to compare to, and therefore comparisons were made to men's rugby. More data and studies are required to confirm these findings.

Keywords: Performance indicators, rugby, women, winning

Introduction

Research has been performed in Women's XV rugby union looking at the match demands using global positioning system (GPS) technology.(1) GPS data give insights into the average distance (and positional variations thereof) the players cover during matches, which assist strength and conditioning coaches in understanding the type and training load the players should be exposed to.(2) However, there is much less information regarding winning and losing in Women's rugby union. Understanding winning and losing performance indicators in women's rugby, with the use of video analysis and tournament-wide coding, will enhance the ability of coaching staff to understand the landscape of their team's play. The coaches will then potentially be able to adapt coaching techniques to specifically address indicators that are predictive of winning or losing and how to improve their player's technique (i.e. tackling technique, or how they approach the breakdown).

Research performed at the World Cup found that women's teams who won (instead of lost) preferred a possession-driven approach in their matches.(3) However, with a sample size of only 8 matches, statistically, it would be difficult to draw accurate conclusions.(3) These performance indicators provide information to the coaches and assist with game plan development and training approaches. A recent Delphi study on women's rugby research, concluded there was only one paper that has investigated technical-tactical match characteristics in women's rugby union, illustrating the vast gap in the research.(4)

It should also be noted that all current studies related to match demands, performance indicators, and injury profiles in women's rugby originate from European countries or Australia (some that include alternate rugby formats, i.e., league and sevens). In those cases, the teams are also professional, compared to the South African league which is currently an amateur / semi-professional league (2021-2022). A recent study in Canadian women's varsity

rugby has described the match events, however it was purely a descriptive study, and did not investigate the indicators and how they affected winning / losing in the teams.(5)

The South African Rugby Union's (SARU) focus on the development and growth of the women's rugby game, has resulted in the Springbok Women's XV qualifying for the 2021 World Cup (but played in 2022 due to COVID-19) and the Sevens team qualifying as a core team for the 2023/2024 World Rugby Sevens series. The feeder for South African national teams is the SARU Women's Premier Division competition. This domestic competition provides a wealth of data for performance indicators to be investigated to understand the various actions that occur during a match (and how often) and how they affect a team's chance of winning or losing. Furthermore, it has been highlighted that women's rugby in South Africa has not been researched, and the specific challenges and opportunities that exist.(6) The landscape of South African women's rugby is different to other parts of the world (or hypothesized to be), and therefore there is a need to investigate the performance indicators that are predictive of winning or losing in this tournament. Understanding the predictors of winning/losing, will also assist in improving the level of play between the teams.

Therefore, the study aimed to i) describe the match activities during the SARU Women's Premier Division competition and then to ii) determine the predictors of winning vs. losing and iii) determine the predictors of the points difference in the same matches.

Methods

Sample

This was a cross-sectional analysis of data freely available on the South African Rugby Union (SARU) website (https://www.sarugby.co.za/match-centre?seriesId=78a91764-7272-4615-9b74-bca43c6b3266&returnUrl=/tournaments/women-s-premier-division/), however, SARU did send excel spreadsheets of these data to the authors to facilitate analysis. These predictors were chosen/included as they are available to all teams and therefore all for widespread use. There are 15 provincial rugby unions that play senior elite women's rugby in South Africa. In 2021 the Women's Premier Division rugby tournament consisted of the top six (6) provincial teams playing in a round-robin format tournament. In 2022, the top seven (7) participating teams played a double round of home and away. There were 3 cancelled matches in 2021. In total, 50 matches were played, but 3 matches' data are missing, thus there were 94 individual team matches included in the dataset (47 matches). 2021 was the first year data were available, hence the sample size of 2021 and 2022.

Data

The indicators included were all the indicators coded by Mobii. Performance indicators were grouped into attacking, defensive, set plays, and errors. Indicators have further explanations in brackets.

