Explaining match outcome and ladder position in the National Rugby League using team performance indicators

Explaining match outcome and ladder position in the National Rugby League using team performance indicators

Woods T. Carl[a*]; Sinclair Wade[a]; Robertson Sam[b]

[a]Discipline of Sport and Exercise Science, James Cook University, Queensland, Australia

[b]Institute of Sport, Exercise & Active Living (ISEAL), Victoria University, Melbourne, Australia

*Corresponding Author

Carl Woods, Discipline of Sport and Exercise Science, James Cook University, Townsville, Queensland, Australia.

Ph: +61 08 4781 6550 Mob: +61 421254329 Email: carl.woods@jcu.edu.au

Abstract

Objectives: To examine the extent at which match outcome and ladder position could be explained using team performance indicators in the National Rugby League (NRL).

Methods: The dataset consisted of 13 performance indicators acquired from each NRL team across the 2016 season (*n* = 376 observations). Data was sorted according to *apriori* match outcome (win/loss) and ladder position (ranked one to 16). Given the binary and categorical nature of the response variables, two analysis approaches were used; a conditional interference classification tree and ordinal regression.

Results: Five performance indicators (‘try assists’, ‘all run meters’, ‘offloads’, ‘line breaks’ and ‘dummy half runs’) were retained within the classification tree, detecting 66% of the losses and 91% of the wins. A significant negative relationship was noted between ladder position and ‘kick metres’ (β (SE) = -0.002 (<0.001); 95% CI = -0.003 – <-0.001) and ‘dummy half runs’ (β (SE) = -0.017 (<0.012); 95% CI = -0.041 – 0.006), while a significant positive relationship was noted for ‘missed tackles’ (β (SE) = 0.019 (0.006); 95% CI = 0.006 – 0.032).

Conclusions: A unique combination of primarily attacking performance indicators provided the greatest explanation of match outcome and ladder position in the NRL. These results could be used by NRL coaches and analysts as a basis for the development of practice conditions and game strategies that may increase their teams’ likelihood of success. Beyond rugby league, this study presents analytical techniques that could be applied to other sports when examining the relationships between performance indicators and match derivatives.

***Keywords**:* performance analysis; classification tree; ordinal regression; team sport

1. Introduction

Similar to many team invasion sports, rugby league requires players to blend a range of multidimensional performance qualities (i.e., physical, technical, and perceptual skills).[1] The premier competition in Australia and New Zealand is the National Rugby League (NRL), which currently consists of 16 teams who compete within a 26-round premiership season. The premiership season serves as a means to rank teams against one another on a 'ladder', where one is the highest rank and 16 is the lowest rank. Within the premiership season, teams are awarded two points for a win and one point each for a drawn match. At its conclusion, the eight highest ranked teams (colloquially termed 'the top eight') then compete against one another in a four week knock out finals series as they attempt to obtain an NRL premiership. Additionally, the top two teams on the ladder at the conclusion of the premiership season are awarded a double chance and home ground advantage within the finals series. Subsequently, the primary objective for each NRL team in the premiership season is to accumulate as many wins as possible, ultimately ranking high on the ladder (i.e., closer to one).

Given the importance of possessing a high number of wins during the premiership season, some research has examined the physical and technical activity profiles of successful and less successful elite (i.e., NRL) and sub-elite (i.e., state-league) rugby league teams.[2-4] Most recently, Kempton et al.[4] investigated the physical and technical profiles of 'successful' (defined via fourth placing on the NRL ladder) and 'less successful' (defined via 16th placing on the ladder) NRL teams. Using linear mixed effects modelling and magnitude based inferences, the authors showed that the successful team recorded lower high-speed running distances (882 (749 – 1014 m)) and engaged in fewer physical collisions (18.6 (16.4 – 20.8)) relative to their less successful counterparts (904 (807 – 1001 m) and 22.2 (20.6 – 23.85), respectively).[4] Additionally, the successful team gained more territory in attack, obtained a greater percentage of ball possession and performed fewer errors when compared to the less successful team.[4] Similar findings have been reported in sub-elite rugby league competitions, with Gabbett[2] noting that the top four state-league teams (out of 12) gained more territory in attack and conceded fewer metres in defence when compared to the bottom four teams. Combined, this work indicates that there are distinctive differences between successful and less successful elite and sub-elite rugby league teams manifested via their technical performance indicator characteristics.

