How to cite this thesis / dissertation (APA referencing method):

Surname, Initial(s). (Date). *Title of doctoral thesis* (Doctoral thesis). Retrieved from http://scholar.ufs.ac.za/*rest of thesis URL on KovsieScholar*

Surname, Initial(s). (Date). *Title of master's dissertation* (Master's dissertation). Retrieved from http://scholar.ufs.ac.za/*rest of thesis URL on KovsieScholar*

Match statistics that discriminate between winning and losing teams in ODI and T20I cricket

by

MARK CHRISTOPHER SCHAEFER

Dissertation submitted in fulfillment of the requirements in respect of the degree

MAGISTER ARTIUM IN HUMAN MOVEMNET SCIENCE

In the department
EXERCISE AND SPORT SCIENCES
In the
SCHOOL OF ALLIED HEALTH SCIENCES
FACULTY OF HEALTH SCIENCES
at the

UNIVERSITY OF THE FREE STATE
BLOEMFONTEIN
31 January 2018

Study Leader: Dr. R. Schoeman

Declaration

I, Mark Schaefer, hereby declare that the work on which this manuscript is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted to any other journal.

Furthermore, the co-authors of the articles in this dissertation, Dr. Riaan Schoeman (supervisor) and Prof. Robert Schall (statistician) hereby give permission to the candidate, Mr. Mark Schaefer to include the articles as part of a Master's dissertation. The contribution (advisory and supportive) of these co-authors was kept within reasonable limits, thereby enabling the candidate to submit this dissertation for examination purposes.

Signed on this 8th day of January 2018

Dr. Riaan Schoeman
Supervisor

Mark Schaefer

Acknowledgments

- The authors thank the faculty of Health Sciences and University of the Free State, for technical and editorial preparation of the manuscript.
- Thanks to my family for their support.
- Thanks to my wife Maxine for all her guidance and motivation.
- Thanks to Dr. Riaan Schoeman and Prof. Robert Schall for their advice and input throughout the writing and planning process.

Summary

Match statistics that discriminate between winning and losing teams in ODI and T20I cricket

Background

Cricket players and teams have a different strategy for batting for the different formats of cricket, namely Twenty-Twenty International (T20I) and One Day International (ODI). Different application of skills is required for each format of cricket can clearly be seen as mostly a different team is selected for each format of the game in professional cricket. Analysis of performance variables such as boundaries hit by batsmen and runs scored during the power play can be used to predict future success or failure of a cricket team based on the match outcome. This study will provide batting statistics that discriminate between winning and losing teams in ODI and T20I cricket. Furthermore, the study will reveal which variables correlate the highest with successful performance within the different formats of the game.

Aims

The aim of this study was twofold, firstly to analyse batting data in ODI cricket that discriminate between winning and losing teams. Secondly to analyse batting data in T20I cricket that discriminate between winning and losing teams.

Method

Sample

Ten international teams were selected for the purpose of this study. The ten teams were selected because they all participate in all three formats of cricket namely ODI, T20I, and test cricket. Six matches from each team's records were randomly selected and observed (3 batting first, 3 batting second). The first aim consisted of conducting analysis of a total of 60 professional ODI cricket matches resulting in 120 records (innings) (both teams involved per match). The second aim consisted of conducting analysis of a total of 60 professional T20I cricket matches resulting in 120

records (both teams involved per match). Drawn matches, and those which employed the Duckworth-Lewis method, were excluded from the study.

Measuring instruments

Retrospective data from the 2014 and 2015 international cricket season was collected from ESPN Cricinfo website.

Data analysis

In this research, a strong and reliable data source is needed which was found in Statsguru. Statsguru is ESPN Cricinfo's cricket statistics maintenance database. The data was then analyzed using the SAS statistical software (SAS, 2013).

Because of the fundamentally different match situation faced by the team batting first and second, respectively, the data were analysed separately for the team batting first and for the team batting second. The outcome of the match is a binary variable (win/lose) since drawn matches were excluded from the analysis. The association of the potential predictor variables with the match outcome was analyzed using univariate logistic regression, fitting each predictor variable, one at a time. The statistical significance of each predictor variable was tested using an exact test (exact conditional logistic regression); the exact P-value is reported. The analysis was carried out using SAS procedure LOGISTIC (see SAS, 2013).

Results

For aim 1 the significant predictors of winning an ODI cricket match when batting first were: runs scored in the first 20 overs (p=0.0019), runs scored in the last 12 overs (p=0.0004), sixes scored (p=0.0017), and the number of runs scored among the top four batsmen (p=0.0015); For aim 1 the significant predictors of winning an ODI cricket match when batting second were: fours scored (p=0.0024), sixes scored (p=0.00277), runs scored between the top order batsmen (p=0.0197), and runs scored between the lower order batsmen (p=0.0222). Variables that predict success in ODI cricket differed for teams batting first and second, respectively. For aim 2 significant predictors of winning a T20I cricket match when batting first were: runs scored in the first 5 overs (p=0.0035), runs scored in the last 7 overs (p=<0.0001),

and sixes scored (p=0.0081); similarly, significant predictors for winning a T20I cricket match when batting second were: Runs scored in the first 5 overs (p=0.0046) fours scored (p=0.0258), runs scored between the top order batsmen (p=0.0034), and runs scored between the lower order batsmen (p=0.0043).

Conclusions

For both aim 1 and 2 data showed that scoring runs in the initial and end part of the innings (first 20 and last 12 overs of an ODI match; first 5 and last 7 of a T20I match), the number of fours and sixes scored, and the number of runs scored between the different batsmen are significantly related to winning a cricket match. The results of this study show that there are variables in cricket that relate positively to success in ODI cricket and can be used as success predictors. These success predictors differ between teams batting first and second.

Keywords

Batting, Cricket One-day International, Twenty-Twenty International, Runs scored, Boundaries scored

References

SAS Institute Inc. *SAS/STAT 9.2 User's Guide, 2nd ed*. Cary, NC: SAS Institute Inc.

Table of contents

List of tables

List of figures

List of Abbreviations

BFL	Batting First Losing
BFW	Batting First Winning
BSL	Batting Second Losing
BSW	Batting Second Winning
ICC	International Cricket Council
ODI	One-Day International
T20I	Twenty-Twenty International

Chapter 1: Introduction

1.1 Introduction

Cricket is a sport loved by many people and spectated by a wide variety of different nationalities (Chadwick & Aurthur, 2010). Cricket as a popular viewing sport can be broken down into variables that influence performance in batting across the various formats of cricket. Cricket boasts a long and far-reaching history (Chadwick & Arthur, 2010). Cricket is a competitive and continuously evolving sport. The science and statistics behind cricket can be used to aid coaches and players in achieving a competitive edge over their opponents. As statistics become more useful in explaining the differences between winning and losing teams in cricket so too does the interest in the statistical analysis of cricket variables. This provides the coaching and conditioning staff with a better understanding of what is required physically and strategically of each player to be successful in T20I and ODI cricket. These differences influence conditioning and team selection strategies (Peterson *et al.,* 2011). Every competitive cricket team or nation is looking for any advantage to keep ahead of the times and other opponents. Cricket has recently added Twenty-Twenty cricket to its roster of formats in an attempt to combat declining popularity. Research into the differences between One Day International (ODI) and Twenty-Twenty International (T20I) cricket is becoming much sought after. Discriminating between

winning and losing teams is of particular interest to researchers and cricket coaches alike.

Cricket is characterised by its three disciplines namely batting, bowling, and fielding. These actions take place on an oval-shaped field that contains a pitch roughly in the centre. Unlike most sports cricket is played over many hours across three different formats. Test, ODI, and T20I cricket all differ in number of overs played per match. This influences each player's strategy as well as the overall team strategy for success. According to Chadwick and Arthur (2010), each format of cricket has conformed to a general style of play and can be defined by certain strategy. Test match cricket is a longer more conservative game. Intense planning and cautious play are the mainstays of a test match. ODI cricket is a shorter game played with less caution and more intensity due to the nature of limited number of overs (50 overs). T20I cricket is more of a spectators' sport than a battle of planning and execution. T20I cricket is the shortest format of cricket (20 overs). Players are more aggressive in an effort to score runs and take wickets in an explosive manner.

Sport analytics play an important role in providing plausible solutions to various problems associated with sports such as cricket (Perera, 2015). Problems such as the planning of appropriate match, technical, physical, tactical, and mental training for each format of cricket can be solved with research. Not only can statistical research aid in the quantifying of individual performance but also team performance and strategy effectiveness.

1.2 Problem Statement

Using quantitative data to better understand individual and team success has fast become a necessity for all competitive cricket teams. Data on performance is gathered by recording and analysing cricket matches and is used to evolve match strategies as well as mental and physical training regimes to ensure success in a match. Much research on the performance measures that contribute to the success of a cricket team has been conducted for both ODI and T20I cricket. Batting variables have been found to be most predictive of eventual success (Shah,

Hazarika & Hazarika, 2017). Although research exists on batting strategy (Preston & Thomas, 2000; Irvine & Kennedy, 2017), batting order (Gill, Swartz, Beaudoin, & De Silva, 2006; Douglas & Tam, 2010), individual measures of batting performance (Mukherjee, 2014), and measures of team performance (Najdan, Robins & Glazier, 2014), there is little research which investigates the performance measures related to success for teams batting first or second. An examination of performance measures for teams batting first and second is warranted since the cricket match situation that a team batting first faces is different to that which the team batting second experiences.

1.3 Purpose statement

The purpose of this study is to investigate the performance variables of batting (which include runs scored by the top four batsmen, sixes scored, fours scored, runs scored during the initial stages of a match and the end of a match) that differentiates between winning and losing teams batting first or second in ODI and T20I cricket.

1.4 Aims

1. To examine how batting performance variables that correlate with success differ based on batting first or second.
2. To differentiate between the batting performance variables of winning and losing teams in each format of the game.
3. To determine the batting performance variables that correlates the highest with success throughout the different formats.

1.5 Research questions

The following questions will be addressed:

1. Is there a significant correlation between winning an international cricket match and the chosen batting variables?
2. Will there be differentiating statistics between winning and losing teams in ODI cricket for batting?

3. Will there be discriminating statistics between winning and losing teams in T20I cricket for batting?
4. Will different variables correlate with success based on batting first or second?
5. Can the batting variables be used to predict future success?

1.6 Structure of dissertation

This dissertation is presented in five parts (see Figure 1). Chapter 1 introduces the problem statement and aims of the study. Chapter 2 is a literature review of cricket which outlines the format of the game and address the influence of statistics on match and training strategy. Chapters 3 and 4 are presented in article format. Article titles are as follows: Chapter 3: *Predictors of batting success for winning and losing teams in One-Day International cricket.* Chapter 4: *Predictors of batting success for winning and losing teams in Twenty-Twenty International cricket.* Chapter 5 represents the overall summary, conclusions, and recommendations. The Harvard method is used for referencing.

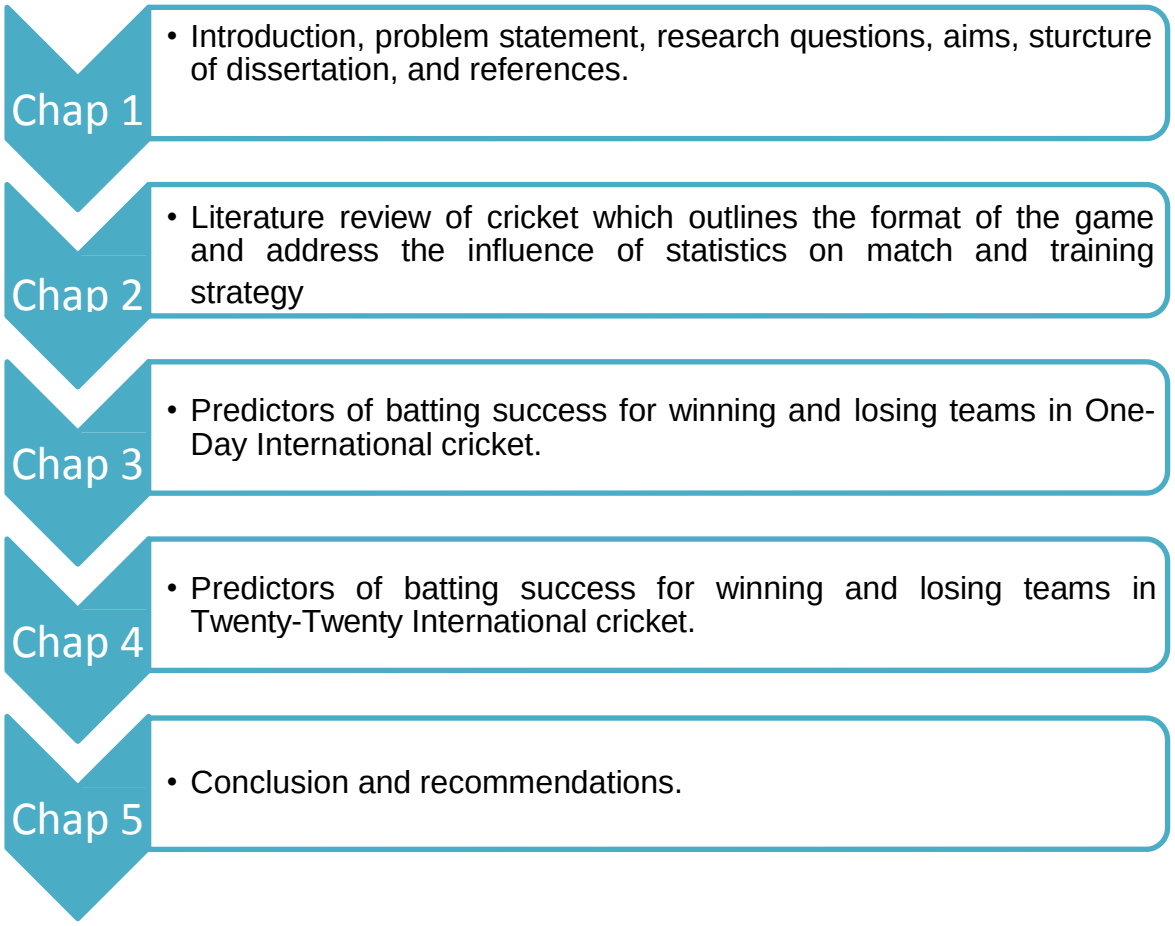

Figure 1. The structure of the dissertation

1.7 Ethical considerations

In this research, a strong and reliable data source is needed which was found in Statsguru. Statsguru is ESPN Cricinfo's cricket statistics maintenance database. In this database, all the match's data are stored with live ball by ball commentary (Munir, Hasan, Ahmed & Quraish, 2015). Data obtained from the One Day International cricket matches was recorded in Microsoft Excel. The data was then analyzed using the SAS statistical software (SAS, 2013). The research study does not involve any contact with the participants nor does it implement an intervention. Ethics clearance was obtained from the University of the Free State where the study was conducted under ethical clearance number UFS-HSD2017/0677.

1.8 References

Chadwick, S. & Arthur, D. (2010). *International cases in the business of sport.* Oxford, UK: Elsevier, pp. 263-264.

Douglas, J.M. & Tam, N. (2010). Analysis of team performances at the ICC World Twenty20 Cup 2009. International Journal of Performance Analysis in Sport, 10, pp. 47-53

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, pp. 350-359

Mukherjee, S. (2014). Quantifying individual performance in Cricket-A network analysis of batsmen and bowlers. Physica A, 393, pp. 624-637

Munir, F., Hasan, K., Ahmed, S. & Quraish, S. (2015). Predicting a T20 cricket match result while the match is in progress. A dissertation presented for the degree of Bachelor in Computer Science

Najdan, M.J., Robins, M.T. & Glazier, P.S. (2014). Determinants of success in English domestic Twenty20 cricket. International Journal of Performance Analysis in Sport, 14, pp. 276-295

Perera, G. H. (2015). Cricket Analytics. Dissertation presented for the degree of Doctor of philosophy. University of Peradeniya.

Peterson, C. J., Pyne, D. B., Portus, M. R. & Dawson, B. T. (2011). Comparason of player movement patterns between 1-day and test cricket. *Journal of strength and conditioning research,* 25(5), pp. 1368-1373.

Shah, S., Hazarika, P. J. & Hazarika, J. (2017). A study on Performance of Cricket Players using Factor Analysis Approach. International Journal of Advanced Research in Computer Science,8(3), pp.656-660

Swartz, T. B., Gill, P. S., Beaudoin, D. & deSilva, B.M. (2006). Optimal batting orders in one-day cricket. Computers & Operations Research, 33, pp. 1939-1950

Chapter 2 Literature review

2.1 Introduction

Cricket is a game of technical finesse consisting of many important variables that may be used to predict performance in batting, bowling, and fielding. Based on long-term performance on variables in batting, bowling or both batting and bowling professional cricket players are selected to play in international cricket matches. It is therefore important that focus is placed on the variables that influence the batting, bowling and fielding performance of a cricket team during a match. Statistical analysis of these performance parameters can aid in determining the variables which are different between winning and losing cricket teams.

The purpose of this chapter is to provide an introduction to the game of cricket, an overview of batting as a success predictor, and the importance of sport analytics in informing coaching and conditioning staff on the demands of international cricket.

2.2 International Cricket

Cricket has three formats: test, one day international and T20. Each format of cricket differs in rules and therefore strategy. Test cricket is played over five days with each team usually batting and bowling twice. The batting team can bat as long as they want providing they have wickets remaining. The days' play ends once 90 overs have been bowled, the light is too bad or the time deadline for the day has been reached. An innings ends when the batting team is bowled out (10 wickets) or the batting team declares. The team with the most runs scored after the two innings are complete wins but a draw can also ensue if the chasing team does not score the required target for victory and is not bowled out before the end of play on the fifth day. The ODI and T20I cricket formats are known as limited overs cricket because in each format the batting side is given a specific number of overs in which to score runs. The ODI cricket format consists of 50 overs per innings while T20 cricket consists of 20 overs per innings. T20I blowers are allowed to bowl a maximum of four overs each whilst ODI bowlers are allowed a maximum of ten overs each. Limited overs cricket not only transformed cricket as a sport but also appeals to a broader variety of spectators. Compared to test cricket, ODI 50 over cricket is viewed by many as a more exciting and manageable dose of cricket that still includes strategy. In comparison to ODI cricket, T20I cricket is a more fast-paced and frantic competition with little scope for strategy further than scoring runs fast (Cannonier, Panda, & Sarangi, 2015). T20 cricket was introduced to create fast-paced and exciting cricket for viewers.