Attacking indicators included the following: tries, conversions attempted, conversions successful, penalty kicks attempted, penalty kicks successful, goal kicking success %, ball carries, meters gained, offloads, clean breaks, attacking breakdown arrivals, general play kicks (whenever a player kicks the ball out of hand during general play), and general play kick errors (a general play kick that is kicked out on the full, taken back into own 22m and then kicked out on the full, kicked dead in goal).

Defensive indicators included: tackles (individual tackles; when a defending player makes contact with a ball carrier with the intent on tackling the attacking player/bringing the player

to the ground), tackles made (individual tackles; when a defending player successfully brings the ball carrier to the ground to complete the tackle), tackles missed (individual tackles; when a defending player loses contact with the ball carrier and the ball carrier is still gaining ground after the tackle attempt), tackles made %, dominant tackles (when the defender successfully attempts a tackle and drives the ball carrier towards his/her own try line, the attacking team loses territory or momentum as a result of the tackle), dominant tackles % (dominant tackles / tackles attempted = %), defensive turnovers won (counter rucks won – breakdowns; ball steals – breakdowns; forced penalties – breakdowns; interceptions won; turnovers forced – tackles including: tackling a ball carrier into touch, tackling a ball carrier, forcing a knock on, tackling a ball carrier forcing the ball carrier to lose possession resulting in the defending team winning possession; turnovers forced – maul including: stealing the ball in a maul, forcing a static maul – ball unplayable), defending breakdown arrivals (whenever a player from the defending team enters a breakdown), and breakdown turnovers won (when any of these occur during a breakdown: counter rucks won; ball steals; forced penalties).
Set play indicators included: scrums, scrums won, scrum penalties conceded, lineouts, lineouts won, and mauls.
Error indicators included: turnovers conceded, penalties conceded, free kicks conceded, yellow cards, and red cards.
All video analysis coding was performed by Mobii (a commercial entity). The Match Coding Team at Mobii consists of multiple Data Capturers. The match is usually split amongst two Data Capturers. Each will code their split of the game and then a two-step post-coding validation/check will be done as per the two steps, 1) check major events (e.g. tackles, ball carries etc), 2) player validation (ensure the player is correctly coded by both coders).
The consensus statement on video analysis in rugby is more detailed than the coding included here, but most align with the consensus.(7)

The team's "performance" was also coded using the match results. Whether the team won or lost (binary), and the points difference [POINTS DIFF](calculated as: "points scored by the team" MINUS "points scored against the team")(continuous) were coded.

Data Analysis
Descriptive statistics were reported for all individual indicators for all team matches reporting N, mean(SD), median, min, max, Q1 and Q3.

Part 1: Win-Lose outcome
The predictors for the binary outcome win-lose was first modeled by performing individual univariate logistic regressions for each individual indicator, where odds ratios (OR's), 95% confidence intervals (95%CI's), R-square values, and p-values were reported.

All individual predictors from the univariate logistic regressions that were marginally significant (at $p<0.10$ level) were entered into a multiple logistic regression using a forward stepwise selection process. These predictors in the multiple logistic regression were tested for collinearity using Spearman correlation coefficients. The final predictive model only used predictors that had the highest predictive value and were not collinear with any other predictors. Additional predictors were added to the model until the ROC (receiver operating characteristic curve) index (AUC – area under the curve) did not improve further. The final predictive model was assessed by the ROC index as well as with the Gini coefficient. When adjusting the final results for teams as a cluster for a specific match (i.e. 'game'), the results

were similar (mixed modelling, using PROC GLIMMIX in SAS with 'game' specified as a cluster and a standard variance covariance structure).