While of value, previous work has not yet investigated the team performance indicator characteristics of higher and lower ranked teams from a national, league-wide perspective (i.e., the entire NRL). This is despite work being conducted in other team[5] and individual[6] sports that adopt ladder systems to rank performances noting differences in performance indicator characteristics between higher and lower ranked teams or players. In part, this may be due to the perceived difficulties encountered when modelling a sequential or ordinal response variable (i.e., ladder position) against a set of defined explanatory variables (i.e., team performance indicators). Ordinal regression may be of use in such instances, as it allows the prediction of an *apriori* response variable whose properties exist in some form of meaningful order or sequence.[7] Further, the extent at which team performance indicators can explain match outcome (win/loss) is yet to be comprehensively investigated in the NRL. This is in contrast to Australian football, where Robertson et al.[8] used decision-tree analysis and logistic regression to identify the performance indicators most associated with match outcome in the Australian Football League (AFL).

Revealing unique combinations of performance indicator characteristics explanatory of higher and lower performing teams may objectively assist coaches with decisions orienting game and training plans or team selection strategies.[8] To this end, the aim of the current study was to examine the extent at which team performance indicators could be used to explain match outcome and ladder position at the conclusion of the NRL premiership season.

2. Methods

Team performance indicators from the 2016 NRL season were extracted from a publically accessible source (http://www.nrl.com/stats) and placed into a custom designed Microsoft Excel spreadsheet (Microsoft, Redmond, USA) for analysis (Supplementary Table 1). These team performance indicators were chosen owing to their use elsewhere,[4] enabling a discussion of findings relative to the performance analysis literature in rugby league. The dataset contained 14 team performance indicators acquired from 26-rounds, equating to 376 observations. Eight games were played per round, with the exception of rounds 12 (four games), 13 (seven games), 15 (four games), 16 (seven games), 18 (four games) and 19 (six games). In these rounds, 'byes' were mandated for certain teams. Drawn matches (n = 2) were excluded from the dataset as the competing teams were awarded equal points for these matches. Ethical declaration was obtained by the relevant Human Research Ethics Committee prior to analysis.

Data were sorted according to match outcome (two levels: win, loss) and ladder position (16 levels) at the conclusion of the 2016 premiership season. Here, a ladder position closer or equal to 'one' was indicative of a higher ranked team, while a ladder positioning closer or equal to '16' was indicative of a lower ranked team.

Prior to the main analyses, a correlation matrix was built to assess the level of collinearity between the team performance indicators (explanatory variables).[9] Descriptive statistics (mean ± standard deviation) were then calculated for each team performance indictor relative to match outcome (win, loss). The effect size and 90% confidence interval of match outcome was calculated using Cohen's d statistic,[10] where an effect size of <0.2 was considered trivial, 0.2 – 0.6 small, 0.6 – 1.2 moderate, 1.2-2.0 large, and >2.0 very large.[11]

To examine the extent at which team performance indicators could be used to explain match outcome, a conditional interference (CI) classification tree was grown using the 'party' package[12] in the *R* computing environment.[13] A CI classification tree is a type of decision induction analysis that estimates a regressive relationship through binary partitioning (splitting) by testing the null hypothesis between a set of explanatory variables and a binary response variable.[12] Here, the team performance indicators were coded as the explanatory variables and match outcome was coded as the binary response variable (1=win, 0=loss). Partitioning is stopped when the null hypothesis cannot be rejected (i.e., $P \geq 0.05$). A benefit of this analysis is that its' fitting algorithm corrects for multiple testing, thus avoiding overfitting.[12] As such, this analysis results in the growth of an unbiased decision tree that does not require pruning.[12]

The relationship between ladder position and team performance indicator characteristics were examined using cumulative link mixed models fitted to the data using the 'ordinal' package[14] in the same computing environment.[14] This type of modelling was chosen as it is a form of ordered regression, useful when modelling a response variable that possesses some type of order or sequence.[7] Confidence intervals of each models' parameter estimates were calculated using the *confint* function, with the 'P values' being estimated using Wald's method.

3. Results

Collinearity was noted between 'all runs', 'all run metres', and 'possession percentage' (r >0.5), with the former being excluded from further analysis. As shown in Table 1, 'try assists', 'line breaks' and 'all run metres' expressed the largest relative effect on match outcome.