As part of fast paced limited overs cricket, the ICC introduced power play overs in 2005 (Silva, Manage, & Swartz, 2015). Both ODI and T20 cricket consist of mandatory overs in which the fielding team is only allowed a certain number of fielders outside the inner ring of the field (30 yard circle around the pitch). ODI power play rules have been changed often since their introduction. Silva, Manage, and Swartz (2015) argue that the power play overs favour the batting team. Their study found that the power play overs only contribute 6.5 runs on average. The power play overs are thought to make run scoring easier and faster. Although run scoring is faster the power play overs also coincide with the loss of wickets as pressure is increased on the batsmen. Researchers should keep power play overs in mind when

examining data longitudinally. On the other hand, T20 power play rules have been consistent over the years. Silva et al (2015) suggested the shortened version, that is T20 cricket, has an optimal placement of the power play overs. The authors did not further study the contribution of runs the power play overs make in T20 cricket.

In some cases, the normal play of cricket is interrupted by weather such as rain. Since cricket is not played in the rain, the umpires and match referee decide whether to delay the game or utilise the Duckworth-Lewis method (Duckworth & Lewis, n.d.). This method, only applied to ODI cricket, is based on the notion that both teams begin the match with a certain amount of resources (300 balls and 10 wickets), when a game is reduced due to rain the number of resources is also reduced. The number of runs needed to win a match or the number of remaining overs is determined using the Duckworth-Lewis method. This ensures the match remains fair and competitive. Comparison of uninterrupted matches with that of matches completed using the Duckworth-Lewis method is cautioned against. Interrupted matches are inconsistent and can be seen as outliers.

Interrupted cricket play aside, the difference in the number of overs between ODI and T20I cricket may affect the team strategies used for winning. An investigation into the variables that differ between each format is worthwhile. Research suggests that there are indeed differences in performance variables between the two formats. Shah, Hazarika and Hazarika (2017) found that the variance that batting and bowling accounted for in ODI (ICC World Cup 2015) and T20 (Indian Premier League 2016) cricket differed. While batting accounted for 56.8% of the variance in ODI, this was as high as 62.5% in T20. Cannonier, Panda and Sarangi (2015) showed that as the data moved from ODI to T20I aggressive batting became increasingly important and the careful retention of wickets became less important. Differences between ODI and T20I formats are also shown by the ICC player ranking for ODI batsmen and T20I batsmen where the same batsmen that are successful in ODI may not be as successful in T20I and vice versa. Evidentially ODI and T20I are different; research conducted to fully understand the differences will aid coaching and conditioning staff and players in preparing for each ODI and T20I cricket separately.

Because cricket matches of different format are different in length, intensity, and rules, players have to adapt to perform effectively based on the demands of each format of cricket. Peterson, Pyne, Portus and Dawson (2011) investigated the differences in physical demands between ODI and test cricket fielding. On completion of the study the authors found slight differences between match demands including but not limited to total distance walking, jogging, and sprinting, total time spent sprinting, striding, and walking, and amount of high intensity efforts. They found that the most substantial difference was the higher physical workload that test match cricket player incurs. This higher physical load is mostly due to test cricket being much longer than ODI cricket. The fact that this study found differences in match demands between two different formats of cricket means that there may be statistical differences between ODI and T20 cricket. The study did not investigate T20 cricket nor did it investigate batting or bowling match demands. These differences could influence conditioning and team selection strategies (Peterson *et al.,* 2011), warranting further research on these differences.

The following section of the literature review addresses batting since it contributes significantly to overall success in 50 and 20 over cricket. Shah, Hazarika and Hazarika (2017) found that batting accounts for much more variance in winning or losing a cricket match for both T20 and ODI cricket. This along with the necessity to keep the study of realistic length this study focuses on batting as a success predictor of cricket.

2.3 Batting

Eleven players are selected to be a part of a cricket team for a specific match. Each player is expected to excel in their specific role as a batsman, bowler, or all-rounder (Amin & Sharma, 2014a). The efficiency of the team depends on the combined individual performances of each player. Most cricket teams have batsmen, batting all-rounders, and a wicket-keeper batsman. A batsman by definition rarely bowls and is expected to perform well in batting, a batting all-rounder is a batsman that has the ability to bowl as well as bat, and a wicket-keeper batsman is the wicket-keeper who is expected to bat well in a match.

2.4 Batting strategy

In both ODI and T20I cricket the batsmen are trying to score as many runs as they can in the limited amount of overs given. In doing this the batsmen expose themselves to the risk of being dismissed. It can then be assumed that most batsmen are dismissed in an attempt to raise the run rate. The possibility of being dismissed increases with aggressive batting which can lead to the fall of wickets (Preston & Thomas, 2000). Avoiding this risk may result in a slow scoring rate at the start of the innings or losing too many early wickets creates pressure for the batting team. Douglas and Tam (2009) suggest that batsmen should be selected based on the ability to score boundaries with minimal risk of losing a wicket in the first six overs of a T20 cricket match. This study can be extended in order to investigate if the same strategy can be generalised to international T20 cricket as well as for both teams batting first and second. Irvine and Kennedy (2017) concluded that specialist batsmen that are capable of consistently scoring boundaries should be utilised during the first six and last five overs of a T20 cricket match. Douglas and Tam (2009), and Irvine and Kennedy (2017) agree on the batting strategy that should be employed in T20 cricket, which encapsulates scoring runs at a high run rate improves the position of the batting team.

Preston and Thomas (2000) investigated the optimal batting strategy in order to achieve victory in one-day cricket. The authors surmised that the optimal strategy for teams batting second is to score runs at a rate that is consistent with the required run rate per over whilst not losing too many wickets in the process. The optimal strategy for teams batting first is to minimize the win conditions for teams batting second. Scoring more runs in the first innings increases the required run rate per over which pressures the team batting second into batting faster to conform to the optimal batting strategy. As the team attempts to score runs faster they increase the probability of a dismissal and decrease the probability of winning. Preston and Thomas (2000) concluded that limited overs batsmen may actually be optimizing in terms of a batting strategy. The authors suggested more research is required into the strategy used by cricket teams batting first and second.

2.5 Batting order

The generally accepted method of implementing an effective batting order is to place the more proficient batsmen towards the top of the batting order whilst placing the less proficient batsmen towards the bottom of the batting order (Gill, Swartz, Beaudoin, & De Silva, 2006). As previously stated the proficiency of a batsman is measured using his batting average which is his total runs scored divided by the number of cricket matches he has played. Using this method of batting order selection has its advantages. Placing the best batsmen at the top of the batting order in one-day cricket provides the more effective batsmen with a greater opportunity to bat for a long period of time (Gill *et al*., 2006). Douglas and Tam (2009) suggest that batsmen that are able to score runs quickly with minimal risk of losing wickets in the first six overs of a T20 cricket match are very important to winning. The selection of batsmen in the order of a batting line up seems to be very important for both T20 and ODI cricket. Bandulasiri, Brown and Wickramasinghe (2016) claimed that the middle order batsmen are the most crucial for winning a cricket match. The authors also emphasise the importance of the first four batsmen as they are responsible for setting up a base for the innings. The study also states that the middle order batsmen are especially important for teams batting second as it is these batsmen that control the innings when chasing a total.

The rise in popularity of T20 cricket has prompted coaching staff to determine if the same batting order strategy used for ODI cricket is effective for T20 cricket. T20 cricket has brought about the need for batsmen that are capable of consistently scoring boundaries throughout a T20 match (Najdan, Robins & Glazier, 2014). Batting order is a part of the strategy that may help win a cricket match in both T20 and ODI. Statistical analysis of performance will reveal how much of a role the top four batsmen play in winning the cricket match. According to the previously mentioned studies, it is recommended batsmen that have a high strike rate (high chance of scoring boundaries) bat higher in the batting order for T20 cricket. Some batsmen are not seen playing T20 cricket as they do not possess the necessary

skills and physicality to score runs at a high strike rate. The role physical strength plays in batting is discussed in the next section.

2.6 Physiological demands of batting

It is important for conditioning staff to better understand the physical demands of playing in a cricket match so they can better train their clients (Peterson, Pyne, Portus, Dawson, 2009). During any given cricket match a batsman can be expected to run maximally between the wickets over an extensive period of time whilst wearing hindering protective gear (Stretch et al., 2000). This, in combination with intensive stroke play, places certain physical demands on batsmen. Appropriate conditioning is required to ensure that the batsmen perform optimally in each cricket match. These demands differ between ODI and T20 cricket as T20 cricket is much shorter in duration than ODI cricket. Therefore, optimal conditioning for ODI cricket differs from the conditioning required for successful performance in T20 cricket. Stretch et al. (2000) found that as a result of the physical demands of a cricket match cricket players look different not only to non-cricket players but also to each other. Peterson *et al*. (2009) investigated the differences in intensity of match play between one-day cricket and T20 cricket. The authors reported significant differences in intensity of a cricket match between one-day and T20 cricket. The authors concluded that T20 cricket was played at a higher intensity than one-day cricket. Although this study shows differences in match intensity between one-day and T20 cricket it did not take into account the effort required to hit the cricket ball. The study only measured intensity as walking, jogging, and sprinting. The strength required to perform general and specific stroke play may contribute to the intensity of a cricket match. This may also differ between ODI and T20 cricket. Stretch *et al* (2000) also suggest that the combination of energy systems, such as the anaerobic and aerobic systems, required for optimal performance in different cricket formats can differ and are also very important to consider.

Statistical analysis can provide further insight into which performance parameters are most important for ODI and T20 cricket, given that there are different demands made on players in each format. Greater insight into the differences in performance

parameters can provide conditioning staff with an advantage in knowing how to prepare their cricket players specifically for the expected match demands placed on a successful batsman.

2.7 Upper body strength and batting performance

Gray, Taliep, and Prim (2010) found that there is a positive correlation between upper body strength and maximum distance a batsman can hit a cricket ball. The findings are especially relevant to T20 cricket. The ability of the team and individual to hit boundaries increases the competitiveness of a team in T20 cricket and may be positively correlated with the success of a team. Upon analysis of these findings, upper body strength is identified as an essential physical trait to a successful cricket batsman (Gray *et al.,* 2010). Although a positive correlation between upper body strength and hitting distance was found in the study, there was no significant correlation between upper body strength and batting performance indicators such as batting average and strike rate. The application of upper body strength through the right technique and timing may be the more effective method of hitting boundaries rather than relying on strength alone.

2.8 Psychological factors in batting

Individual player performance is not only affected by physical demands, but by psychological factors as well. Weissensteiner, Abernethy, Farrow and Gross (2012) evaluated the importance of mental toughness in cricket. Mental toughness was defined as a state of prolonged focus that is affected by certain emotional factors such as self-belief. The study found that the more experienced and more skilled batsmen scored higher on the mental toughness tests that were administered. Additionally, Lemmer (2015) implicated psychological factors in explaining why some teams lose a game from a very commanding position. This study referred to this phenomenon as strangling. Commentators and spectators of cricket in South Africa may refer to this as ‘choking’. Mental toughness in combination with skill and experience is essential for success in cricket.

2.9 Differences between batting first and second

Finally, the situation faced by a batsman is fundamentally different when batting first and second (Bhattacharjee & Lemmer, 2016). Bhattacharjee & Lemmer (2016) used pressure as a measure of the differences between batting in the first and second innings of a cricket match. There are many factors that can create pressure and change the nature of the match situation between innings. Pitch conditions can affect a cricket match. Weather conditions affect the cricket pitch which can lead to the team batting first or second having an advantage. As a cricket match continues some cricket payers may experience mental and physical tiredness. Davis, Perera and Swartz (2015) restricted their study to the first innings of a cricket match because, they argue, the game changes in the second innings. Due to the overall different situation faced by teams batting first and second this study will analyse statistics separated into teams batting first and second.

Studies conducted by Bhaskar (2009) and Dawson, Morley, Paton, and Thomas (2009) investigated the probability of winning a One-Day cricket match when winning the toss and electing to bat first or second. The studies found that the team winning the toss and electing to bat first had an advantage in day-night cricket matches and a disadvantage in day cricket matches. The authors separated batting first from batting second irrespective of the toss. The authors analysed the advantage of batting first or second as it pertains to winning a cricket match thereby providing data on the strategy that could be implemented in order to win a cricket match when batting first or second. Winning the toss in cricket is important but cricket teams cannot afford to rely on winning the toss in order to gain an advantage in the match. This is why a strategy for batting first and second should be explored and planned.

2.10 Sport analytics

Statistics have become very popular in recent years in many sports including cricket. The analysis of statistics aids in understanding which factors greatly affect performance. This knowledge leads to more specific training programmes and a more educated selection of players for the team. For example, statistical analyses of performance have been shown to aid coaches and players in correcting technique,

planning for specific situations, improving overall match strategy and preparedness, and predicting success.

Typically, the analysis of recorded and captured cricket statistics such as batting averages, bowling economy, batting strike rate, and wickets taken are used to rate the efficiency of individual players and in turn the entire team (Stretch *et al.,* 2000). Stretch et al. (2000) ranked cricket players based on the analysis of differential statistics. The ICC's ranking system as well as PageRank player ranking systems (Stretch et al., 2000) influences team selection. A number of researchers have shown how cricket statistics can be used to evaluate the quality of players. Amin and Sharma (2014b) rated the efficiency of cricket players based on statistical analysis. Gill, Swartz, Beaudoin and De Silva's (2006) research exemplified how to optimise batting order. How to develop a successful batting strategy for winning has been examined by Preston and Thomas (2000) and Irvine and Kennedy (2017). The number of studies examining the use of statistics in improving success shows that sport analytics is very beneficial to the advancement of cricket as a competitive sport. Nevertheless, there are limited studies which analyse the performance of teams batting first or second. It may be the case that batting first or second will require different strategies at the team and individual level given that teams batting second have a runs target to meet, whereas those batting first do not.

The manner in which data is collected, captured, and analysed is as important as the results they yield. A number of methods have been used for data collection in cricket; one such method is time-motion analysis, for example, used by Duffield and Drinkwater (2008). With the use of Global Positioning Systems (GPS) the authors were able to track movements performed by cricket players at different intensities. The advantage of this method is that a GPS system provides specific information about each player's movements during a match or practice. This aids conditioning and coaching staff in understanding the physical demands of a match and practice. The disadvantage is that a time motion analysis is without funding an expensive study to conduct.

Another method of data collection is the observation of matches. The data collector will observe a match in real time or via video recordings and capture data according to the relevant variables for the study. Fortunately, large amounts of observational data are publicly available. For example, Statsguru, ESPN Cricinfo's cricket statistics maintenance database, saves data relating to all ODI and T20 matches of cricket. In this database system, all the match data are stored with live ball by ball commentary (Munir, Hasan, Ahmed & Quraish, 2015). Cricinfo is a reputable website for data collection and is used by authors such as Shah et al, (2015) and Douglas and Tam (2010) for analysis.

A number of statistical analyses have been used in sports analytics. For example Douglas and Tam (2010) used inferential statistics such as the Students t-test and the Cohen's d-test in order to investigate statistical differences between various performance parameters and winning. The advantage of this was a very insightful look into the main differences between winning and losing cricket teams. The disadvantage of this study is that although it did investigate the differences between winning and losing the study did not create any analysis that could be used for predicting success in future cricket matches. Lemmer, Bhattacharjee, and Saikia (2014) state that predicting the outcome of a match in any sport is difficult due to the inconsistent results when teams play each other more than once. The univariate logistic regression is a powerful statistical tool for analysing the differences between winning and losing cricket teams as well as creating a base for future prediction models (Peng, Lee, & Ingersoll, 2002).

2.11 Summary

This review explained that there are many differences between ODI and T20I cricket. There are many different influencing factors with regards to batting first and second in a cricket match. The variables used to predict success will be different between teams batting first and second. Analysis of statistics for winning and losing ODI and T20I cricket teams will reveal which variables influence the success of a cricket team in a match and may even indicate the type of cricket player that is most valuable to a winning team. Further research will identify which training methods best prepare

cricket players for the specific match demands of a winning team in different situations.

Conditioning staff have to take the physical demands of the last match into consideration as it can have an effect on the next match, especially if the next match is a different format of cricket (Peterson *et al*., 2011). Analysis of variables that include but are not limited to the number of fours hit by batsmen, sixes hit by batsmen, and runs scored by the top four batsmen may be used to predict future success or failure of a team in each format of the game. Determining the statistical differences and correlations between winning and certain variables provide the coaching and conditioning staff with a better understanding of what is required of each player to be successful in each format of cricket. The review of the literature also shows that more research is needed in this area.

2.12 References

Amin, G. R. & Sharma, S. K. (2014a). Cricket team selection using data envelopment analysis. *European Journal of Sport Science,* 14, 369-376.

Amin, G. R. & Sharma, S. K. (2014b). Measuring batting parameters in cricket: a two stage regression- OWA method. *Measurement,* 53, 56-61.

Bandulasiri, A., Brown, T. & Wickramasinghe, I. (2016). Factors affecting the result of matches in the one day format of cricket. *Operations research and decisions*, 4, 22-32

Bhaskar, V. (2009). Rational Adversaries? Evidence from Randomised trials in One Day Cricket. *The Economic Journal,* 1-23.

Cannonier, C., Panda, B. & Sarangi, S. (2015). 20-Over Versus 50-Over Cricket: Is There A Difference? *Journal of Sports Economics,* 16(7), 76-783.

Chadwick, S. & Arthur, D. (2010). *International cases in the business of sport.* Oxford, UK: Elsevier, 263-264.

Dawson, P., Morley, B., Paton, D. & Thomas, D. (2009). To bat or not to bat: An examination of match outcomes in day-night limited overs cricket. *Journal of Operational Research Society,* 60, 1786-1793.

Douglas, J.M. & Tam, N. (2010). Analysis of team performances at the ICC World Twenty20 Cup 2009. *International Journal of Performance Analysis in Sport*, 10, 47-53

Duckworth, F. & Lewis, T. (n.d.). ESPN cricinfo. www.cricinfo.com. 10/04/2018

Duffield, R. & Drinkwater, E.J. Time-motion analysis of Test and One-day international cricket centuries. Journal of Sports Sciences, 26(5), 457-464

Gill, P. S., Swartz, T. B., Beaudoin, D. & de Silva, B. M. (2006). Optimal batting orders in one-day cricket. Computers & Operations Research, 33, 1939-1950.

Gray, J., Taliep, M. S. & Prim, S. K. (2010). Upper body muscle strength and batting. *Journal of strength and conditioning research,* 24(12), 3484-3487.