Part 2: Points Difference outcome (Two-step factor analysis)
First step: performed a Principal Component Analysis. This is also a variable reduction procedure which makes no assumption about the underlying causal model.
In the exploratory factor analysis of 27 rugby match variables, a varimax rotation was used, which resulted in 6, orthogonal (uncorrelated) factors, as a best fit solution for the data (Table 4). Approximately 66% of the total variance is accounted for by these 6 factors combined. Five of the match variables did not load onto any of the 6 factors, however they were entered as individual items in the subsequent ordinal multivariate logistic regression (Step 2). Step 2: perform an ordinal multivariate logistic regression where the response outcome (absolute value of the points difference for a match) had 3 ordered levels. The 3 levels were: Level 1 = unbalanced match (points difference >25), Level 2 = balanced match (points difference 9 - 25) and Level 3 = close match (points difference < 8); these levels were decided upon after considering statistical models (enough samples in each group) and rugby scoring (1= >3 converted tries; 2= 2-3 converted tries; 3= 1 converted try/try and a penalty). With respect to these 3 outcome levels, probabilities modelled were cumulated over the lower ordered response levels. By using this ordering, the lower categories will represent the more favorable match results (i.e. larger points difference). The abovementioned 6 factors from Step one were entered as 6 covariates in the ordinal multivariate logistic model. None of the 5 individual items (mentioned in Step 1), that were entered into this model were significant, so consequently removed. The score chi-square for testing the proportional odds assumption was 9.0311, which is not significant (p=0.1718), indicates that the proportional odds assumption is reasonable. The ordinal multivariate logistic model could not accommodate clustering for 'game'.
All statistical analyses were performed using SAS(v9.4). Significance was at the 5% level, unless stated otherwise.

Results

Description of match activities
In total there were 94 individual team matches (2021=54; 2022=40) (47 matches), where none of the matches ended in a draw. The number of matches per team varied in 2021 due to cancellations (a team played between 8-10 matches) and 6 each in 2022. There were three red cards during the tournaments (2 in 2021; 1 in2022), however due to the low numbers these were not included in the analysis. The descriptive information for the match activities (performance indicators) of each team match is shown in Table 1 (details per winning and losing team can be found in Supplementary Table 1).

Table 1: Performance indicators occurring in all team matches

	All Team Matches (n=94)					
	Min	Q1	Mean (SD)	Median	Q3	Max
Attacking						
Tries	0	1	3.1 (2.8)	2	4	15
Conversions Attempted	0	1	3.0 (2.7)	2	4	15
Conversions Successful	0	0	1.3 (1.6)	1	2	8
Penalty Kicks Attempted	0	0	0.9 (1.1)	1	1	4
Penalty Kicks Successful	0	0	0.5 (0.8)	0	1	4
Goal Kicking Success (%)	0	0	39.5 (30.6)	47.5	60	100
Ball Carries	37	91	116.6 (33.6)	114	138	218
Metres Gained	241	515	747.8 (285.8)	705.5	932	1734
Offloads	4	11	16.3 (7.0)	16	21	37
Clean Breaks	0	1	3.0 (3.1)	2	4	21
Attacking Breakdown Arrivals	56	139	182.0 (55.3)	184	211	361
General Play Kicks	1	9	12.9 (5.7)	12	17	29
General Play Kick Errors	0	0	0.8 (0.9)	0	1	3
Defensive						
Tackles	53	147	183.3 (56.1)	181	222	361
Tackles Made	46	119	146.8 (44.3)	146	172	286
Tackles Missed	6	24	36.9 (18.0)	34.5	46	87
Tackles Made (%)	63	77	80.7 (6.0)	81	85	95
Dominant Tackles	0	5	9.0 (5.0)	8	12	28
Dominant Tackles (%)	0	3	5.1 (2.8)	5	7	18
Defensive Turnovers Won	0	0	1.0 (1.2)	1	2	4
Defending Breakdown Arrivals	15	47	63.3 (22.6)	61	80	131
Breakdown Turnovers Won	0	3	5.0 (2.7)	5	7	13
Mauls	0	1	2.3 (2.3)	2	4	12
Set Plays						
Scrums	2	7	9.6 (3.4)	9.5	12	19
Scrums Won	2	6	8.8 (3.5)	8	11	18
Scrums Won %	50	83.3	90.8 (11.9)	100	100	100
Scrum Penalties Conceded	0	0	1.2 (1.4)	1	2	6
Lineouts	3	8	9.7 (3.2)	9	12	17
Lineouts Won	1	4	6.6 (2.9)	6	9	14
Lineouts Won (%)	25	55.6	66.2 (16.7)	66.7	75	100
Errors						
Turnovers Conceded	5	10	14.1 (4.7)	14	18	24
Penalties Conceded	3	11	14.4 (4.5)	14	18	26
Free Kicks Conceded	0	0	0.7 (0.9)	1	1	4
Yellow Cards	0	0	0.6 (0.9)	0	1	4