******INSERT TABLE ONE ABOUT HERE******

Of the 376 observations, the CI classification tree successfully classified 124 of the 188 recorded losses (66%) and 171 of the 188 recorded wins (91%). Five of the 13 modelled team performance indicators were included within the CI tree (Figure 1), these being 'try assists' (root node), 'all run metres', 'line breaks', 'dummy half runs', and 'offloads'. Nine terminal nodes were grown; numbers 4, 5, 8, 9, 10, 12, 15, 16, and 17.

******INSERT FIGURE ONE ABOUT HERE******

Following the branches to the right of the root node (>2 try assists), node number 11 partitioned the data on 'all run metres' at a count of 1340m. Of the 24 observations in terminal node 12, the probability of winning was lower (25%) than the probability of losing (75%). Node number 13 partitioned the data on 'try assists' at a count of 4. Of the 60 observations in terminal node number 17, the probability of winning was higher (98.3%) than the probability of losing (1.7%). This combination of performance indicators provided the greatest probability of winning. A count of ≤4 'try assists' partitioned the data based on 'offloads' (node number 14); branching into terminal nodes 15 and 19 at a count of 9.

Following the branches to the left of the tree (≤2 try assists), node number 2 partitioned the data on 'all run metres' at a count of 1450m. An accumulation of ≤1450 'all run metres' was then partitioned based on 'line breaks' at a count of 4 (node 3). Of the 80 observations in terminal node 4, the probability of losing was higher (95%) than the probability of winning (5%). This combination of performance indicators provided the lowest probability of winning. The probability of winning was slightly increased if a team accrued >4 'line breaks' (terminal node 5).

From the 13 team performance indicators modelled, three expressed a significant relationship with ladder position (Table 2). A significant positive relationship was observed between 'missed tackles' and ladder position (β (SE) = 0.019 (0.006); 95% CI = 0.006 – 0.032), with the count of this indicator generally increasing as ladder position moved further away from one. Additionally, a significant negative relationship was observed between 'kick metres' and ladder position (β (SE) = -0.002 (<0.007); 95% CI = -0.003 – <-0.001) and 'dummy half runs' and ladder position (β (SE) = -0.017 (<0.012); 95% CI = -0.041 – 0.006), with the count of these indicators generally decreasing as ladder position moved further away from one. Comparatively, the remaining 10 team performance indicators were unable to meaningfully explain ladder position.

******INSERT TABLE TWO ABOUT HERE******

4. Discussion

This study examined the extent at which team performance indicators could explain match outcome and ladder position at the conclusion of the 2016 NRL premiership season. Results demonstrated that five performance indicators successfully explained match outcome; classifying 66% of the losses and 91% of the wins. It was the unique combination of 'try assists' and 'all run metres' that provided the greatest probability of winning, while a unique combination of 'try assists', 'all run metres' and 'line breaks' provided the lowest probability of winning. Further, 'missed tackles' expressed a significant positive relationship with ladder position, while 'kick metres' and 'dummy half runs' expressed a significant negative relationship with ladder position, both generally decreasing as ladder position moved further away from one. These results yield practical applications for coaches and performance analysts in the NRL by offering an objective framework of use in the design of practice conditions and game strategies.

The identification of attacking performance indicators in the explanation of match outcome complements the observations of Kempton et al.,[4] who noted that a successful rugby league teams gained more territory in attack. Coupled, these results show that higher functioning (e.g. winning) NRL teams possess superior attacking strategies manifested via 'all run metres', 'try assists', 'offloads' and 'dummy half runs' while perhaps being better equipped at maintaining a greater percentage of ball possession relative to their lower functioning (e.g. losing) counterparts.[14] Practically, coaches could use this information to design game strategies that focus upon the efficiency of their teams attacking phases. For example, attacking strategies oriented on 'offloads' and 'dummy half runs' may stretch an opponent's defensive line and incur 'line breaks', resulting in the accumulation of a large 'all run' meterage and 'try assist' count; the amalgamation of all may increase a team's probability of winning.

The rules of rugby league game-play dictate that a team can only accrue six ‘tackles’ while in possession of the ball until it is relinquished to the opposition. To counteract this ruling and gain more territory in attack, teams ‘kick’ the ball toward their opposition’s goal line on their sixth tackle. It was of interest to note that higher ranked teams appeared to kick the ball further than their lower ranked counterparts; suggesting that they may possess rostered players with more pronounced kicking skills. This finding has also been reported in sub-elite rugby league competitions, with Gabbett[2] noting that the top four teams gained more territory with their kicks relative to the bottom four teams. Although relinquishing ball possession, longer kicks push an opposition closer to their goal line. This is an important consideration, as Kempton et al.[15] noted that ball possession closer to an opponent’s goal line (within 20 m) was likely to increase the likelihood of scoring a try. Given this, it would be of value for future work to examine the placement of kicks performed during game-play, as this may offer a deeper insight into the explicit offensive strategies successful teams implement to optimise their likelihood of scoring.