Irvine, S. & Kennedy, R. (2017). Analysis of performance indicators that most significantly affect International Twenty20 cricket. I*nternational Journal of Performance Analysis in Sport*, 17(3), 350-359

Lemmer, H.H., Bhattacharjee, D. & Saikia, H. (2014). A consistency Adjusted Measure for the Success of Prediction Methods in Cricket. *International Journal of Sports Science & Coaching*, 9(3), 497-512

Munir, F., Hasan, K., Ahmed, S. & Quraish, S. (2015). Predicting a T20 cricket match result while the match is in progress. A dissertation presented for the degree of Bachelor in Computer Science.

Najdan, M.J., Robins, M.T., Glazier, P.S. (2014). Determinants of success in English domestic Twenty20 cricket. International Journal of Performance Analysis in Sport, 14, 276-295

Peng, C.Y. J., Lee, K. L. & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research,* 96(1), 1-14.

Peterson, C. J., Pyne, D. B., Portus, M. R. & Dawson, B. T. (2011). Comparison of player movement patterns between 1-day and test cricket. *Journal of strength and conditioning research,* 25(5), 1368-1373.

Preston, I. & Thomas, J. (2000). Batting strategy in limited overs cricket. *The statistician,* 95-106.

Shah, S., Hazarika, P. J., Hazarika, J. (2017). A study on Performance of Cricket Players using Factor Analysis Approach. International Journal of Advanced Research in Computer Science,8(3), 656-660

Silva, R.M., Manage, A.B.W., & Swartz, T.B. (2015). A Study of the Powerplay in One-Day Cricket. European Journal of Operational Research, 244, 1-20

Stretch, R. A., Bartlett, R. & Davids, K. (2000). A review of batting in men' s cricket. *Journal of sport sciences,* November, Volume 18, 931-949.

Weissensteiner, J. R., Abernethy, B., Farrow, D. & Gross, J. (2012). Distinguishing psychological characteristics of expert cricket batsmen. *Journal of Science and Medicine in Sport,* Volume 15, 74-79.

Chapter 3: Article 1: Predictors of batting success for winning and losing teams in One-day International cricket

To be submitted to *South African Journal for Research in Sport, Physical Education and Recreation* (Appendix A). Note that the referencing and formatting guidelines of the journal are followed.

Abstract

Predictors of batting success for winning and losing teams in One Day International cricket

[1]Mark Schaefer, [1]Riaan Schoeman, [2]Robert Schall,
[1]Department of Exercise and Sport Sciences, [2]Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa

The aim of this study was to evaluate batting variables in ODI cricket and identify predictors that discriminate between winning and losing batting teams. Understanding the batting variables that predict the success of an ODI cricket team will aid coaching staff in team selection, batting order, and overall match strategy. Match data from the 2014 and 2015 ODI cricket season was recorded from cricinfo. A total of 60 ODI cricket matches were observed. Significant predictors of winning an ODI cricket match when batting first were: runs scored in the first 20 overs (p=0.0019), runs scored in the last 12 overs (p=0.0004), sixes scored (p=0.0017), and the number of runs scored between the top four batsmen (p=0.0015); similarly, significant predictors for winning an ODI cricket match when batting second were: fours scored (p=0.0024), sixes scored (p=0.00277), runs scored between the top order batsmen (p=0.0197), and runs scored between the lower order batsmen (p=0.0222). Variables that predict success in ODI cricket differed for teams batting first and second, respectively.

Key words: Batting, One-day International, Runs scored, Boundaries.

3.1 Introduction

Eleven players are selected to be a part of a cricket team for a specific match and each player is expected to excel in their specific role as a batsman, bowler, or all-rounder (Amin & Sharma, 2014a). The efficiency of the team depends on the combined individual performances of each player. Most cricket teams have batsmen, batting all-rounders, bowlers, and a wicket-keeper batsman.

Cricket is a game of technical finesse consisting of many important factors that may influence batting performance. Recorded cricket statistics, such as batting averages and batting strike rate, are used to rate the efficiency of individual players and, in turn, the entire team (Stretch *et al.*, 2000). Mukherjee (2014) suggested that the quantification of individual cricket performances based on batting averages is of vital importance for team selection in a cricket match. The current study investigated other performance variables, such as runs scored in the first 20 overs of an ODI (One- Day International) cricket match. These variables may be used in conjunction with individual batting averages for team selection and match strategy. Studies such Amin and Sharma (2014a), Gill, Swartz, Beaudoin, and De Silva (2006), Preston and Thomas (2000), Irvine and Kennedy (2017), and Douglas and Tam (2010) provide extensive evidence that statistical analysis of performance variables will aid in determining how important specific variables (runs scored in the first 20 and last 12 overs, sixes scored, fours scored, runs scored between the top four batsmen, runs scored between the middle three batsmen, and runs scored between the lower four batsmen.) are to winning a match. There is a lack of recent research on the identification of specific statistical differences between the winning and losing team in a cricket match, especially when dividing the winning and losing teams into batting first and second. Lemmer, Bhattacharjee, and Saikia (2014) stated that predicting the outcome of a match in any sport is difficult. Therefore, it is necessary to review the important aspects of batting in order to better understand why the analysis of cricket statistics is a research field worth pursuing.

Although most of the research on cricket shows that batting is one of the most important aspects of winning a cricket match, Stuelcken, Portus, and Mason (2005) concluded that batting is often overlooked in the scientific literature. Success in batting can be characterised by many variables. These variables conventionally include batting average and strike rate, as well as the number of fours and sixes hit by each individual batsman. Batting order and team

batting strategy are variables which have been shown to influence a match. The generally accepted method of implementing an effective batting order is to place the more proficient batsmen towards the top of the batting order while placing the less proficient batsmen towards the bottom of the batting order (Gill, Swartz, Beaudoin & de Silva, 2006). Placing the best batsmen at the top of the batting order provides the more effective batsmen with a greater opportunity to bat for a longer period of time (Gill *et al.,* 2006). Other cricket studies have investigated individual batting performance, and movement variables using time-motion analysis (Peterson *et al.*, 2011; Rudkin & O'Donoghue, 2008; Amin & Sharma, 2014b). The information from these studies can be used to rank individual cricket batsmen. Barr & Kantor (2014) also emphasise the importance of using batting average and strike rate to measure batting performance. Measuring individual batting performance is important but some researchers, such as Mukherjee (2013) suggest that the manner in which batsmen perform as part of a team is also just as important. Team batting strategy is a worthwhile topic which is being overshadowed by individual batting statistics.

Determining predictors of winning in ODI cricket can provide the coaching staff with a better understanding of what is required of each player to be successful in this format of cricket. Therefore, this study will identify performance variables by means of statistics that discriminate between winning and losing teams. Furthermore, the study will reveal which variables correlate the highest with successful outcomes within this format of the game for teams batting first or second.

3.2 Methodology

3.2.1 Sample

A total of 60 professional ODI cricket matches between 2014 and 2015 were captured resulting in 120 records (two innings per match). Ten teams were selected for the purpose of this study, namely South Africa, Australia, New Zealand, India, Bangladesh, Zimbabwe, England, Sri Lanka, Pakistan, and West Indies. These teams all participate in all three formats of cricket namely test, ODI, and T20I cricket. Six matches from each team's records were randomly selected and observed. Of the six matches, three matches were won by the team batting first, and three matches were won by the team batting second. Drawn matches, and

those which employed the Duckworth-Lewis method were excluded from the study. Data were collected for all performance variables of concern in this study.

3.2.2 Data collection procedure

A quantitative study with an observational design by means of retrospective data analysis was used to determine batting performance variables that correlate the highest with the success of a team in ODI cricket. Data from the years 2014 and 2015 international cricket season was collected from Cricinfo (accessed 2015). All data was captured in Microsoft Excel 2007 and subsequently converted into a SAS data set. Data were statistically analyzed to evaluate and compare the variables during for winning and losing ODI cricket teams. The following variables were analyzed in this study: runs scored in the first 20 overs of a cricket match, runs scored in the last 12 overs of a cricket match, the number of fours and sixes scored in a cricket match, and runs scored by the top four, middle three, and bottom four batsmen in a cricket match to establish batting order contributions. In this research, a strong and reliable data source is needed which was found in Statsguru. Statsguru is ESPN Cricinfo's cricket statistics maintenance database. In this database, all the match's data are stored with live ball by ball commentary (Munir, Hasan, Ahmed & Quraish, 2015). ESPN crcinfo is considered to a reliable as it is used in professional cricket as well as being referenced by many published authors.

3.3 Statistical analysis

Data obtained from the One Day International cricket matches were recorded in Microsoft Excel. The data was then analyzed using the SAS statistical software (SAS, 2013).

Because of the fundamentally different match situation faced by the team batting first and second, respectively, the data were analysed separately for the team batting first and for the team batting second. The outcome of the match is a binary variable (win/lose) since drawn matches were excluded from the study. The association of the potential predictor variables with the match outcome was analyzed using univariate logistic regression, fitting each predictor variable, one at a time. The statistical significance of each predictor variable was tested using an exact test (exact conditional logistic regression). Furthermore, an odds ratio and associated 95% confidence interval is reported which reflects the effect (that is, the

increase in the odds of winning) associated with an increase of one unit of the predictor variable. The analysis was carried out using SAS procedure LOGISTIC (SAS, 2013).

3.4 Results

3.4.1 Batting first and second

Table 1 and Table 2 represent the mean values of each potential predictor variable, respectively for the winning and losing teams; Table 1 presents the data for teams batting first and Table 2 for teams batting second. Statistical significance was set at the 95% confidence interval ($p<0.05$). Values within the 95% confidence interval indicate a significant relationship between the performance variable and winning an ODI cricket match.

Table 1 indicates significance for runs scored in the first 20 overs, runs scored in the last 12 overs, runs scored by top order batsmen and the number of sixes scored. It is clear that the contribution of the middle and lower order batsmen to winning a game when batting first is not significant. Table 1 also show greater significance for sixes scored rather than fours scored.

Table 1. Team batting first: mean values of potential predictors for winning and losing teams

Team potential predictor of success	Means		p value
	Win	Lose	
Runs scored in the first 20 overs	89.9	71.9	0.0002*
Runs scored in the last 12 overs	108.4	61.2	<0.0001*
Fours scored	20.7	16.9	0.1830
Sixes scored	14.7	5.4	<0.0001*
Runs scored by the top order batsmen	186	106.7	0.0002*
Runs scored by the middle order batsmen	88.6	73.9	0.3240
Runs scored by the lower order batsmen	24.7	35.3	0.1207

* Statistically significant p-value from logistic regression analysis

Table 2 shows that there is a greater reliance on boundaries scored in order to win a game when batting second. The number of fours scored is more important for teams batting second than those batting first. Runs scored by the top order ($p=0.0239$) and lower order ($p=0.0108$) show high significance unlike for teams batting first.

Table 2. Team batting second: mean values of potential predictors for winning and losing teams

Team potential predictor of success	Means		p value
	Win	Lose	
Runs scored in the first 20 overs	91.7	90.1	0.7958
Runs scored in the last 12 overs	37.1	33.6	0.7018
Fours scored	20.9	13.8	0.0028*
Sixes scored	5	9.4	0.0384*
Runs scored by the top order batsmen	151.2	115.2	0.0239*
Runs scored by the middle order batsmen	54.6	66.3	0.3385
Runs scored by the lower order batsmen	11.9	25.1	0.0108*

* Statistically significant p-value from logistic regression analysis

Tables 3 and 4 represent the results of the statistical analysis completed using a univariate logistic regression. Table 3 represents the results for teams batting first. Table 4 represents the results for team batting second.

Table 3. Univariate logistic regression: Predictors of match outcome ODI data; Team batting first

Predictor Variable	[a]95% CI	Odds ratio	p Value	Test Statistic[b]
Runs scored first 20 overs	1.029 to 1.117	1.067	0.0002	12.1582
Runs scored last 12 overs	1.022 to 1.071	1.043	<0.0001	18.5247
Fours scored	0.986 to 1.089	1.034	0.1830	1.8221
Sixes scored	1.082 to 1.338	1.183	<0.0001	14.7809
Runs scored by the top order batsmen	1.006 to 1.024	1.014	0.0002	12.2078
Runs scored by middle order batsmen	0.996 to 1.015	1.005	0.3240	1.0051
Runs scored by lower order batsmen	0.963 to 1.004	0.984	0.1207	2.4351

Note:[a]Profile likelihood confidence interval; [b]Exact conditional score test

Table 3 represents the calculated odd ratio with 95% CIs. The odds ratio for a variable in the logistic regression represents how the odds change with a 1-unit increase of that variable. Table 3 shows that the odds ratio is highest for sixes scored. This means that for every increase in sixes scored by one six, increases the odds of winning a match by 1.183.

In contrast to Table 3, Table 4 indicates that the highest recorded odds ratio is for fours scored. Tables 2 and 4 show a reliance on fours scored for ODI cricket teams batting second.

Table 4. Univariate logistic regression: Predictors of match outcome ODI data; Team batting second

Predictor Variable	[a]95% CI	Odds ratio	p Value	Test Statistic[b]
Runs scored first 20 overs	0.918 to 1.026	1.003	0.7958	0.0719
Runs scored last 12 overs	0.988 to 1.018	1.003	0.7018	0.1525
Fours scored	1.032 to 1.175	1.096	0.0028	8.7028
Sixes scored	0.860 to 0.994	0.931	0.0384	4.2968
Runs scored by the top order batsmen	1.001 to 1.020	1.010	0.0239	5.0167
Runs scored by middle order batsmen	0.983 to 1.005	0.994	0.3385	0.9411
Runs scored by lower order batsmen	0.929 to 0.991	0.962	0.0108	6.1956

Notae:[a]Profile likelihood confidence interval; [b]Exact conditional score test

3.4.2 Runs scored in the first 20 and last 12 overs

Teams batting second and winning (BSW) scored on average the most runs (M=91.1) in the first 20 overs of a cricket match innings. Teams batting first and winning (BFW) scored on average the most runs (M=108.4) in the last 12 overs of an innings.

According to the univariate logistic regression runs scored in the first 20 overs are positively related to winning the match for teams batting first (p=0.0002). That is to say, the higher the number of runs scored in the first 20 overs by the ODI team batting first the higher the probability that that team has of winning the cricket match. This increase in probability of success is shown in Figure 2.

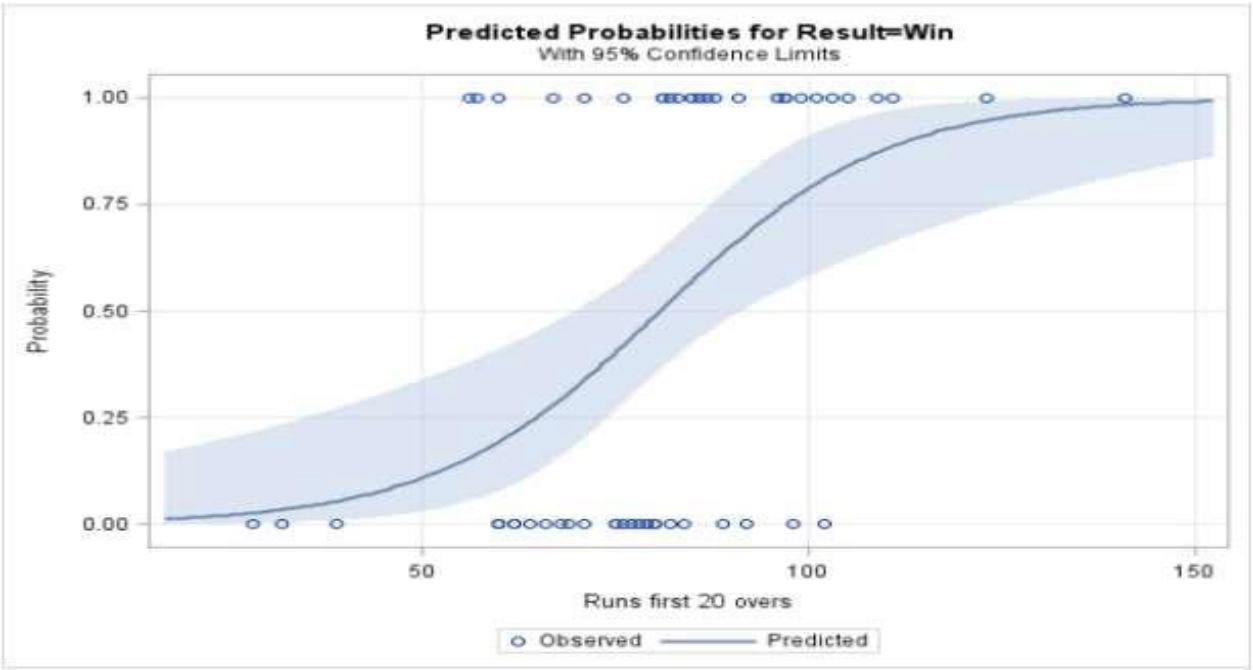

Figure 2. Probability graph for runs scored in the first 20 overs when batting first

The small increase in probability of success as a function of the runs scored in the first 20 overs for teams batting second - shown in Figure 3 shows that this variable is not significantly related to winning (p=0.7958).

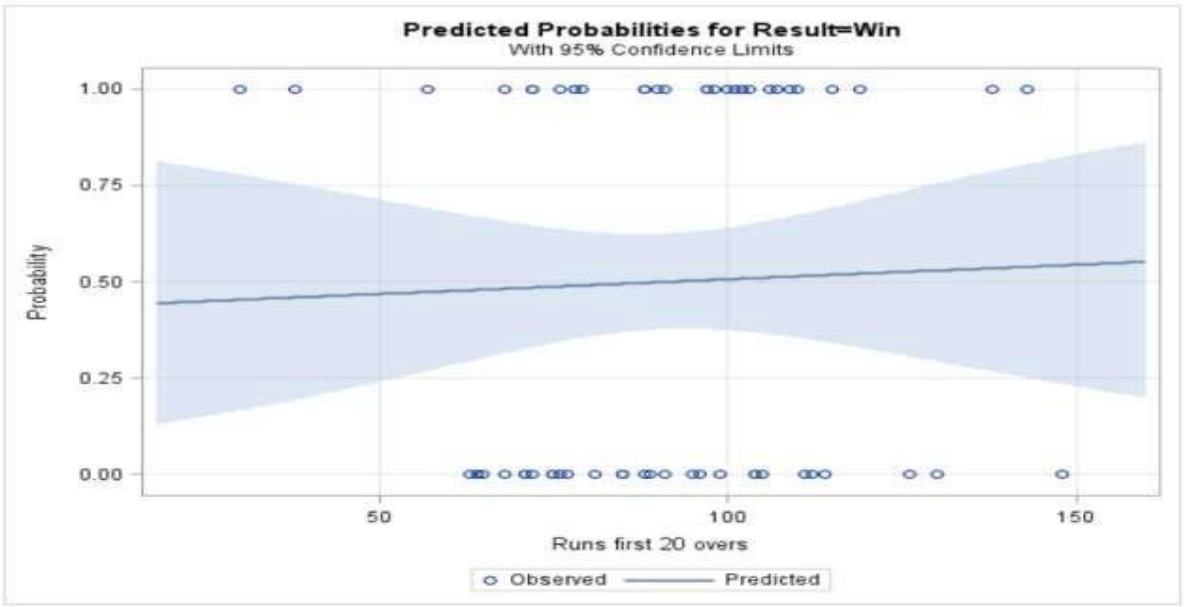

Figure 3. Probability graph for runs scored in the first 20 overs when batting second

According to the univariate logistic regression runs scored in the last 12 overs is positively related to winning the match for teams batting first (p=<0.0001). This means that for every run scored in the last 12 overs of an ODI cricket match the higher the probability of winning the match. Figure 4 represents this resulting increase in probability of success for teams batting first. Runs scored in the last 12 overs of an ODI cricket match is not significantly related to winning the match for teams batting second (p=0.7018) (figure 5).