Table 2: The odds of a team winning (vs losing) for each attacking, defensive, set play and error performance indicator (univariate logistic regression)

		Univariate logistic regression	
	R^2	OR (95%CI)	p-value
Attacking			
Tries	0.44	3.46 (2.04-5.87)	**<0.0001****
Conversions Attempted	0.44	3.59 (2.10-6.15)	**<0.0001****
Conversions Successful	0.37	5.02 (2.46-10.27)	**<0.0001****
Penalty Kicks Attempted	0.00	1.12 (.076-1.66)	0.5564
Penalty Kicks Successful	0.02	1.45 (0.85-2.47)	0.1691
Goal Kicking Success (%)	0.10	1.02 (1.01-1.04)	**0.0025****
Drop Kicks Attempted	0.01	0.32 (0.03-3.18)	0.3301
Ball Carries	0.04	1.01 (1.00-1.03)	**0.0562***
Metres Gained	0.13	1.00 (1.00-1.01)	**0.0013****
Offloads	0.01	1.03 (0.97-1.09)	0.3784
Clean Breaks	0.23	1.73 (1.30-2.30)	**<0.0001****
Attacking Breakdown Arrivals	0.01	1.00 (1.00-1.01)	0.2777
General Play Kicks	0.01	1.04 (0.96-1.11)	0.3458
General Play Kick Errors	0.01	1.28 (0.82-1.98)	0.2748
Defensive			
Tackles	0.06	0.99 (0.98-1.00)	**0.0221****
Tackles Made	0.03	0.99 (0.98-1.00)	0.1149
Tackles Missed	0.12	0.96 (0.93-0.98)	**0.0017****
Tackles Made (%)	0.10	1.12 (1.04-1.22)	**0.0042****
Dominant Tackles	0.03	1.07 (0.99-1.17)	0.1079
Dominant Tackles (%)	0.09	1.28 (1.07-1.52)	**0.0061****
Defensive Turnovers Won	0.01	1.13 (0.80-1.60)	0.4839
Defending Breakdown Arrivals	0.01	0.99 (0.97-1.01)	0.3430
Breakdown Turnovers Won	0.00	1.01 (0.87-1.18)	0.8787
Mauls	0.00	1.02 (0.85-1.22)	0.8195
Set Plays			
Scrums	0.03	0.90 (0.79-1.02)	**0.0903***
Scrums Won	0.03	0.90 (0.80-1.02)	**0.0924***
Scrum Penalties Conceded	0.02	1.22 (0.90-1.67)	0.2024
Lineouts	0.01	0.94 (0.83-1.07)	0.3594
Lineouts Won	0.00	1.02 (0.89-1.18)	0.7475
Lineouts Won (%)	0.07	1.04 (1.01-1.07)	**0.0107****
Errors			
Turnovers Conceded	0.02	1.06 (0.97-1.16)	0.1928
Penalties Conceded	0.01	0.96 (0.88-1.06)	0.4248
Free Kicks Conceded	0.04	1.64 (0.98-2.73)	**0.0600***

Yellow Cards	0.07	0.50 (0.28-0.88)	**0.0157****

*marginally significant (0.05<=p<0.10)
** significant (p<0.05)