A unique finding of this study was the identification of the positive relationship between ‘missed tackles’ and ladder position. This suggests that higher ranked NRL teams possess more comprehensive defensive strategies when compared to their lower ranked counterparts. A potential strategy higher ranking teams employ while in defence is to tackle in pairs or groups; colloquially referred to as ‘gang tackling’.[16] This strategy is likely to counteract the traditional attacking strategy of ‘charging’ (i.e., running directly at a defensive line to physically barge through) by negating potential physical discrepancies between an attacking and defending player. This type of defensive football would likely require collective team behaviour, with players needing to spread at speed following the tackle given the potential holes ‘gang tackling’ may incur along a defensive line. Although most teams are likely to engage in this strategy, higher ranked teams may be more equipped at performing this efficiently given the reduction in ‘missed tackles’ noted in this study. Concomitantly, our results showed that higher ranked teams accrued a greater count of ‘dummy half runs’; an attacking strategy commonly employed against an unstructured defence.[16] Thus, higher ranked teams may not only spread at speed following a 'gang tackle' but they appear more equipped at identifying and exacerbating holes in an opponent's defensive line when employing the same defensive tactic. The reduction in missed tackles recorded by higher ranked teams may also be explained by physiological differences. Gabbett[17] reported a negative augmentative relationship between tackling technique and physiological fitness in rugby league. Translated to game play, it is possible that a decline in tackling technique would increase the number of 'missed tackles'.[17] Given this, NRL coaches should ensure rostered players possess the collective capabilities required to 'gang tackle', as well as possessing the required physiological characteristics to negate the influence of fatigue on tackling technique.

Despite providing an objective framework for the development of practice conditions and game strategies, it is important to note that lower performing teams may not possess players capable of accruing a high count of 'try assists', 'all run metres', 'offloads', 'kick metres' and/or 'dummy half runs'. Accordingly, these results yield implications for the identification and recruitment of talent into the NRL. Recruitment managers of lower ranked teams may actively seek juniors who possess superior attacking attributes, while concomitantly looking to poach players from other NRL teams who possess the aforementioned skill qualities during offseason trade periods. The addition of such players may afford a coach with the 'tools' needed to build a competitive game strategy, optimising their likelihood of success. Given these recommendations, the utility of skill tests that measure the aforementioned technical qualities should continually be promoted with the rugby league talent pathway, as their integration may assist with the identification of prospective NRL players.[18]

Despite the practical utility of this work, it is not without limitations. Firstly, the playing draw in the NRL is not equal, indicating that certain teams will potentially play each other more than once. It is therefore possible that higher ranked teams incidentally play lower ranked teams multiple times throughout the premiership season, leading to a competitive advantage manifested in their team performance indicator characteristics. Secondly, this study did not account for locational or environmental factors that could potentially influence match outcome. Previous work has shown that situational and contextual factors such as match location and environmental conditions could either positively or negatively impact on a team's perceptual,[19] technical and physical performance output.[20] Thus, future work may wish to consider such factors when investigating the performance indicator characteristics of higher and lower performing NRL teams. Lastly, despite being data widely used by NRL teams, we have to assume face reliability of the notional analyses, as work is yet to establish test-retest reliability. Additionally, the performance indicators presented in this study are discrete in nature and do not elucidate chains or sequences of play. For example, the effectiveness of a kick may be driven by the team's position on their fifth tackle, the speed of the 'play the ball' or the quality of the dummy half pass. Accordingly, future work may consider devising indicators that combine multiple actions or conversely investigate the chain of play that led to a specific action, as both may offer deeper insights into the unique profiles of higher and lower performing NRL teams. Nonetheless, this study presents a unique insight into the technical profiles of higher and lower functioning NRL teams, offering an enticing platform for which future work can progress.

5. Conclusion

This study demonstrates that higher performing NRL teams in 2016 premiership season generated distinctive performance indicator characteristics when compared to their lower performing counterparts. It was a combination of 'try assists', 'all run metres', 'line breaks', 'dummy half runs' and 'offloads' that provided the greatest explanation of match outcome, while 'missed tackles', 'kick metres' and 'dummy half runs' expressed significant relationships with ladder position.