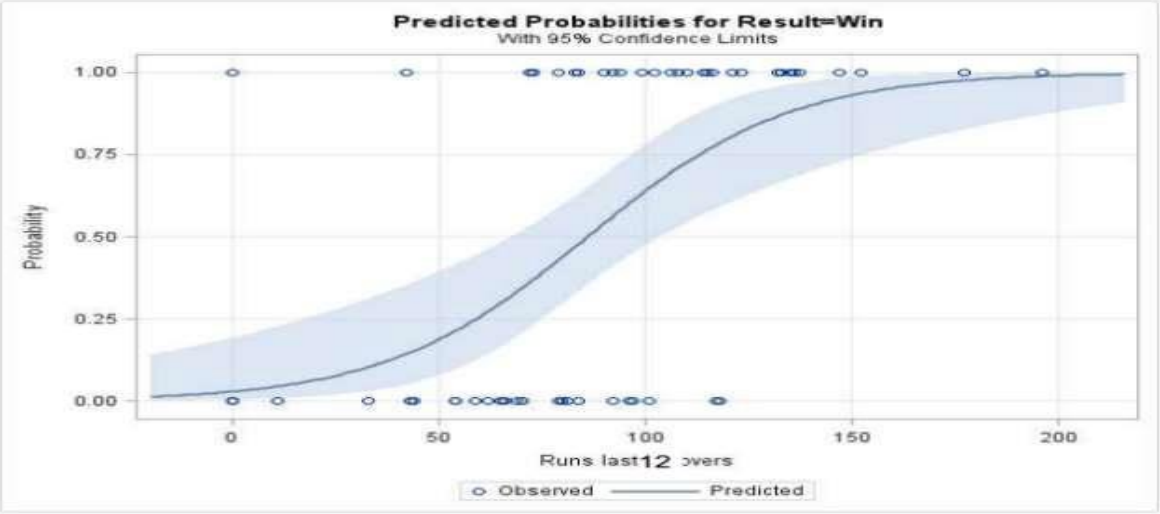

Figure 4. Probability graph for runs scored in the last 12 overs when batting first

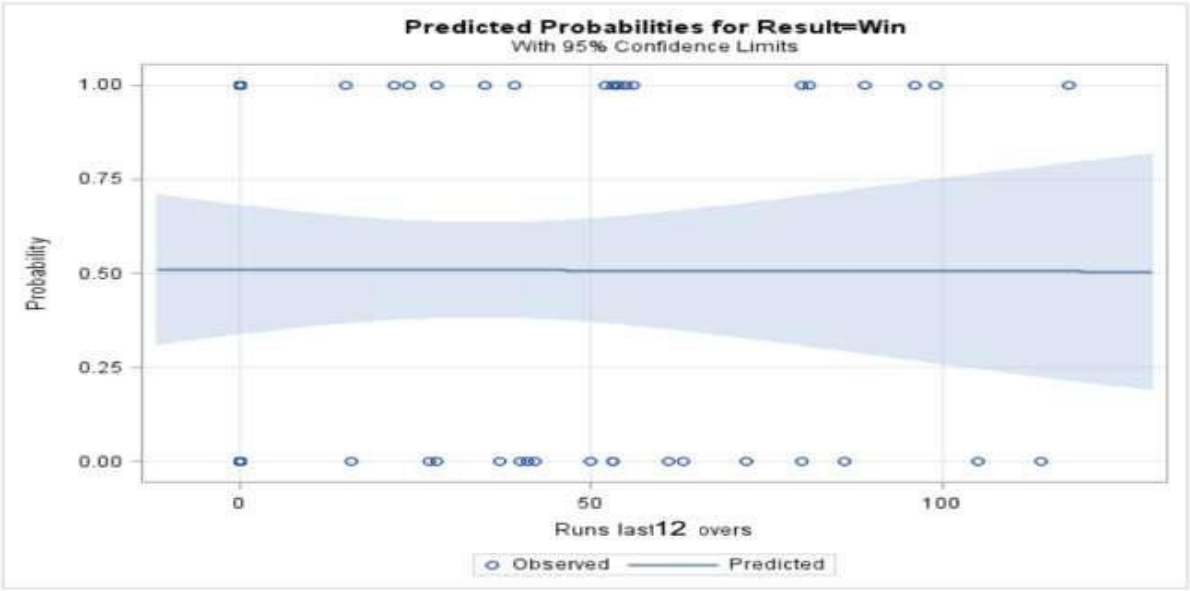

Figure 5. Probability graph for runs scored in the last 12 overs batting second

3.4.3 Boundaries

Tables 1 and 2 represent the mean number of different boundaries scored during a cricket match for both the winning and losing teams. The results between the winning and losing teams batting first and second are compared. Winning cricket teams that bat first and second, score on average more fours than the losing teams. As reported in Tables 1 and 2 the mean number of sixes scored was 14.7 during the recorded cricket matches is much higher for BFW than it is for teams batting first and losing (BFL) was 5.4, Teams batting second and wining (BSW) hit 5 sixes, and teams batting second and losing (BSL) hit 9.4 sixes. According to the univariate logistic regression the number of fours scored is positively related to winning the cricket match for teams batting second (p=0.0028) as shown in Figure 6 that represents the increase in probability of success.

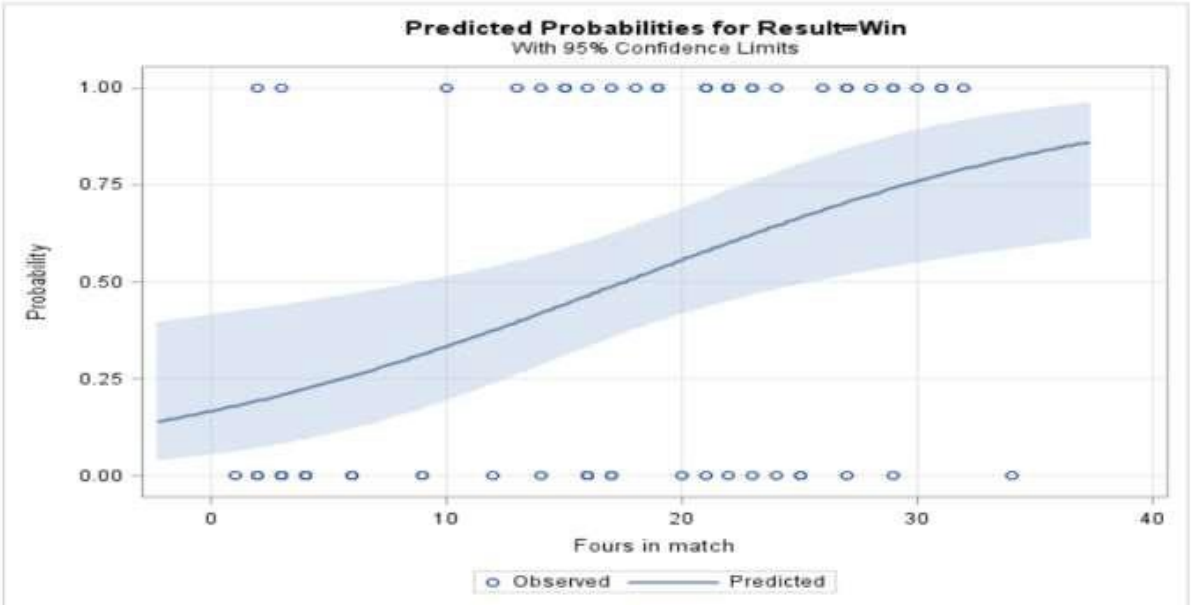

Figure 6. Probability graph for fours scored when batting first

The univariate logistic regression shows no significant relation between winning an ODI cricket match and the number of fours scored by the ODI team batting first (Figure 7) (p=0.1830). The amount of sixes scored throughout an ODI cricket match is positively related to winning an ODI cricket match for teams batting first (p=<0.0001). That is to say, the higher the number of sixes hit by the ODI team batting first the higher the probability that that team has of winning the cricket match. This increase in probability of success is shown in Figure 8. Interestingly the amount of sixes scored is negatively related to wining an ODI

cricket match for teams batting second (p=0.0384). That is to say, the higher the number of sixes scored by the ODI team batting second the lower the probability the team has of winning the cricket match. This decrease in probability of success is shown in Figure 8 and 9.

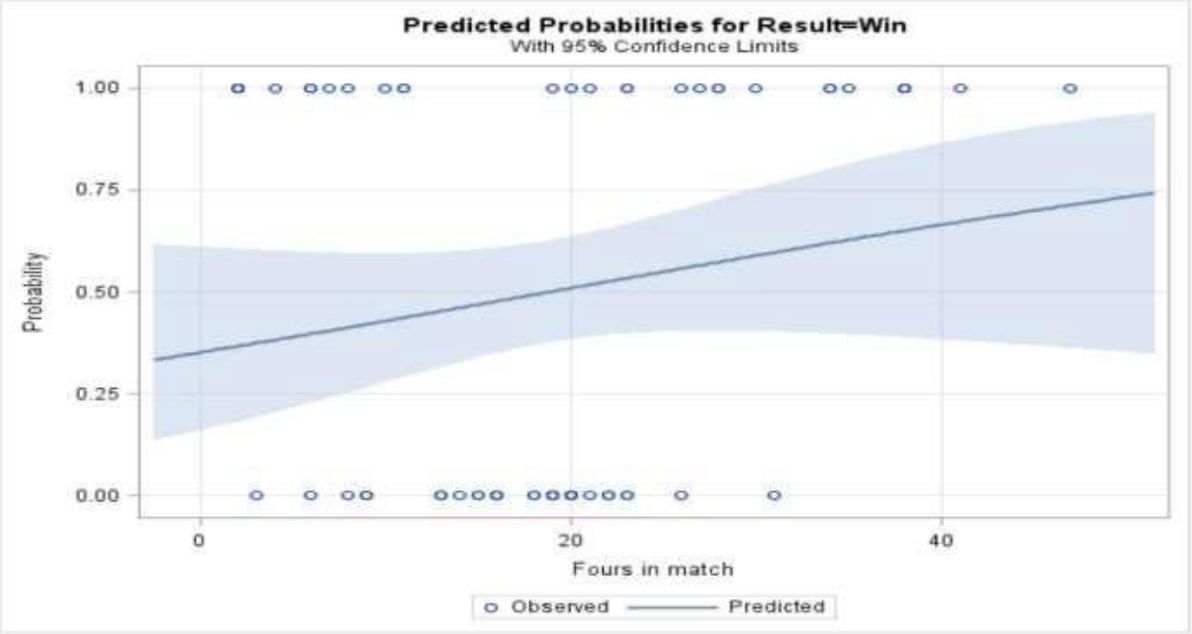

Figure 7. Probability graph for fours scored when batting second

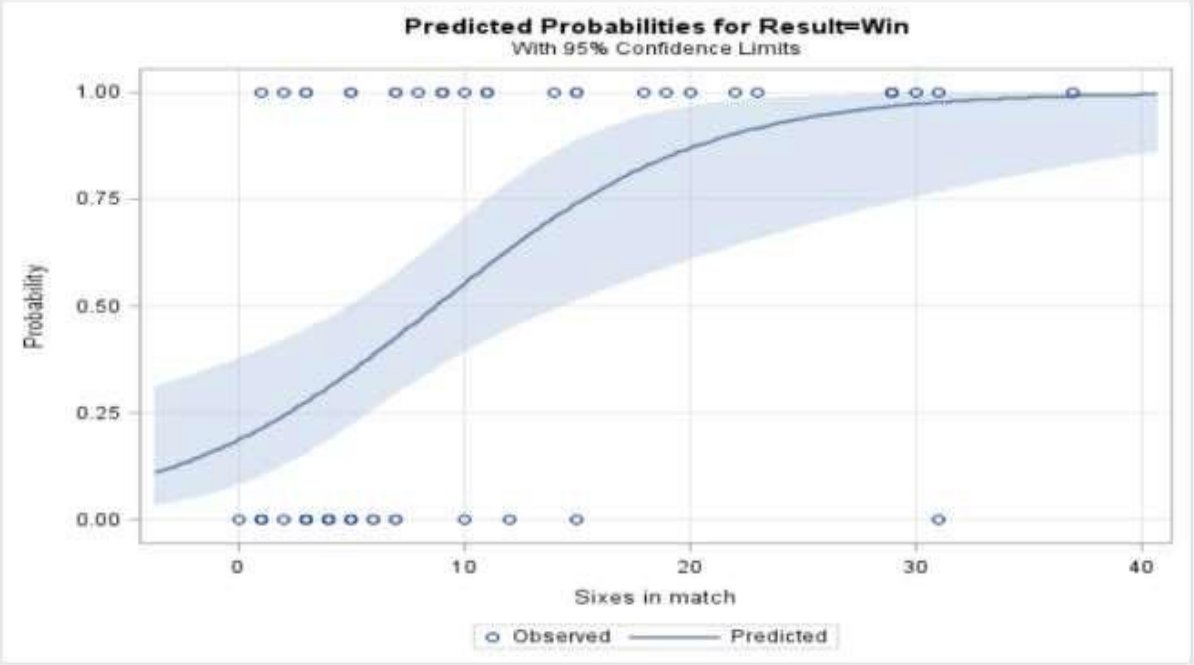

Figure 8. Probability graph for sixes scored batting first

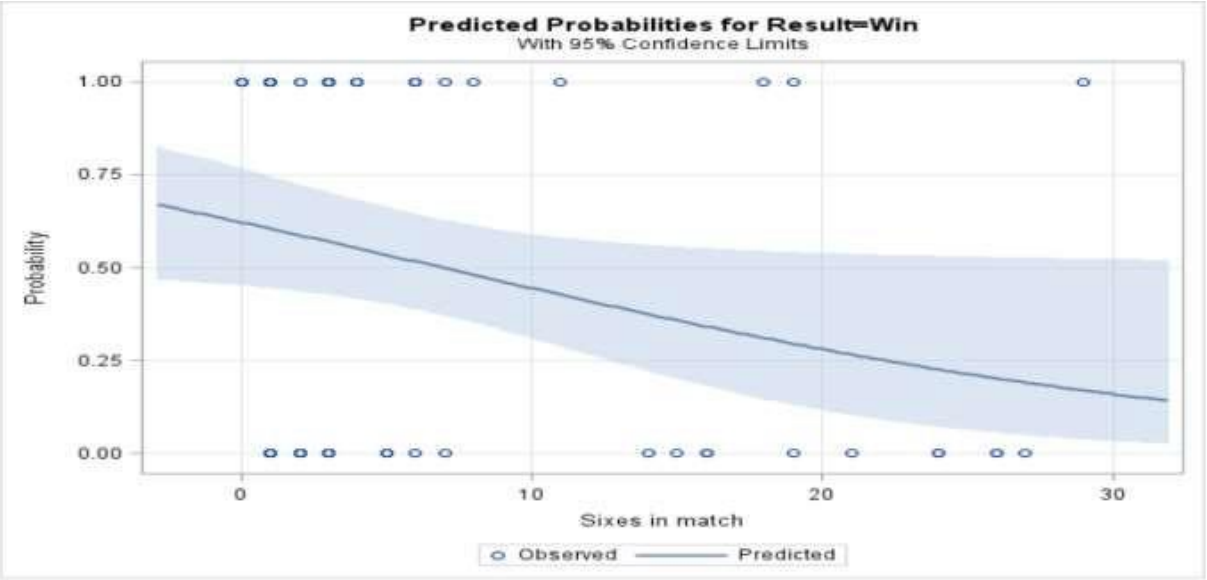

Figure 9. Probability graph for sixes scored when batting second

3.4.4 Batting order

Tables 1 and 2 represented the mean number of runs scored by batsmen who batted at different positions in the batting order. The average runs scored are represented for both teams batting first and second for winning and losing.

Teams BFW scored on average more runs (186) among the top four batsmen than the teams BFL (106.7), BSW (151.2), and BSL (115.2). Teams BFW scored on average more runs (88.6) amongst the middle order batsmen than the teams BFL (73.9), BSW (54.6), and BSL (66.3). Teams BFL scored on average more runs (35.3) amongst the lower order batsmen than the teams BFW (24.7), BSW (11.9), and BSL (25.1).

According to the univariate logistic regression the amount of runs scored by the top four batsmen is positively related to winning an ODI cricket match for teams batting first ($p=0.0002$). That is to say, the higher the number of runs scored by the top four batsmen in the ODI team batting first the higher the probability that team has of winning the cricket match as presented in figure 10.