Part1: Win-Lose predictors

The individual (univariate) odds of a team winning compared to losing for each of the various performance indicators are shown in Table 2 (further information regarding coefficients are in Supplementary Table 2). Numerous performance indicators were either significant or at least marginally significant in predicting the odds of a team winning in the univariate analysis. Out of the attacking variables, tries, conversions (attempted and successful), goal kicking success, meters gained and clean breaks were all significant in increasing the odds of winning. Defensive indicators that were significant were all tackle-related (all tackles, tackles missed, tackles made %, dominant tackles %), and depending on the tackle type, either increased the odds or decreased the odds of winning. There were only 4 yellow cards reported, consequently, this error indicator was only marginally significant.

In the final predictive model, tries (OR=6.48, increased the odds >6 times), dominant tackles % (OR=1.45, 45% increased odds) and goal kicking success % (OR=1.03, 3% increased odds) increased the odds of winning, whilst a yellow card (OR=0.27, 73% decreased odds) decreased the odds of winning. Tackles made % and lineouts won % were not significantly associated at the 5% level in the multiple logistic model, however they did improve the predictive ability of the model (ROC index of 0.9525 vs 0.9340) and were therefore included in the final model.

Table 3: The odds of a team winning (vs losing) for each attacking, defensive, set play and error performance indicator (multiple logistic model)

	Final Predictive Logistic Regression model	
	OR (95%CI)	p-value
Attacking		
Tries	6.48 (2.51-16.77)	0.0001
Goal Kicking Success (%)	1.03 (1.00-1.06)	0.0245
Defensive		
Tackles Made (%)	1.13 (0.98-1.31)	0.0823
Dominant Tackles (%)	1.45 (1.03-2.04)	0.0333
Set Plays		
Lineouts Won (%)	1.03 (0.98-1.08)	0.2090
Errors		
Yellow Cards	0.27 (0.09-0.81)	0.0199

ROC index (AUC): 0.9525
Gini coefficient=0.905

Part 2: Points difference analysis

The distribution of the points difference outcome was normally distributed (min=1 ; Q1=7 ; mean=20.34; Q3=28; maximum=79). Supplementary table 3 reports the 6 factors created by the Principal Component Analyses procedure, as well as the loadings of individual match variables to each of the 6 factors. The six factors were:

Factor 1: [tries, conversions attempted, conversions successful, clean breaks] relates to: tries, and clean breaks which can often lead to a try (an attacking factor)
Factor 2: [ball carries, metres gained, offloads, attacking breakdown arrivals, turnovers conceded] relates to: attack plays, and conceding a turnover
Factor 3: [scrums, scrums won, general play kicks, general play kick errors] relates to: scrums, and general play kicks, which can often occur post scrum by the scrumhalf
Factor 4: [defensive turnovers won, defensive breakdown arrivals, breakdown turnovers won] relates to: the breakdown
Factor 5: [penalty kicks attempted, penalty kicks successful, goalkicking successful %] relates to: penalty kicks
Factor 6: [tackles made %, dominant tackles %, lineouts won %] relates to: tackles and lineouts, and therefore set plays and defence

Five match variables did not load on any of the 6 factors, they were mauls, scrum penalties conceded, penalties conceded, free kicks conceded and yellow cards. Table 4 reports the ORs (95%CI) for the 6 factors from the ordinal multivariate logistic regression. Only Factors 1 and 6 were significant. Factor 1 (OR=1.016) indicates that Factor 1 is predictive of being in the lower ordered response levels, suggesting that match variables that relate to clean breaks, tries and conversions thereof, are predictive of an unbalanced match, or larger points difference. Factor 6 (positive tackles, lineout wins; OR=0.958) favoured a close match, or smaller points difference. It is important to note that although both Factors 1 and 6 are significant, they are not highly significant. See Supplementary Figure 1 for the spread of the data within each response outcome level.