6. Practical Applications

- NRL coaches may look to develop practice conditions and game strategies that afford a high count of 'try assists', 'all run metres', 'line breaks', 'dummy half runs' and 'offloads' to increase their likelihood of achieving a successful match outcome.
- Devising defensive strategies that minimise 'missed tackle' counts may assist with a higher ladder positioning for an NRL team.
- NRL recruitment mangers may utilise these results to identify suitable players who possess the requisite skill sets to assist with team success.

7. Acknowledgements

The authors would like to acknowledge the many analysts who coded the data over the course of the 2016 NRL season.

8. References

1. Sirotic AC, Coutts AJ, Knowles H, et al. A comparison of match demands between elite and semi-elite rugby league competition. *J Sports Sci* 2009; 27(3):203-211
2. Gabbett TJ. Effects of physical, technical, and tactical factors on final ladder position in semiprofessional rugby league. *Int J Sports Physiol Perform* 2014; 9(4):680-688
3. Hulin BT, Gabbett TJ, Kearney S, et al. Physical demands of match play in successful and less-successful elite rugby league teams. *Int J Sports Physiol Perform* 2015; 10(6):703-710
4. Kempton T, Sirotic AC, Coutts AJ. A comparison of physical and technical performance profiles between successful and less-succesful professional rugby league teams. *Int J Sports Physiol Perform* 2016; 6(1):1-19
5. Woods CT. The use of team performance indicator characteristics to explain ladder position at the conclusion of the Australian Football League home and away season. *Int J Perf Analysis Sport* 2016; 16(3):837-847
6. Clarke SR. Rating non-elite tennis players using team doubles competition results. *J Operational Res Soc* 2011; 62(7):1385-1390

7. Harrell F. *Regression Modelling Strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*, Switzerland, Springer International Publishing, 2015
8. Robertson S, Back N, Bartlett JD. Explaining match outcome in elite Australian Rules football using team performance indicators. *J Sports Sci* 2016; 34(7):637-644
9. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012; 24(3):69-71.
10. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155-159.
11. Hopkins WG. A new view of statistics. Sportscience. http://www.sportsci.org/resource/stats
12. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: A conditional inference framework. *J Comput Graph Stat*. 2006;15(3):651-674
13. R Core Team. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria
14. Christensen RHB. Package ‘ordinal’: regression models of ordinal data. Retrieved from https://cran.r-project.org/web/packages/ordinal/ordinal
15. Kempton T, Kennedy N, Coutts AJ. The expected value of possession in professional rugby league match-play. *J Sports Sci* 2015; 34(7):1-6
16. Gabbett TJ, & Abernethy B. Dual-task assessment of a sporting skill: influence of task complexity and relationship with competitive performances. *J Sports Sci* 2012; 30(16):1735-1745
17. Gabbett TJ. Influence of fatigue on tackling technique in rugby league players. *J Strength Cond Res* 2008; 22(2):625-632
18. Hendricks Sm Lambert M, Masimla H, et al. Measuring skill in rugby union and rugby league as part of the standard team testing battery. *Int J Sports Sci Coach* 2015; 10(5):949-965

19. McGuckin TA, Sinclair WH, Sealey RM, et al. Players' perceptions of home advantage in the Australian rugby league competition. *Percept and Motor Skills* 2015; 121(3):666-674.
20. Kempton T, Coutts AJ. Factors affecting exercise intensity in professional rugby league match-play. *J Sci Med Sport* 2016; 19(6):504-508

Figure 1. The CI classification tree illustrating the probability of wins and losses in the NRL. *Note*: 'n' denotes the number of observations in each node. The first y value denotes the probability of losing and the second y value denotes the probability of winning (e.g. 0.7 = 70%).