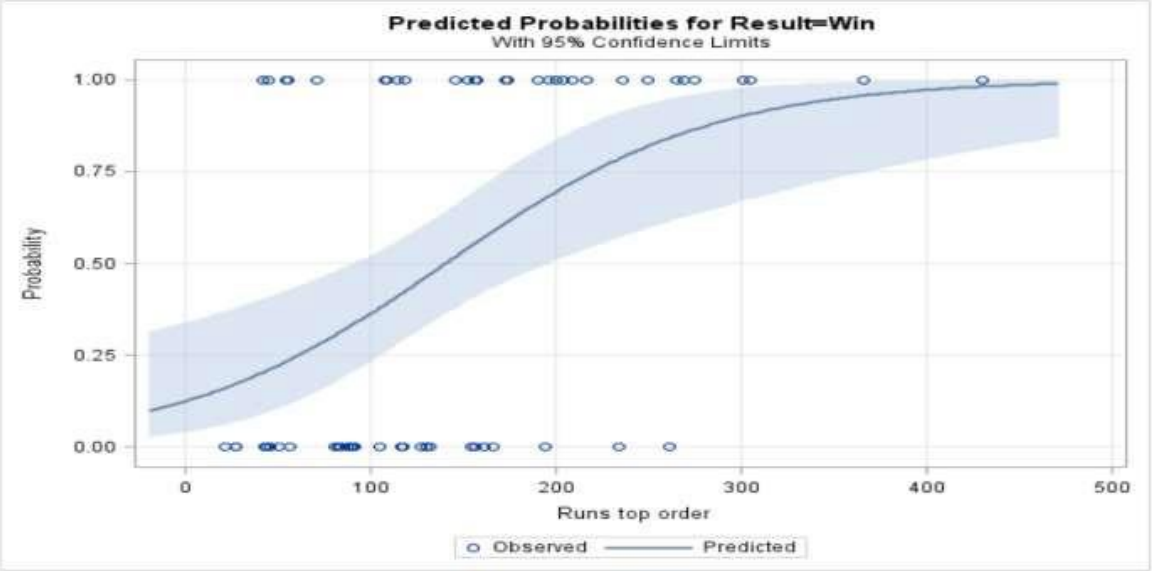

Figure 10. Probability graph for runs scored by the top four batsmen when batting first

The ODI cricket team BSW a cricket match is also positively related to the number of runs scored by the top four batsmen in the cricket team (p=0.0239). ODI cricket teams BSW is represented by Figure 11.

Figure 11. Probability graph for runs scored by the top four batsmen when batting second

An ODI cricket team BFW a cricket match is not significantly related to the number of runs scored by the middle three batsmen (p=0.3240) (Figure 12). This is also representative of ODI cricket teams batting second (p=0.3385) (Figure 13).

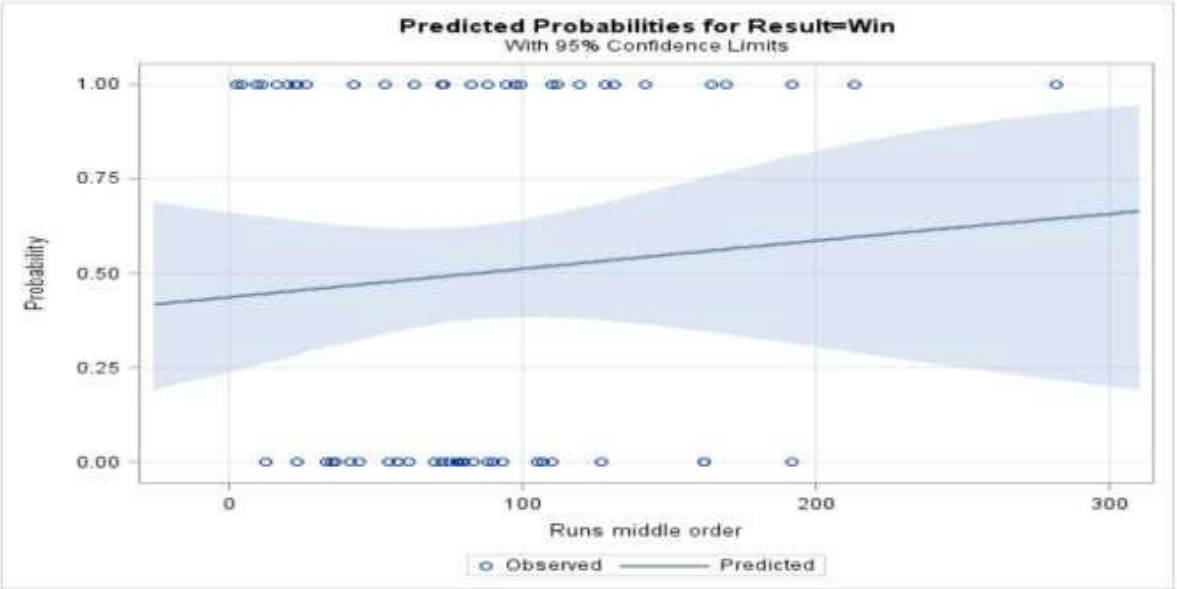

Figure 12. Probability graphs for runs scored by middle order batsmen when batting first

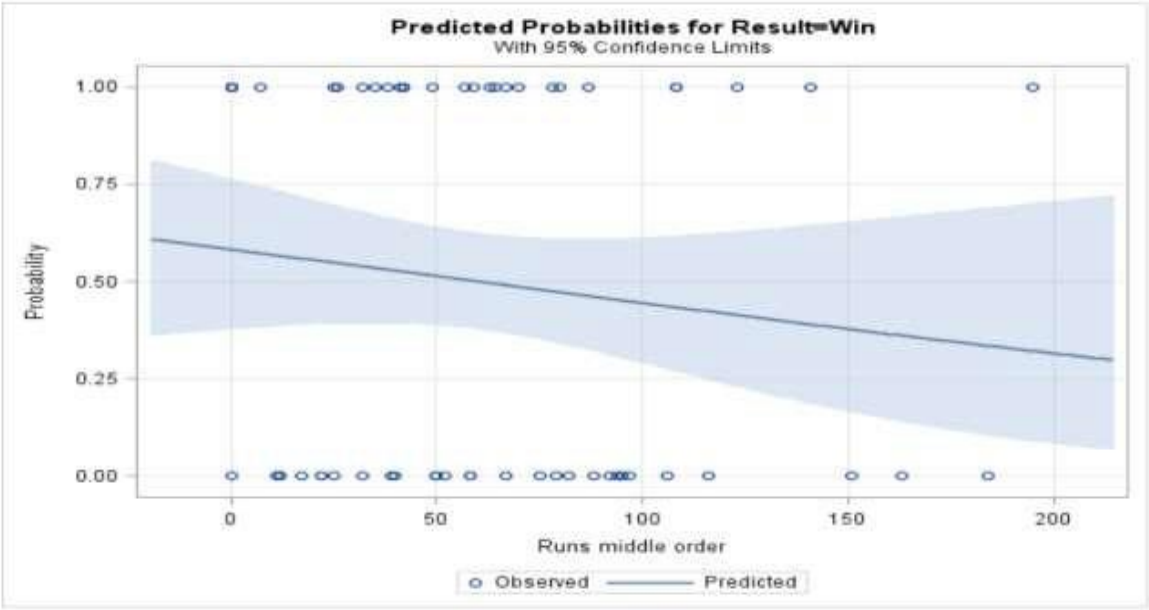

Figure 13. Probability graph for runs scored by middle order batsmen when batting second

According to the univariate logistic regression the amount of runs scored by the lower four batsmen is not significantly related to winning a match for teams batting first (p=0.1207) (Figure 14). However, the findings show that a team BSW a cricket match is negatively related to the number of runs scored by the lower four batsmen (p=0.0108). That is to say, the higher the number of runs scored by the lower four batsmen the lower the probability that that team has of winning the match. This decrease in probability of success is shown in Figure 15.

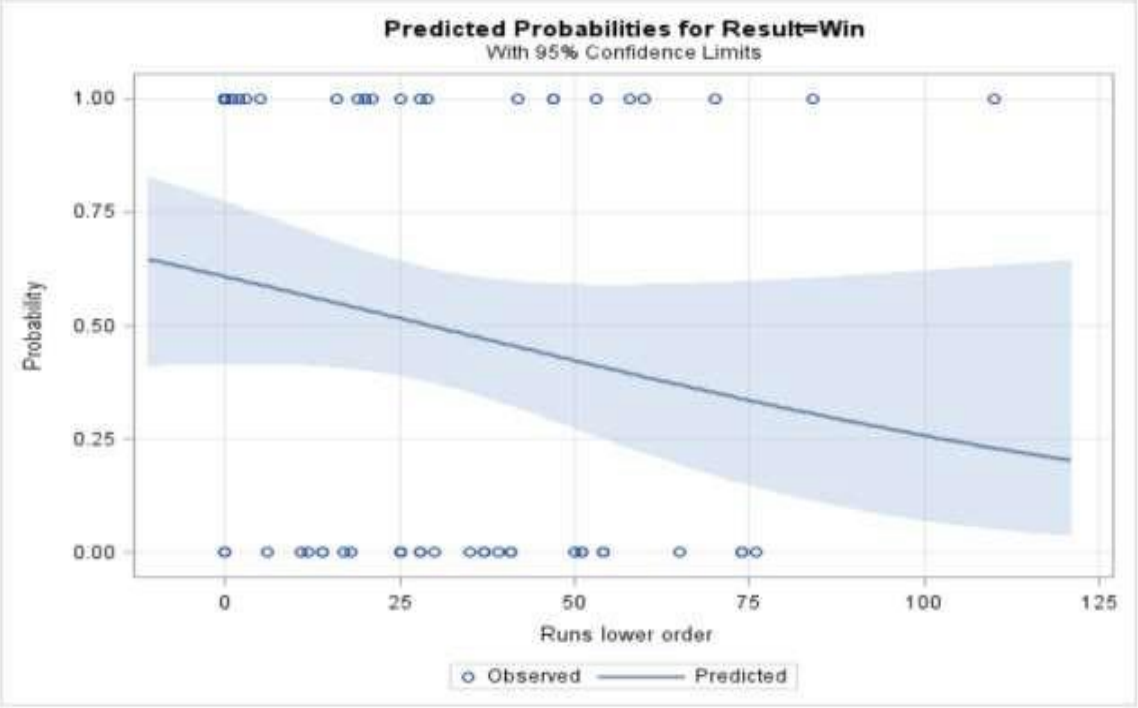

Figure 14. Probability graph for runs scored by lower order batsmen when batting first

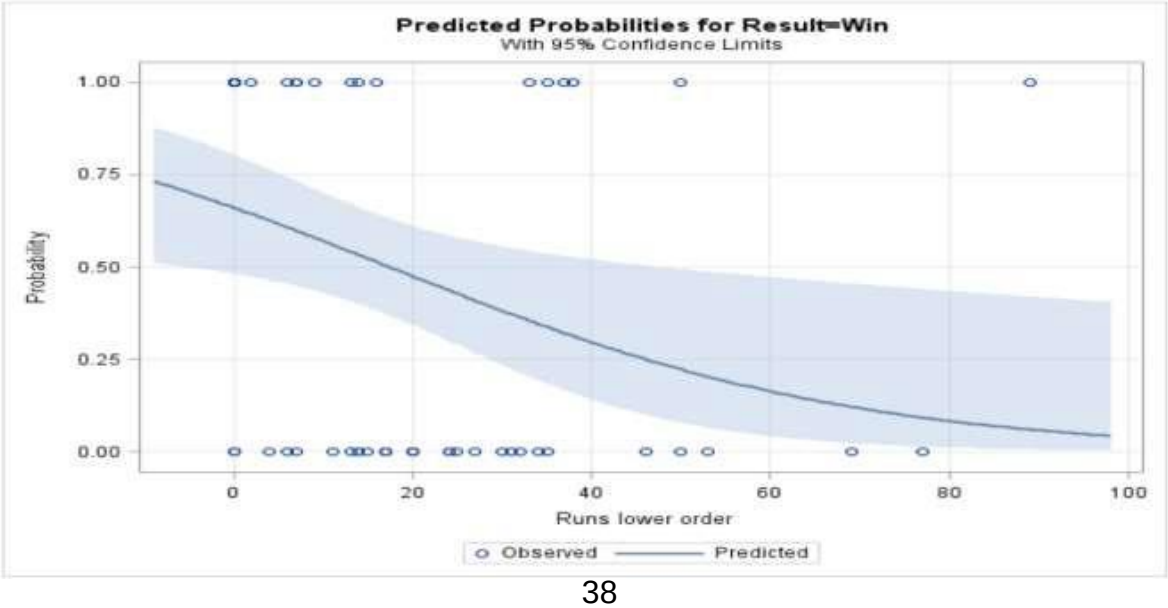

Figure 15. Probability graph for runs scored by lowered order batsmen batting second

3.5 Discussion

3.5.1 Batting first and last overs

The results of this study show that there are batting variables such as number of boundaries scored, the batting order, and the influence of the first 20 overs and last 12 overs in ODI cricket that significantly relate to the success of a cricket team in an ODI match. This study shows that variables that determine success in an ODI cricket match differ for teams who bat first and teams who bat second.

Norman and Clarke (2010) suggested that batsmen are content to survive the initial part of an ODI innings and take incremental steps towards riskier batting. The findings in this study show that the strategy for ODI teams batting first is to start the innings slowly without losing too many wickets in order to set up a good foundation for the later part of the innings. This is supported by the tendency of winning ODI cricket teams batting first to score on average a higher number of mean runs in the last 12 overs (108.4 runs) of a cricket match than in the first 20 overs (89.9 runs) of a cricket match. Batsmen who have been at the crease for a certain number of overs are more confident and can time the ball perfectly when trying to score runs quickly by hitting boundaries. The results of this study suggest that runs scored in the last 12 overs of a cricket match for teams batting first is very important for winning a cricket match. Therefore, the more runs scored in the last 12 overs, the higher the probability of success in an ODI cricket match.

3.5.2 Fours and sixes

The analysis shows that an increase in the number of sixes hit in a cricket match increases the probability of winning for ODI cricket teams batting first. On the other hand, for teams batting second, it is an increase in the number of fours scored which increases the probability of winning a match. This finding suggests that teams batting first have the freedom to play more aggressively and score more sixes as the teams batting first are not restricted by a runs total scored by the other team. Batting aggressively may be a risky strategy but can lead to bowlers being under pressure. Bowlers then tend to bowl more bad deliverers in their effort to get the batsmen out. This can lead to even more runs and sixes scored. Teams batting second

rely less on scoring sixes, but attempt to build a more consistent innings in order to reach the runs total scored by the team batting first. The data shows that the more sixes the team batting second scores the lower the probability that team has of winning a match. This may be due to teams batting second being too aggressive in stroke play and losing too many wickets too early. Teams batting second and losing may also have lost the match because the teams were chasing a high runs total. Therefore, losing teams batting second hit more sixes in the pursuit of the high runs total but ultimately fail to win the cricket match.

3.5.3 Batting Order

This study shows that the higher the number of runs scored by the top four batsmen in both teams batting first and second the higher the probability of winning the match. The data suggests that the responsibility of scoring runs as a part of the first four batsmen is very important to winning a match regardless of batting first or second. Norman and Clarke (2010) suggested that batting order should be open to a more strategic discussion during a cricket match. Each batsman and strengths and should move in the batting order based on if the current situation suits the batsman's strengths. This study suggests that the top four batsmen are the most important and should be prepared for every eventual situation. According to this study the higher the number of runs scored by the lower order batsmen in teams batting second the higher the probability of losing the ODI cricket match.

3.6 Conclusion

In this study data was analysed using a univariate logistic regression in order to determine whether or not scoring runs in the first 20 and last 12 overs, the number of fours and sixes scored, and the number of runs scored between the different batsmen are significantly related to winning a cricket match.

The results of this study show that there are variables such as number of boundaries scored, the batting order, and the influence of runs scoring in the first 20 and last 12 overs in cricket that relate positively with success in ODI cricket and can be used as success predictors. These success predictors differ between teams batting first and second. In contrast to other studies, this study focusses on variables in ODI cricket that reflect the performance of the entire team. From this study we can conclude that the top order (four) batsmen are in most cases more important to winning the cricket match than others. Further research into the roles of the

significant variables of the current study is warranted. The practical application of the results of the current study is discussed in the following section.

3.7 Practical application

The significance of scoring runs and more specifically scoring sixes for the team batting first suggests that being more aggressive in stroke play and general batting approach when batting first puts the bowlers under pressure. This in turn puts the bowling team's batsmen under pressure even before they have their chance to chase the runs total. This study highlights the importance of the batting order in ODI cricket. It could be argued that ODI cricket teams may only need to employ the services of four specialist batsmen and thereafter complete the team with all-rounders and bowlers. An argument can also be made that if an ODI cricket team's wicket-keeper does not bat in the top four, that batsmen is surplus to demand in the cricket team. The information derived from this study implies that a strategy of batting aggressively when batting first is advantageous. For teams batting second, the strategy seems to be a consistent well-structured run chase where the batsmen are conservative in stroke selection.

3.8 References

AMIN, G. R. & SHARMA, S. K. (2014b). Measuring batting parameters in cricket: a two stage regression- OWA method. *Measurement,* 53: 56-61.

BARR, G.D.I. & KANTOR, B.S. (2004). A criterion for Comparing and Selecting Batsmen in Limited Overs Cricket. *The Journal of the Operational Research Society*, 55(12): 1266-1274.

DOUGLAS, J.M. & TAM, N. (2010). Analysis of team performances at the ICC World Twenty20 Cup 2009. *International Journal of Performance Analysis in Sport*, 10: 47-53.

GILL, P. S., SWARTZ, T. B., BEAUDOIN, D. & DE SILVA, B. M. (2006). Optimal batting orders in one-day cricket. *Computers & Operations Research*, 33: 1939-1950.

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.

LEMMER, H.H., BHATTACHARJEE, D. & SAIKIA, H. (2014). A consistency Adjusted Measure for the Success of Prediction Methods in Cricket. *International Journal of Sports Science & Coaching*, 9(3): 497-512.

MUKHERJEE, S. (2013). Complex analysis network in cricket: community structure, players role and performance index. *World Scientific*, 16(8): 1-20.

MUKHERJEE, S. (2014). Quantifying individual performance in Cricket — A network analysis of batsmen and bowlers. *Physica A*, 393: 624-637.

MUNIR, F., HASAN, K., AHMED, S. & QURAISH, S. (2015). Predicting a T20 cricket match result while the match is in progress. A dissertation presented for the degree of Bachelor in Computer Science.

NORMAN, J. M. & CLARKE, S. R. (2010). Optimal batting orders in cricket. *The Journal of the Operational Research Society*, 61: 980-986.

PETERSON, C. J., PYNE, D. B., PORTUS, M. R. & DAWSON, B. T. (2011). Comparison of player movement patterns between 1-day and test cricket. *Journal of strength and conditioning research*, 25(5): 1368-1373.

PRESTON, I. & THOMAS, J. (2000). Batting strategy in limited overs cricket. *The Statistician*, pp. 95-106.

RUDKIN, S. T. & O'DONOGHUE, P. G. (2008). Time-motion analysis of first-class cricket. *Journal of Science and Medicine in Sport*, 11: 604-607.

SAS INSTITUTE INC (2013). *SAS/STAT 13.1 User's Guide*. Cary, NC: SAS Institute Inc.

STRETCH, R. A., BARTLETT, R. & DAVIDS, K. (2000). A review of batting in men's cricket. *Journal of sport sciences,* 18: 931-949.