Table 4: The odds of a factor being predictive of a larger match points difference

Effect	**Odds Ratio (OR; 95%CI)**
Factor 1	*1.016 (1.001-1.031)*
Factor 2	0.993 (0.984-1.001)
Factor 3	1.000 (0.946-1.057
Factor 4	1.006 (0.987-1.026)
Factor 5	1.000 (0.980-1.021)
Factor 6	*0.958 (0.928-0.989)*

Response Outcome Level 1 = unbalanced match (points difference >25) (N = 28),
Response Outcome Level 2 = balanced match (points difference 9 - 25) (N = 32),
Response Outcome Level 3 = close match (points difference $\leq$ 8) (N = 34)
OR>1 predictive of lower response level (unbalanced match) OR<1 predictive of higher response level (close match)
ROC index for this model is 0.6770. R-squared = 0.1664. Wald Chi square = 14.94 (p-value = 0.0207).

Discussion

In this study, we aimed to describe the activities performed by each team during a match and then determine the predictors for winning or losing and for points difference. The main findings of the paper were 1) the four factors predictive of winning or losing: tries, dominant tackles %, goal kicking success % and yellow cards; 2) the factors predictive of points difference were the combination of clean breaks, tries and the conversions (larger points difference); and positive tackles and lineout wins were predictive of close matches (smaller points difference).

This study's descriptive data provides novel data on the match activities that occurred during the women's provincial rugby union competition. The performance indicators (match activities) used in various other studies vary, however many are consistent and allow for comparison. For this tournament, teams, on average, made 147 individual tackles (winning teams=140; losing=154) and missed 37 tackles (winning=31 [22.1% missed]; losing=43 [27.9% missed]). Having said that, in Canadian varsity women's rugby, a team performed on average 280 tackles per match,(5) potentially indicating that the more amateur a competition, the higher the number of tackles. For penalties, data from an elite women's study which considered 8 women's matches, showed that winning teams conceded 8 penalties, and losing teams conceded 9.3.(3) This was much lower than our study, where on average, a team conceded 14 penalties, regardless of winning or losing. One potential reason for this difference, could be the level of play, where the SA tournament is amateur /semi-professional, compared to elite and professional tournaments. Total ball carries were also much higher in our cohort (117 overall, 110-123 for losing/winning) compared to a women's cohort from the Rugby World Cup (range of 73-98 carries).(3) Using these data, coaches can understand the match demands in terms of match events occurring and the need to condition their players appropriately.

Part1: Win-Lose predictors

Using the descriptives of the game, the predictors of winning and losing were identified. For the SARU Women's Premier Division, four performance indicators were identified as predictive of winning / losing (whilst another two improved the model but were not significant). Tries (>6x increased odds), dominant tackles % (45% increased odds), goal kicking success % (3% increased success), and a yellow card (73% decreased odds) were the predictive variables associated with winning vs. losing in this study. Whilst tries were the largest contributor, dominant tackles % is the biggest confounder for this, and is the driving force for the increase in OR to >6x compared to if you only have tries in the model (OR=3.46). The only data available for team success in women's rugby emanates from the sevens format, where restarts, passing effectiveness, line breaks (none of which were in our dataset), tries, and missed tackles were associated with winning/losing.(10) Other variables were also identified, however, they were variables not available in our dataset either and more specific to sevens.(10) In men's rugby (including rugby league), tries, successful conversions, and penalty goals have also been shown to be predictive of winning.(11, 12) As these match activities immediately result in points, this finding is not unexpected. In our women's focused data set, however, some other variables that differed include turnovers won, lineouts won/lost,(13), and scrum penalties.(11, 12, 14) Interestingly, whilst lineouts won % was not significant, it was a variable that added to the predictive ability of the model (an increased odds of 3% for each increase in %), which does match that of the men's data. The variations in tackle coding could account for some of the deviations, so in our data, tackles made % was predictive, whereas in other studies, tackles completed % was predictive.(12) Even though differences exist, there are some similarities between other studies and ours, but each competition and gender should be looking at their own data and attempting to gain an advantage based on the type of game played in their region. Coaches need to allocate more training time to tackle technique (ensuring a decrease in missed tackles), and goal kicking. These indicators would most likely also be improved with a focus on fitness (seeing as these teams perform >140 tackles in a match, fitness will be key to ensuring tackles are not missed and executed well). The level of play should again be considered, as amateur teams will be more likely to miss tackles, and the higher level of play/player will assist in executing skills. It should be noted that during these two seasons, the standard of play across the various teams was similar allowing for a homogenous sample,