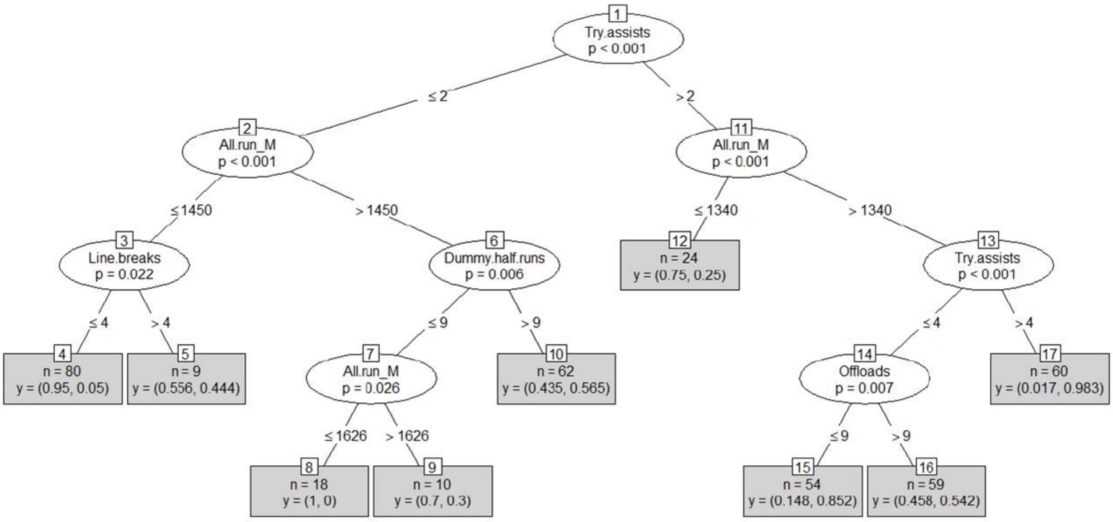

Table 1. Descriptive and effect size statistics relative to match outcome

Performance indicator	**Wins**	**Losses**	***d* (90% CI)**	**Size**
All run metres	1620.2 ± 190.7	1422.2 ± 209.8	0.98 (0.80 – 1.16)	‘Moderate’
Line breaks	5.1 ± 2.4	2.9 ± 1.7	1.04 (0.86 – 1.22)	‘Moderate’
Try Assists	3.7 ± 1.8	1.8 ± 1.1	1.25 (1.06 – 1.43)	‘Large’
Offloads	10.0 ± 4.3	10.5 ± 4.6	0.11 (-0.05 – 0.28)	‘Trivial’
Tackles	310.7 ± 38.2	338.9 ± 36.0	0.75 (0.58 – 0.93)	‘Moderate’
Missed tackles	24.8 ± 7.5	30.4 ± 8.4	0.69 (0.51 – 0.86)	‘Moderate’
Errors	8.6 ± 2.6	9.9 ± 2.7	0.48 (0.31 – 0.66)	‘Small’
Total kicks	19.4 ± 4.0	18.5 ± 3.6	0.25 (0.08 – 0.42)	‘Small’
Kick metres	475.7 ± 121.6	427.8 ± 111.3	0.40 (0.23 – 0.58)	‘Small’
Dummy half runs	11.9 ± 4.2	10.1 ± 4.9	0.41 (0.24 – 0.58)	‘Small’
Possession percentage	52.3 ± 0.3	48.1 ± 0.3	0.93 (0.75 – 1.10)	‘Moderate’
Tackle breaks	30.4 ± 8.3	24.7 ± 7.5	0.70 (0.53 – 0.88)	‘Moderate’
Penalties conceded	6.9 ± 2.2	6.7 ± 2.8	0.05 (-0.11 – 0.22)	‘Trivial’

Table 2. Parameter estimates of the cumulative link mixed models fitted to ladder position

Performance indicator	**Estimate**	**SE**	**LCI**	**UCI**	***P***
All run metres	<-0.001	<0.001	-0.001	<0.001	0.214
Line breaks	-0.071	0.037	-0.144	0.001	0.054
Try assists	0.007	0.045	-0.081	0.095	0.876
Offloads	0.014	0.012	-0.009	0.038	0.240
Tackles	0.001	0.001	-0.001	0.004	0.311
Missed tackles*	0.019	0.006	0.006	0.032	0.002
Errors	0.214	0.021	-0.020	0.603	0.321
Total kicks	0.050	0.026	<-0.001	0.101	0.054
Kick metres*	-0.002	<0.001	-0.003	<-0.001	0.002
Dummy half runs*	-0.017	0.012	-0.041	0.006	0.014
Possession percentage	-2.047	1.452	-4.893	0.798	0.158
Tackle breaks	0.007	0.007	-0.007	0.022	0.319
Penalties conceded	0.004	0.023	-0.041	0.049	0.860

Note: 'Estimate' denotes the beta coefficient estimate; 'SE' denotes the standard error of the coefficient; 'LCI' denotes the lower 95% confidence interval of the estimate; 'UCI' denotes the upper 95% confidence interval of the estimate; * denotes significance (P <0.05).