STUELCKEN, M. C., PORTUS, M. R. & MASON, B. R. (2005). Off-side front foot drives in men's high performance cricket. *Sports Biomechanics*, 4(1): 17-36.

Chapter 4: Article 2: Predictors of batting success for winning and losing teams in Twenty-Twenty International cricket

---

To be submitted to The *African Journal for Physical Activity and Health Sciences* (AJPHES) (Appendix B). Note that the referencing and formatting guidelines of the journal have been followed with the exception of the placement of tables and figures. While the journal requires these to be attached to the end of the article when submitting, they are included in text for ease of reading of the dissertation.

Abstract

Predictors of batting success for winning and losing teams in Twenty-Twenty International cricket

[1]Mark Schaefer, [1]Riaan Schoeman, [2]Robert Schall,
[1]Department of Exercise and Sport Sciences, [2]Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa

The aim of this study was to describe batting variables in T20I cricket and identify predictors that discriminate between winning and losing batting teams. Understanding the batting variables that predict the success of a T20I cricket team will aid coaching staff in team selection, batting order, and overall match strategy. Match data from the 2014 and 2015 T20I cricket season was recorded from cricinfo. A total of 60 T20I cricket matches were used in the study. Significant predictors of winning a T20I cricket match when batting first were: runs scored in the first 5 overs (p=0.0035), runs scored in the last 7 overs (p=<0.0001), and sixes scored (p=0.0081); similarly, significant predictors for winning a T20I cricket match when batting second were: Runs scored in the first 5 overs (0.0046) fours scored (p=0.0258), runs scored between the top order batsmen (p=0.0034), and runs scored between the lower order batsmen (p=0.0043). Variables such as the number of boundaries scored, the batting order, and the influence of runs scoring in the first 5 and last 7 overs that predict success in T20I cricket differed for teams batting first and second, respectively.

Keywords: Twenty-Twenty; Predictors; Winning and Losing

4.1 Introduction

Modern society has become a more fast-paced results-driven lifestyle. The younger generation of sports spectators have pressured the ICC to innovate a more audience-focused approach to cricket in later years (Cannonier, Panda, & Sarangi, 2015). T20I cricket was introduced in 2003 in an attempt to hold the attention of modern society. Not only did T20I cricket hold the attention of current cricket fans but it also drew in a new type of supporter. Younger people both male and female have come to enjoy the fast-paced spectator sport that is T20I cricket. To make T20I more result driven, when a match is tied at the end of the two batting innings the two teams break the tie with a super over or bowl off. Some rules such as the no-ball free hit rule in T20 cricket has crossed formats to ODI cricket perhaps in an attempt to make ODI cricket more entertaining.

Like ODI cricket, T20I cricket is played between two nations consisting of 11 players each over limited overs. T20I cricket is played over 20 overs. T20I is by far the shortest version of international cricket. This shortened version of cricket played at the highest level of competitiveness makes for a great viewing spectacle. Cannonier, Panda, and Sarangi (2015) suggest that T20I cricket rules have been adjusted to provide batsmen with certain advantages over batsmen in ODI cricket. Norman and Clarke (2010) suggested that cricket pitches are prepared in favour of batsmen for T20 and ODI matches. Batting in T20I is seen by supporters to be the main attraction of cricket and drives the result of cricket. Hughes and Bartlett (2010) state that a using batting as a measure of performance in a cricket match is difficult in the absence of a measure of bowling as a performance variable. In a contradictory study Shah *et al*., (2015) suggested that because the found variance of batting in both ODI and T20 cricket was much higher than that of bowling, batting should be emphasised as the more important success predictor.

Bhattacharjee and Lemmer (2016) describe batting strategy in limited overs cricket as an attempt to reach a runs total before depleting all wickets and overs on hand (before all batsmen are dismissed or 20 overs have passed). This definition gives the impression of a cautious and calculated approach to batting. However, T20I cricket is a sometimes reckless in an attempt to reach the highest runs total possible regardless of wickets fallen or in hand. Entertaining the crowd in T20 cricket is paramount whilst winning the match is perhaps slightly less important. Davis, Perera and Swartz (2015) agree that T20I batsmen behave

more aggressively throughout a T20 match. It is suggested that batsmen are able to do this as the 10 available wickets are enough to utilise aggressively in 20 overs.

With the rise of T20I cricket and the differences in strategy, team selection and batting order would also change. Norman and Clarke (2010) stated that with the introduction of limited overs cricket especially T20 cricket the rate at which batsmen scored runs is becoming more important. This idea was also held by Douglas and Tam (2009) as well as Irvine and Kennedy (2017). Not only is the number of runs a batsman scores on average important but so too is how fast the batsmen could score those runs. This became an important aspect of team and batting order section. Batting slowly in a T20I cricket match puts pressure on the batting team and forces an increase in aggressive batting (Bhattacharjee & Lemmer, 2016). This leads to increased risk of dismissal and losing the match. Batsmen that could score runs quickly without much time to adjust to batting conditions are selected more for T20I cricket and elevated up the batting order (Irvine & Kennedy 2017). The shift in thinking what makes a good batsman effective with the introduction of T20I cricket can be seen by statements made in the study by Lemmer (2011). The author states that when two batsmen have scored the same amount of runs the one that has scored the runs faster has performed better.

The aforementioned studies cumulatively agree that T20I cricket is won by batsmen batting aggressively throughout the match. Scoring runs by means of fours or sixes is the main objective for batsmen in T20I cricket. T20I cricket is still relatively new compared to test match and ODI cricket. More research is required to determine the most effective batting strategy, batting order, and team selection. This study investigates certain variables in an effort to identify possible success predictors of T20I cricket.

4.2 Methodology

4.2.1 Sample

A total of 60 professional T20I cricket matches were captured resulting in 120 records (two innings per match). Ten teams were selected to be captured for the purpose of this study, namely South Africa, Australia, New Zealand, India, Bangladesh, Zimbabwe, England, Sri Lanka, Pakistan, and West Indies. These teams all participate in all three formats of cricket

namely test, ODI, and T20I cricket. Six matches from each team's records were randomly selected and observed. Of the six matches, three matches were won by the team batting first, and three matches were won by the team batting second. Drawn matches, and those which employed the Duckworth-Lewis method were excluded from the study. Data was captured for all performance variables of concern in this study.

4.2.2 Data collection procedure

A quantitative study with an observational design by means of retrospective data analysis was used to determine batting performance variables that correlate the highest with the success of a team in T20I cricket. Data from the years 2014 and 2015 international cricket season was collected from ESPN Cricinfo. All data was captured in Microsoft Excel 2007 and subsequently converted into a SAS data set. Data was statistically analysed to evaluate and compare the variables during the 2014 and 2015 season for winning and losing T20I cricket teams.

The following variables were analysed for the purpose this study: runs scored in the first five overs of a T20I cricket match, runs scored in the last seven overs of a T20I cricket match, the number of fours and sixes scored in a T20I cricket match, and runs scored by the top four, middle three, and bottom four batsmen in a T20I cricket match to establish batting order contributions. In this research, a strong and reliable data source was found in Statsguru. Statsguru is ESPN Cricinfo's cricket statistics maintenance database. Cricinfo being used in professional cricket as well as being referenced by published authors. In this database, all the match's data are stored with live ball by ball commentary (Munir, Hasan, Ahmed & Quraish, 2015).

4.3 Statistical analysis

Data obtained from the Twenty-Twenty International cricket matches was captured in Microsoft Excel. The data was then analysed using the SAS statistical software (SAS, 2013).

Because of the fundamentally different match situation faced by the team batting first and second, respectively, the data were analyses separately for the team batting first and for the

team batting second. The match outcome is used as a binary variable (win/lose). The association of the potential predictor variables with the match outcome was analysed using univariate logistic regression, fitting each predictor variable, one at a time. The statistical significance of each predictor variable was tested using an exact test (exact conditional logistic regression); the exact P-value is reported. Furthermore, an odds ratio and associated 95% confidence interval is reported which reflects the effect (that is, the increase in the odds of winning) associated with an increase of one unit of the predictor variable. The analysis was carried out using SAS procedure LOGISTIC (SAS, 2013).

4.4 Results

4.4.1 Batting first and second

Table 5 and Table 6 represent the mean values of each potential predictor variable, for the winning and losing teams respectively; Table 1 presents the data for teams batting first and Table 6 for teams batting second. The p-value from a univariate logistic regression is reported. Table 5 represents the importance of scoring well in the beginning and at the end of a T20I innings. This is made evident by both runs scored in the first five and last seven overs being significant. Scoring sixes is also significant whilst scoring fours is not. Table 6 indicates that scoring runs in the initial part of the batting innings is more important than scoring runs at the end of the innings. Runs scored by the top order batsmen is also significant further showing that the initial part of the inning batting second is important. Runs scored by the lower order is negatively correlated with winning.

Table 5. Team batting first: mean values of potential predictors for winning and losing teams

Team potential predictor of success	Means		p value
	Win	Lose	
Runs scored in the first 5 overs	36.8	29.8	0.0035*
Runs scored in the last 7 overs	67.1	51.5	<0.0001*
Fours scored	9.7	9.3	0.7559
Sixes scored	11.2	7.2	0.0081*
Runs scored by the top order batsmen	116	96.3	0.0661
Runs scored by the middle order batsmen	43.8	35.1	0.1806
Runs scored by the lower order batsmen	6.8	7.9	0.7095

Note* Statistically significant p-value from logistic regression analysis

Table 6. Team batting second: mean values of potential predictors for winning and losing teams

Team potential predictor of success	Means		p value
	Win	Lose	
Runs scored in the first 5 overs	40.8	33.4	0.0046*
Runs scored in the last 7 overs	44.8	47.1	0.6534
Fours scored	10.6	7.3	0.0258*
Sixes scored	8.8	9.2	0.8211
Runs scored by the top order batsmen	110.9	81.1	0.0034*
Runs scored by the middle order batsmen	28.7	40.3	0.0830
Runs scored by the lower order batsmen	5.4	14.4	0.0043*

Note* Statistically significant P-value from logistic regression analysis

Tables 7 and 8 present the results of the statistical analysis completed using a univariate logistic regression. Table 7 represents the results for teams batting first. Table 8 represents the results for teams batting second. Table 7 represents sixes scored as the highest odds ratio. For every increase in sixes scored by one six the odds of winning a T20I cricket match increases by 1.142. Table 8 shows the runs scored in the last seven overs of a match, has the highest odds ratio. In terms of an increase in odds of winning a match, scoring runs in the last seven overs is the most important.

Table 7. Univariate logistic regression: Predictors of match outcome T20I data; Team batting first

Predictor Variable	[a]95% CI	Odds ratio	p Value	Test Statistic[b]
Runs scored first 5 overs	1.026 to 1.166	1.087	0.0035	8.3022
Runs scored last 7 overs	1.038 to 1.136	1.080	<0.0001	14.3817
Fours scored	0.924 to 1.121	1.017	0.7559	0.1141
Sixes scored	1.036 to 1.275	1.142	0.0081	6.9451
Runs scored by the top order batsmen	0.999 to 1.027	1.012	0.0661	3.3787
Runs scored by middle order batsmen	0.994 to 1.039	1.015	0.1806	1.8274
Runs scored by lower order batsmen	0.944 to 1.038	0.991	0.7095	0.1548

Note:[a]Profile likelihood confidence interval; [b]Exact conditional score test

Table 8. Univariate logistic regression: Predictors of match outcome T20I data; Team batting second

Predictor Variable	[a]95% CI	Odds ratio	p Value	Test Statistic[b]
Runs scored first 5 overs	1.030 to 1.174	1.095	0.0046	7.6761
Runs scored last 7 overs	1.966 to 1.021	1.994	0.6534	0.2124
Fours scored	0.061 to 1.245	1.117	0.0258	4.9771
Sixes scored	0.902 to 1.082	0.988	0.8211	0.0628
Runs scored by the top order batsmen	1.007 to 1.037	1.021	0.0034	8.3362
Runs scored by middle order batsmen	0.960 to 1.002	0.982	0.0830	3.0217
Runs scored by lower order batsmen	0.879 to 0.979	0.933	0.0043	7.7742

Note:[a]Profile likelihood confidence interval; [b]Exact conditional score test

4.4.2 Runs scored in the first five and last seven overs

Teams batting second and winning (BSW) scored on average the more runs (M=40.8) in the first five overs of an innings. Teams batting first and winning (BFW) scored on average the more runs (M=67.1) in the last seven overs of an innings.

According to the univariate logistic regression runs scored in the first five overs are positively related to winning the match for teams batting first (p=0.0035). The higher the number of runs scored in the first five overs by the T20I team batting first the higher the probability that that team has of winning the match. This increase in probability of success is shown in Figure 16. The increase in probability of success as a function of the runs scored in the first five overs for teams batting second - shown in Figure 17 shows that this variable is also significantly related to winning (p=0.0046).

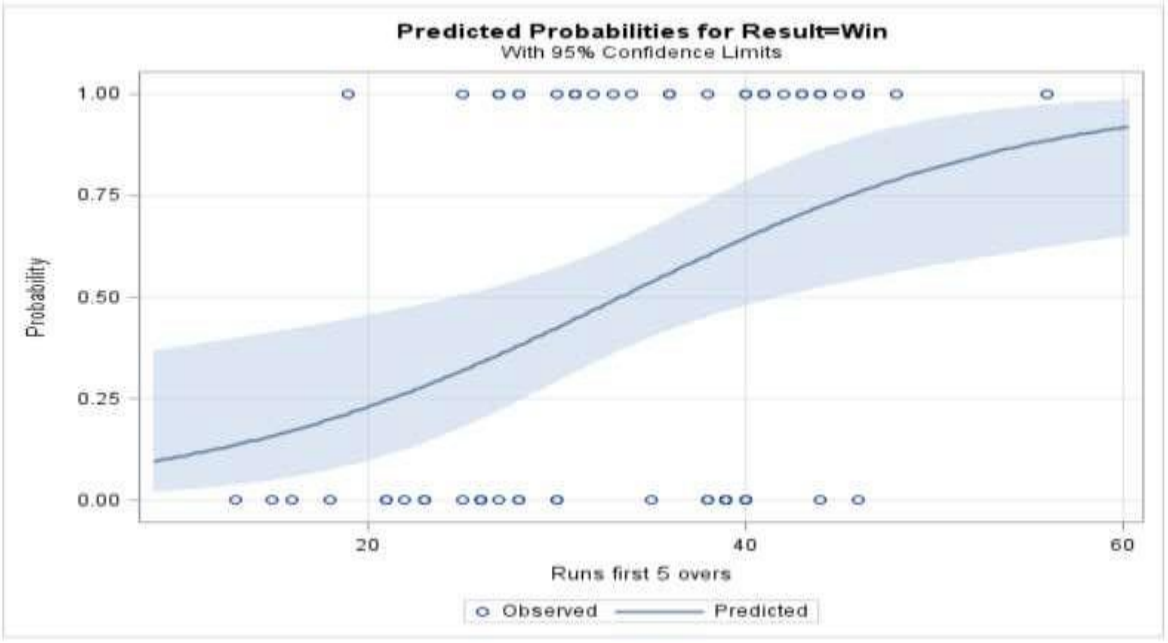

Figure 16. Probability graph for runs scored in the first five overs when batting

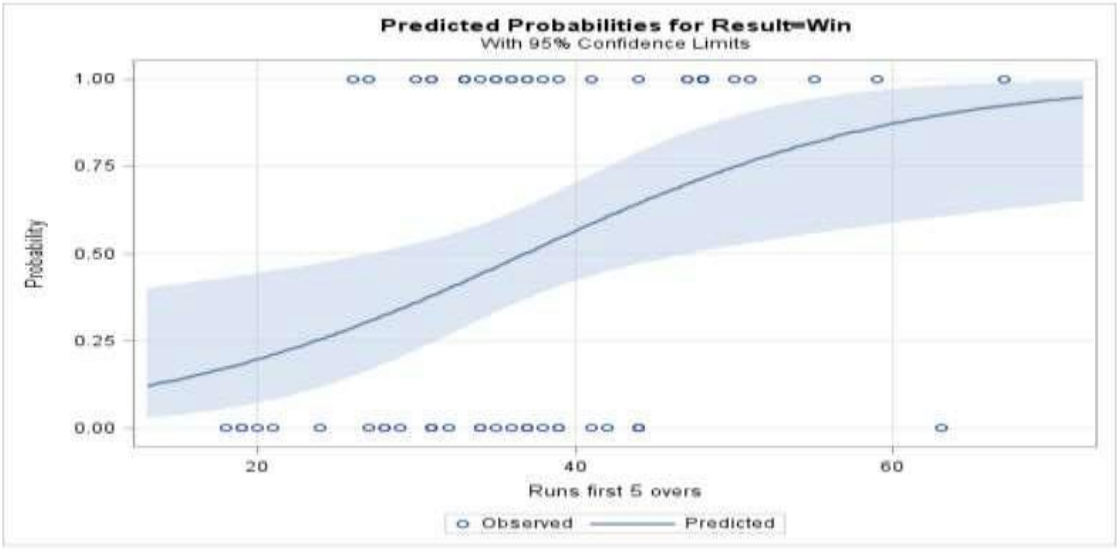

Figure 17. Probability graph for runs scored in the first five overs when batting second

According to the univariate logistic regression runs scored in the last seven overs is positively related to winning the match for teams batting first (p=<0.0001). For every run scored in the last seven overs of a match the higher the probability of winning the match. Figure 18 represents this resulting increase in probability of success for teams batting first. Runs scored

in the last seven overs of a match is not significantly related to winning the match for the team batting second (p=0.6543) (figure 19).