and to our knowledge this is the largest sample of matches used in a model in women's rugby. Tactical aspects, such as how to bring a player to ground, whether or not the team wants to drive the player back or bring to ground, will also affect dominant tackle numbers and potentially missing tackles. Further information regarding tackle technique should be incorporated into future studies to determine what predictive ability it has on winning / losing. Another aspect to consider would be to include injuries occurring and how that affects subsequent ability to win or lose a match.

Part 2: Points difference analysis

While winning and losing is key to match performance, in many competitions, points difference can be the difference between making the playoffs or not. The bonus points allocated for winning by large margins or losing by smaller margins can affect teams in knock-out tournaments. The only cluster factor that was predictive of a larger points difference (>25 points difference), included clean breaks, tries and the conversion thereof. Intuitively, tries and conversions translate into immediate points, and therefore it makes sense that the points difference would increase. On the other hand, positive tackles and lineout wins, were predictive of a close match ($\leq$8 points difference). In previous studies in men's rugby, lineouts stolen, tackles completed % have been shown to contribute to points against, and line breaks contribute to points for.(12) In rugby league, meters gained and line breaks were associated with increased points difference.(15) It is understandable that conversions led to an increase, as they are only available after a try of 5 points has been scored, similar to goal kicking success % where if you are successful, it will result in points.

Given this finding, coaches should attempt to incorporate broken play in their drills, in an attempt to train players how to capitalize on errors (whether theirs or the opposition). Another interesting finding was that scrums (regardless of winning or losing it), increase points for the team. From these data, set pieces and tackling are important aspects coaches need to prioritise. These variables are important to consider in the context of the matches and competitions they are played in. The results in Part 2 support the findings of Part 1, and should be read collectively when attempting to implement strategies for winning, and winning with a large points difference.

Implementation

The findings of this paper should assist coaches and players in understanding aspects of the game that are key to performance. Tackle dominance was an important factor for winning and losing, and emphasizes the need for effective tackling in the women's game. Another factor for coaches to focus on, was goal kicking success, whilst intuitively this would be important, coaches should spend extra time with their kickers. The final factor for consideration are yellow cards, discipline is always going to be an important part of the game, and using the information on how this effects your chances of winning should be communicated effectively to players. Using the points difference outcome, the role of set pieces is prominent. The effective coaching of lineouts in the women's game will assist in team success. Again, the tackle was an important aspect of the final score, correct tackle technique, usually resulting in effective and dominant tackles should be a top priority for coaching staff, both for performance and player safety.

Strengths and Limitations

This is one of the first studies to describe the performance indicators in women's rugby union and the predictors of winning and losing. Hopefully, with multiple seasons data and adjusting for season, and updating databases and including rule changes, predictors may change and coaches should be informed as to how this changes predictive variables. We hope to replicate the study and determine if there are potentially new predictors or if predictors changed over the years. Future work, when more data are available, should also include a possible Structural Equations Model modeling the measurement model for the factors simultaneously with the structural part predicting the the points difference outcome. A further limitation includes that the current points difference modeling (Part 2) did not accommodate a cluster effect. It should be noted that the coding was performed by a commercial entity and not according to the consensus statement research standards, and future studies that have the resources to code matches should use the consensus standards (but we acknowledge using publicly available coding/match statistics are more easily available and allow for all teams to access these data and analyse without an in-house video analysis team).(7)

Conclusion

In summary, numerous performance indicators were found to be predictive of winning/losing and associated with points difference in the matches. For winning compared to losing, tries, goal kicking success %, dominant tackles %, increased the odds of winning, while yellow cards decreased the odds. With regards to points difference, clean breaks, tries and conversions were predictive of unbalanced matches (larger points difference), and positive tackles and lineout wins were predictive of close matches (smaller points difference).