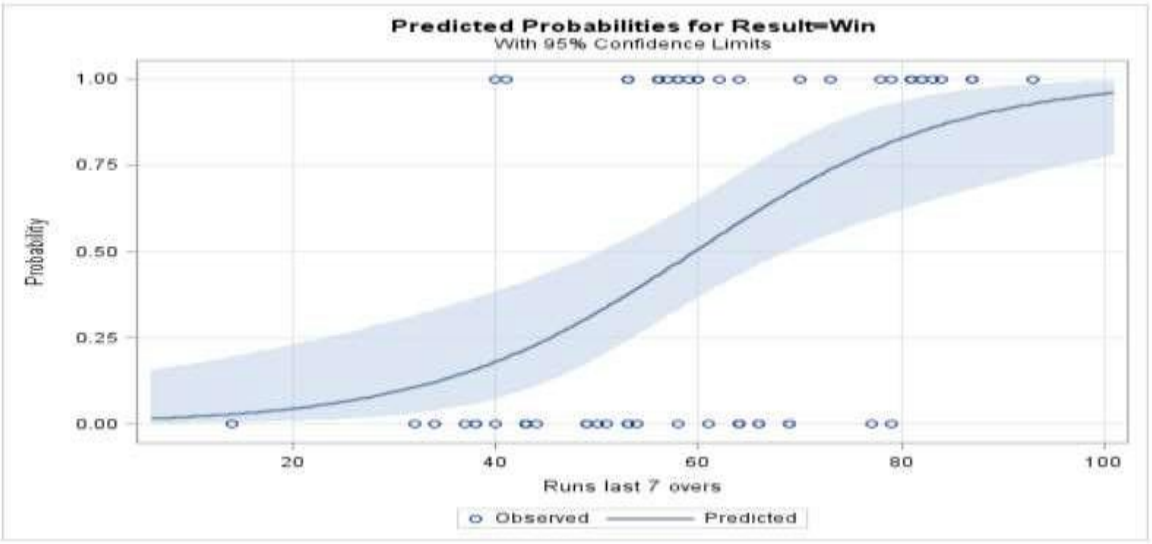

Figure 18. Probability graph for runs scored in the last seven overs when batting first

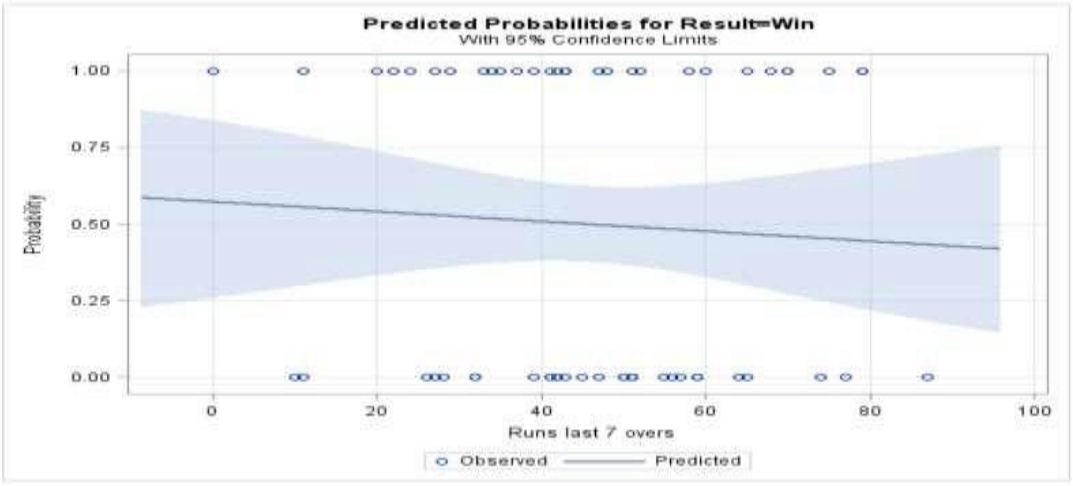

Figure 19. Probability graph for runs scored in the last seven overs when batting second

4.4.3 Boundaries

Tables 5 and 6 represent the mean number of different boundaries scored during the match for both winning and losing teams. Winning teams batting first (9.7) and second (10.6), score on average more fours than the losing teams batting first (9.2) and second respectively (7.3). As reported in tables 5 and 6 the mean number of sixes scored (11.2) during the recorded cricket matches is much higher for BFW than it is for teams batting first and losing (BFL)

(7.2), batting second and wining (BSW) (8.8), and teams batting second and losing (BSL) (9.2).

According to the univariate logistic regression, the number of fours scored is positively related to winning the match for teams batting second (p=0.0258). Figure 21 represents the increase in the probability of success. The univariate logistic regression shows no significant relation between winning a match and the number of fours scored by the team batting first (Figure 20) (p=0.7559). The number of sixes scored throughout a match is positively related to winning a T20I cricket match for teams batting first (p=0.0081). That is to say, the higher the number of sixes hit by the T20I team batting first the higher the probability that that team has of winning the cricket match. This increase in the probability of success is shown in Figure 22. The amount of sixes scored is not significantly related to winning a for teams batting second (p=0.8211) (Figure 23).

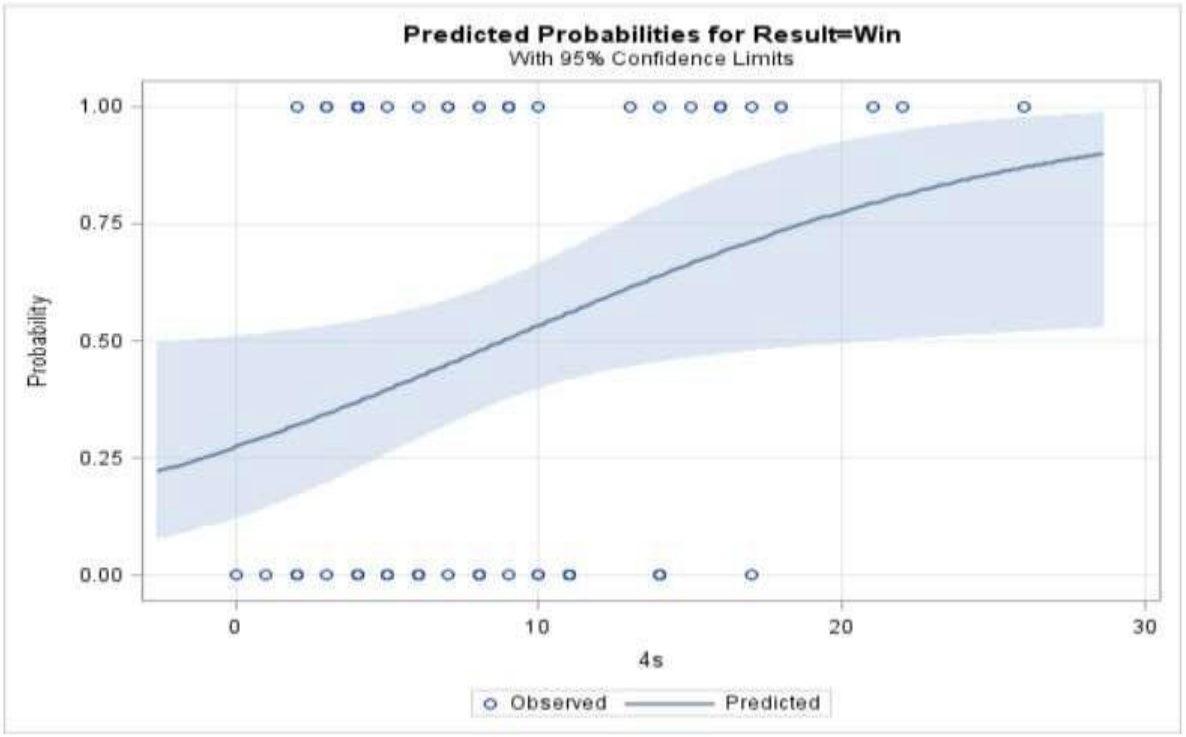

Figure 20. Probability graph for fours scored when batting first

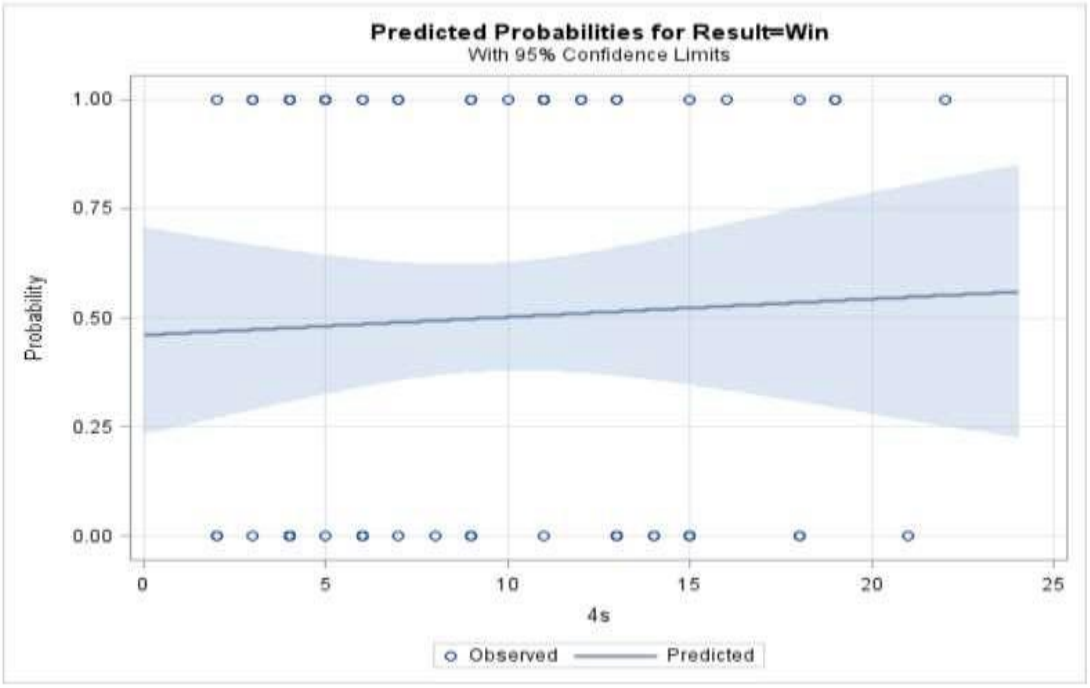

Figure 21. Probability graph for fours scored when batting second

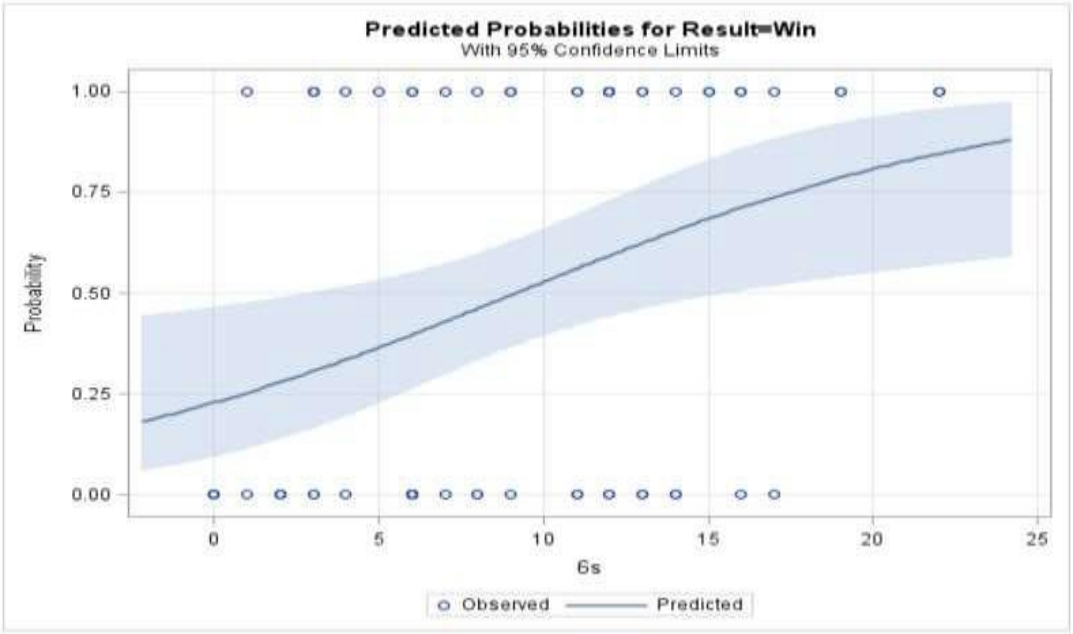

Figure 22. Probability graph for sixes scored batting first

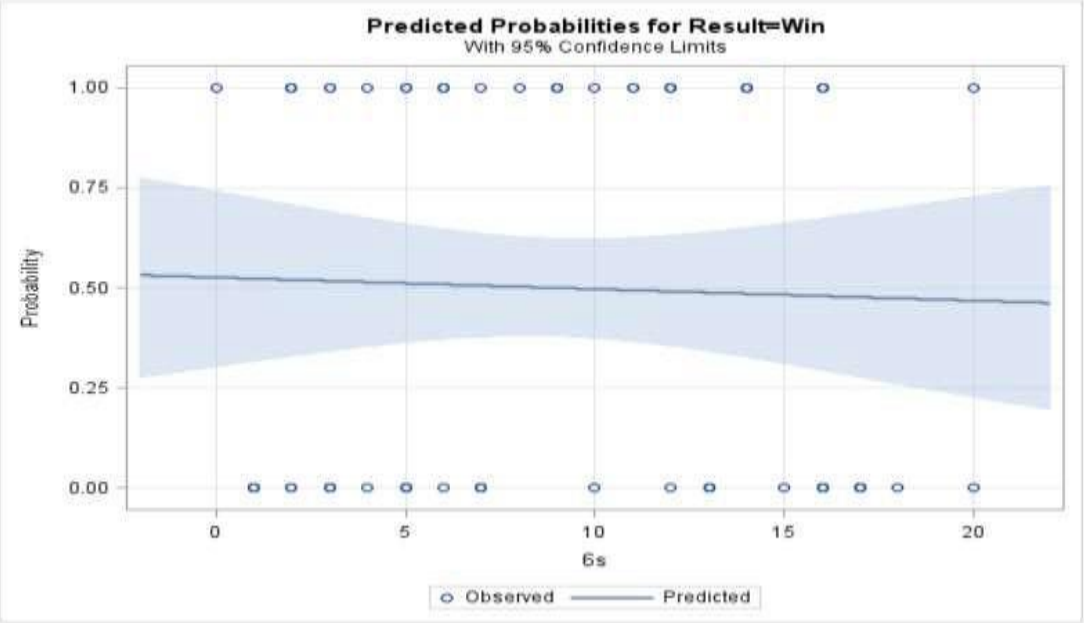

Figure 23. Probability graph for sixes scored when batting second

4.4.4 Batting order

Tables 5 and 6 represented the mean number of runs scored by batsmen who bat at different positions in the batting order. The average runs scored are represented for both teams batting first and second for winning and losing teams.

Teams BFW scored on average more runs (116) among the top four batsmen than the teams BFL (96.3), BSW (110.9), and BSL (81.1). Teams BFW scored on average more runs (43.8) amongst the middle order batsmen than the teams BFL (35.1), BSW (28.7), and BSL (40.3). Teams BSL scored on average more runs (14.4) amongst the lower order batsmen than the teams BSW (5.4), BFW (6.8), and BFL (7.9).

According to the univariate logistic regression the number of runs scored by the top four batsmen (p=0.0661) (Figure 24), middle three batsmen (p=0.1806), and lower four batsmen (p=0.7095) is not significantly related to winning a match. According to the univariate

logistic regression the amount of runs scored by the top order batsmen batting second (p=0.0034) is significantly related to winning a T20I cricket match. That is to say that the higher the number of runs scored between the top four batsmen batting second the higher the probability of winning the match. The increase in probability is shown by figure 25. The results of Figure 26 and 27 show that runs scored by the middle three batsmen (p=0.0830) is not significantly related to winning a match for teams batting second. However, the findings show that a team batting second in a cricket match is negatively related to the number of runs scored by the lower four batsmen (p=0.0043) (figure 28). The higher the number of runs scored by the lower four batsmen the lower the probability that that team has of winning the cricket match. This decrease in the probability of success is shown in Figure 29.

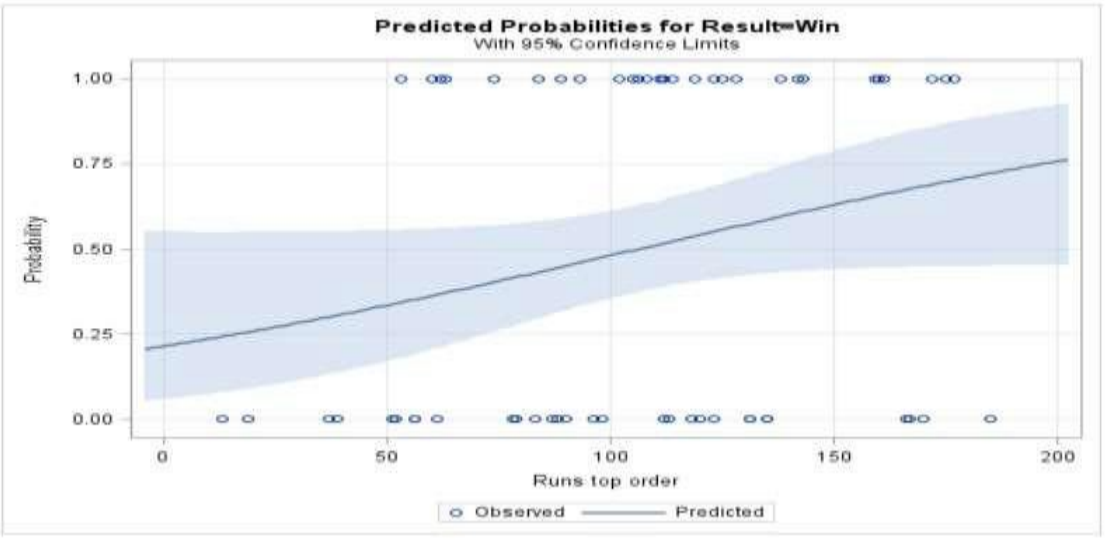

Figure 24. Probability graph for runs scored by the top four batsmen when batting first

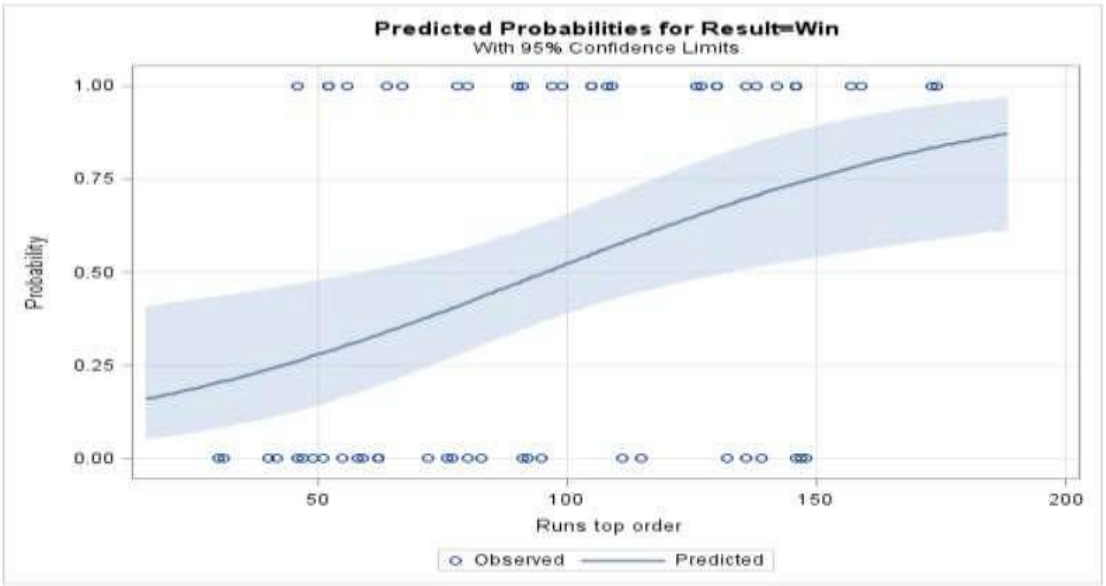

Figure 25. Probability graph for runs scored by the top four batsmen batting second