These performance indicators varied from those previously shown in men's and the limited data available in women's rugby. More data are needed to validate the statistical models (and try new models) to replicate the study in other competitions of varying levels and geographical locations.

Coaches should attempt to add additional focus on tackling (dominance aspects, better technique to avoid missed tackles), and set pieces. Furthermore, attacking pieces in the form of clean breaks were associated with increased points.

References

1. Clarke AC, Anson J, Pyne D. Physiologically based GPS speed zones for evaluating running demands in Women's Rugby Sevens. J Sports Sci. 2015;33(11):1101-8.
2. Suarez-Arrones L, Portillo J, Pareja-Blanco F, Sáez de Villareal E, Sánchez-Medina L, Munguía-Izquierdo D. Match-play Activity Profile in Elite Women's Rugby Union Players. J Strength Condition Res. 2014;28(2):452-8.
3. Hughes A, Barnes A, Churchill SM, Stone JA. Performance indicators that discriminate winning and losing in elite men's and women's Rugby Union. Int J PerformAnalys Sport. 2017;17(4):534-44.
4. Heyward O, Emmonds S, Roe G, Scantlebury S, Stokes K, Jones B. Applied sports science and sports medicine in women's rugby: systematic scoping review and Delphi study to establish future research priorities. BMJ Open Sport Exer Med. 2022;8(3):e001287.
5. West SW, Shill IJ, Clermont C, Pavlovic N, Cairns J, Seselja B, et al. Same name, same game, but is it different? An investigation of female rugby union match events in Canadian Varsity players. Int J Sports Sci Coach. 2022;17(5):1119-27.
6. Paul L, Isaacs N, Naidoo D, Parker N, Cantwell L, Hendricks S. Women's rugby in the South African context: challenges and opportunities. Br J Sports Med. 2023:bjsports-2023-107019.
7. Hendricks S, Till K, Hollander Sd, Savage TN, Roberts SP, Tierney G, et al. Consensus on a video analysis framework of descriptors and definitions by the Rugby Union Video Analysis Consensus group. Br J Sports Med. 2020;54(10):566-72.
8. Alcantara IM, Naranjo J, Lang Y. Model selection using PRESS statistic. Computational Statistics. 2023;38(1):285-98.
9. Harrell FE. Regression modeling strategies. Bios. 2017;330(2018):14.
10. Barkell FJ, O'connor D, Cotton GW. Characteristics of winning men's and women's sevens rugby teams throughout the knockout Cup stages of international tournaments. Int J Perform Analys Sport. 2016;16(2):633-51.
11. Ortega E, Villarejo D, Palao JM. Differences in game statistics between winning and losing rugby teams in the six nations tournament. J Sports Sci Med. 2009;8(4):523-7.
12. Schoeman R, Schall R. Team performance indicators as predictors of final log position and team success in Aviva Premiership, Guinness Pro 14, French Top 14 and Super Rugby. Int J Perform Analys Sport. 2019;19(5):763-77.
13. Bennett M, Bezodis NE, Shearer DA, Kilduff LP. Predicting performance at the group-phase and knockout-phase of the 2015 Rugby World Cup. Euro J Sport Sci. 2021;21(3):312-20.
14. Scott GA, Bezodis N, Waldron M, Bennett M, Church S, Kilduff LP, et al. Performance indicators associated with match outcome within the United Rugby Championship. J Sci Med Sport. 2023;26(1):63-8.
15. Parmar N, James N, Hughes M, Jones H, Hearne G. Team performance indicators that predict match outcome and points difference in professional rugby league. Int J Perform Analys Sport. 2017;17(6):1044-56.