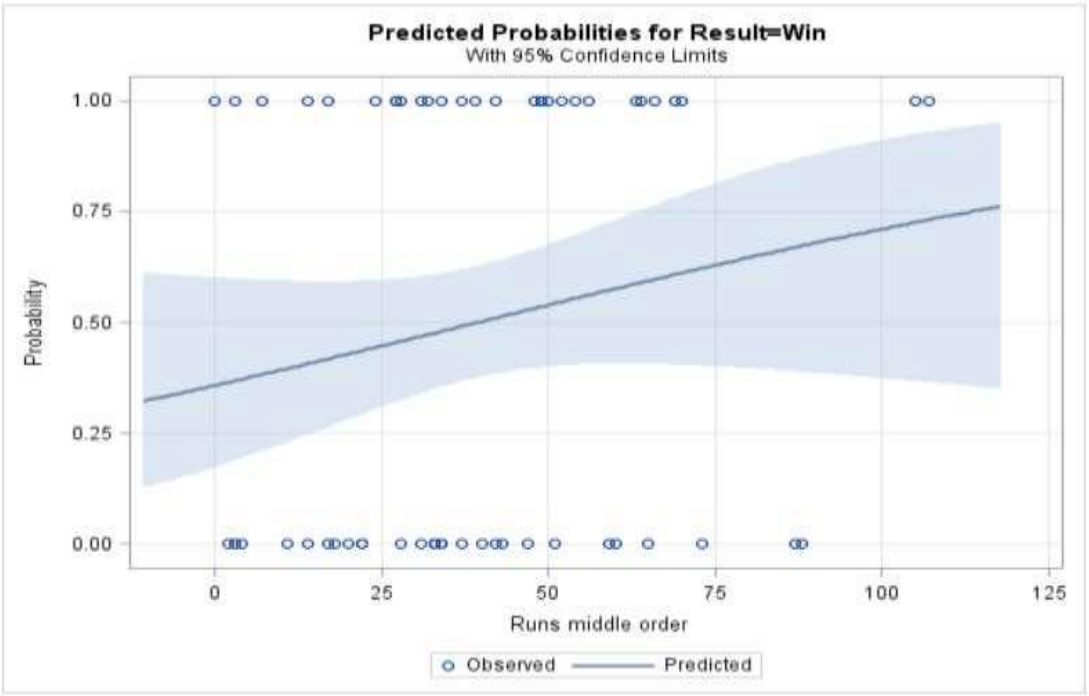

Figure 26. Probability graph for runs scored by the middle order batsmen when batting first

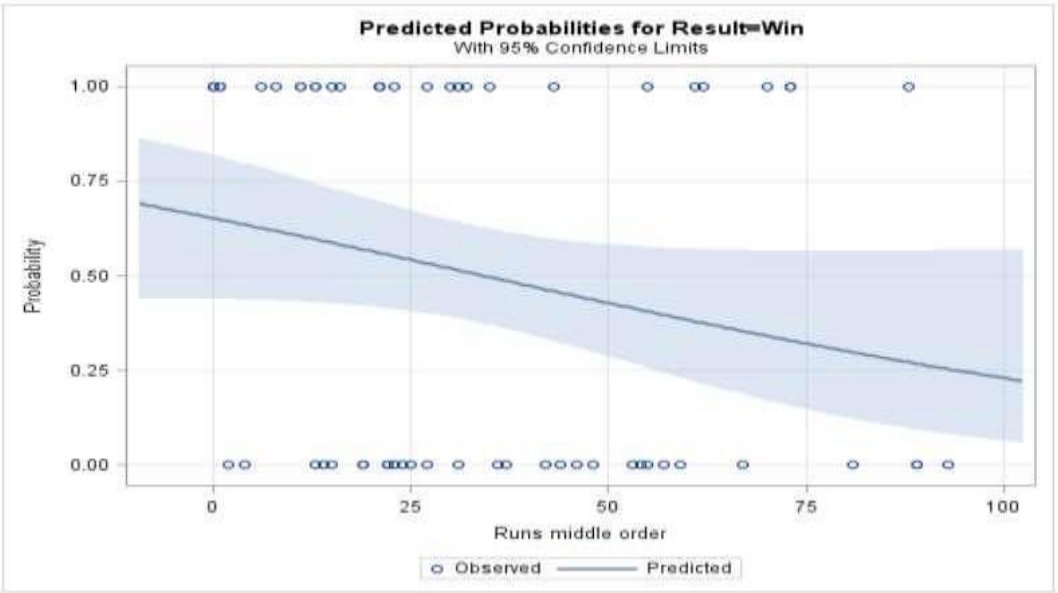

Figure 27. Probability graph for runs scored by the middle order batsmen when batting second

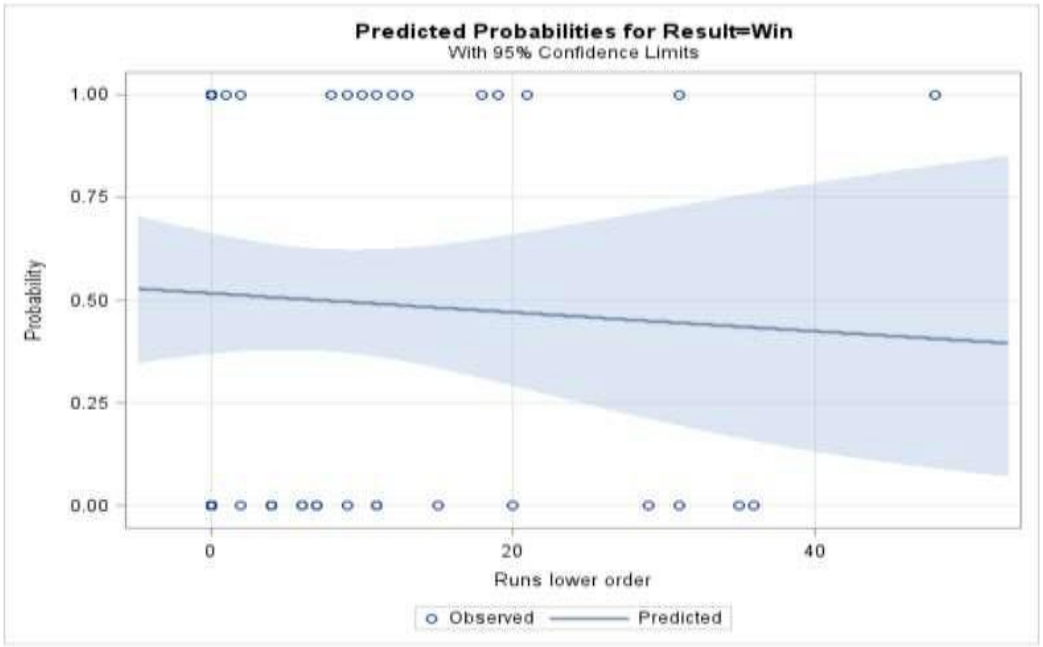

Figure 28. Probability graph for runs scored by the lower order when batsmen batting first

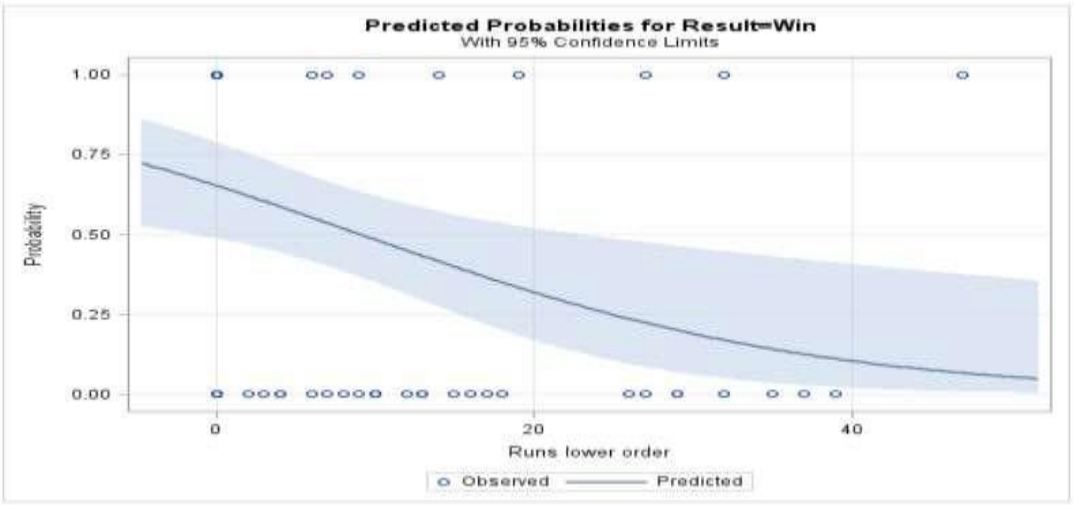

Figure 29. Probability graph for runs scored by the lower order batsmen when batting second

4.5 Discussion

4.5.1 Batting first and last overs

The results of this study indicate that there are batting variables in T20I cricket that correlate with winning a match. Certain variables are significant for both teams batting first and second whereas certain variables are significant for teams batting first and not for batting second vice versa. The statistical differences between teams batting first and second area evidence that a different strategy should exist to win a match batting first and second.

The results section revealed scoring runs in the first five overs of a T20I cricket match is important for both teams batting first and second. Scoring more runs in the initial part of the innings and having a good starting foundation greatly improves a team's probability of winning a match. This is true for both teams batting first and second. According to the current study, the batting strategy in the first five overs of a match should be to score as many runs as possible as it sets up a good foundation to win the match. This is especially true for teams batting second as indicated by Bhattacharjee and Lemmer (2016). The current study found that when batting first it is important to remain aggressive throughout a match but for the team batting second is seems to be a better strategy to score fast in the initial part of the batting innings and then bat more carefully towards the end of the innings to ensure victory. Davis, *et al.*, (2015) suggest that constant aggression when batting in a T20 is the best strategy. This further highlights the need for research that examines the differences between teams batting second and first.

The results of the study suggest that for teams batting first runs scored in the last seven overs of a match is more important than runs scored in the first five overs of the match. The strategy for teams batting first in a T20I may be to score moderate runs more carefully in the first five overs and to score as fast as possible with no regard for being dismissed in the last seven overs.

4.5.2 Fours and sixes

For teams batting first scoring more sixes in the match increases the probability of success. For teams batting second scoring more fours in a match increases the probability of success. This suggests that it is more important for teams batting first in a T20I cricket match to be aggressive in the type of shots they play in order to score as much runs during their innings as possible leading them to score more sixes on average. For teams batting second in a match, it may be more important to play safely whilst batting which would mean that scoring fours and not sixes is the priority.

4.5.3 Batting order

The results revealed that runs scored by the top, middle, and lower order batsmen is of no importance to teams batting first. According to this study batting order strategy for teams batting first is at this moment irrelevant.

For teams batting second scoring runs amongst the top order batsmen of significance and an important aspect of a strategy to win the match. Interestingly the more runs scored amongst the lower four batsmen decreases the probability of winning a match. This finding indicates that scoring runs amongst the top order batsmen and in the first five over for teams batting first is of great importance. Being in a position where the lower order batsmen are scoring the bulk of the runs is losing strategy.

4.6 Conclusion

This study utilised data analysed with a univariate logistic regression in order to determine if the following variables: runs scored in the first five overs, runs scored in the last seven overs, fours scored, sixes scored, and runs scored between the different batting groups correlated with batting success in a match. The results of the study indicate that certain variables do correlate with batting success in a match. These variables are further investigated to determine the influence on batting strategy. Scoring runs in the first five overs of a match is significant form both teams batting first and second. The amount of sixes hit during a match is significant for teams batting first and the amount of fours scored significant for teams

batting second. In the following section the real world applications of the current study is explored.

4.7 Practical application

The findings suggest that batting aggressively and scoring runs quickly in the first five overs is essential to winning a match. Interestingly scoring runs amongst the top order batsmen for teams batting first is of no significance. The amount of runs scored between the top order batsmen when batting second is an important aspect to winning the match. This further emphasises the need for a good aggressive runs scoring start for teams batting second. For teams batting first scoring runs quickly at the end of the innings is a winning strategy. This is further evidence suggesting that when batting first the team should score as fast as possible without regard for being dismissed. When batting second the team should score quickly to catch up to the required scoring rate and play safely from there.

4.8 References

Bhattacharjee, D. & Lemmer, H. H. (2016). Quantifying the pressure on the teams batting or bowling in the second innings of limited overs cricket matches. *Internation Journal of Sports Science and Coaching,* 11(5), 683-692.

Cannonier, C., Panda, B. & Sarangi, S. (2015). 20-Over Vesus 50-Over Cricket: Is There A Difference? *Journal of Sports Economics*, 16(7), 76-783.

Davis, J., Perera, H. & Swatz, T. B. (2015). A Simulator for TWENTY20 Cricket. *Australian and New Zealand Journal of Statistics*, 57(1), 55-71.

Douglas, J.M. & Tam, N. (2010). Analysis of team performances at the ICC World Twenty20 Cup 2009. International Journal of Performance Analysis in Sport, 10, pp. 47-53

Hughes, M.D. & Bartlett, R.M. (2010). The use of performance indicators in performance analysis. Journal of Sport Sciences, 20(10), pp. 739-754

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, pp. 350-359

Lemmer, H. H. (2011). The single match approach to strike rate adjustments in batting performance measures in cricket. *Journal of Sport Science and Medicine,* 10, 630-634.

Munir, F., Hasan, K., Ahmed, S. & Quraish, S. (2015). Predicting a T20 cricket match result while the match is in progress. A dissertation presented for the degree of Bachelor in Computer Science.

Norman, J. M. & Clarke, S. R. (2010). Optimal batting orders in cricket. *The Journal of the Operational Research Society,* 61, 980-986.

SAS Institute Inc (2013). SAS/STAT 13.1 User's Guide, SAS Institute Inc, Cary, NC.

Shah, S., Hazarika, P. J., Hazarika, J. (2017). A studyon Performace of Cricket Players using Factor Analysis Approach. International Journal of Advanced Research in Computer Science,8(3), pp.656-660

Chapter 5 Conclusion and recommendations

5.1 Introduction

International cricket is a complex sport rich with data for research. The increasing demand for statistical analysis in cricket makes this study worthwhile. Coaches and coaching staff are seeking any advantage that provides the team with a winning edge over the opponent. Cannonier, Panda, and Sarangi (2015) talk about how cricket has transformed over the years with the inclusion of limited overs cricket, especially T20I cricket. Cricket has become a faster paced and dynamic sport. As most research focuses on quantifying batting performance by means of measuring individual performance, this study focuses on quantifying batting performance by means of team performance. Certain batting variables were chosen and analysed in an attempt to investigate if the variables can be used as success predictors in ODI and T20I cricket.

In this chapter the conclusions of the dissertation are presented. First the conclusion from article 1, whether certain batting variables could predict success in ODI cricket is presented. This is followed by conclusions of article 2, whether certain batting variables could predict success in T20I cricket. Finally, comparisons are made between the batting variable predictors of success in ODI and T20I cricket.

5.2 Conclusion article 1: batting variables which predict success in ODI cricket

In this study data was analysed using a univariate logistic regression in order to determine whether or not scoring runs in the first 20 and last 12 overs, the number of fours and sixes scored, and the number of runs scored between the different batsmen are significantly related to winning an ODI cricket match.

The results of this study show that there are variables in cricket that relate positively with success in ODI cricket and can be used as success predictors. These success predictors differ between teams batting first and second.

The significance of scoring runs and more specifically scoring sixes for the team batting first suggests that being more aggressive in stroke play and general batting

approach when batting first puts the bowlers under pressure. This in turn puts the bowling team's batsmen under pressure even before they have their chance to chase the runs total. This study highlights the importance of the batting order in ODI cricket. From this study it can be concluded that the top four batsmen are in most cases more important to winning the cricket match than others. It could be argued that ODI cricket teams may only need to employ the services of four specialist batsmen and thereafter complete the team with all-rounders and bowlers. An argument can also be made that if an ODI cricket team's wicket-keeper does not bat in the top four, that batsmen is surplus to demand in the cricket team.

The information derived from this study implies that a strategy of batting aggressively when batting first is advantageous. For teams batting second, the strategy seems to be a consistent well-structured run chase where the batsmen are conservative in stroke selection.

5.3 Conclusion article 2: batting variables which predict success in T20I cricket

This study utilised data analysed with a univariate logistic regression in order to determine if the following variables: runs scored in the first five overs, runs scored in the last seven overs, fours scored, sixes scored, and runs scored between the different batting groups correlated with batting success in a T20I cricket match.

The results of the study indicate that certain variables do correlate with batting success in a T20I cricket match. These variables are further investigated to determine the influence on batting strategy.

Scoring runs in the first five overs of a T20I cricket match is significant for both teams batting first and second. This finding suggests that batting aggressively and scoring runs quickly in the first five overs is essential to winning a T20I cricket match. Interestingly scoring runs amongst the top four batsmen for teams batting first is of no significance. Scoring runs amongst the top four batsmen batting second is an important aspect to winning. This further emphasises the need for a good aggressive

runs scoring start for teams batting second. For teams batting first scoring runs quickly at the end of the innings is a winning strategy. The amount of sixes hit during a T20I cricket match is significant for teams batting first and the amount of fours scored significant for teams batting second. This is further evidence suggesting that when batting first the team should score as fast as possible without regard for dismissal. When batting second the team should score quickly to catch up to the required scoring rate and play safely from there.

5.4 Comparison of ODI and T20I predictors of success

The above study provided statistical analysis on two different formats of cricket. The study investigated if certain variables could be used as success predictors in cricket. The study found that there are variables that correlate with success and may be used as success predictors. The two articles investigated ODI and T20I separately. From the two studies certain similarities and differences between the results of the study for ODI and T20I can be seen. This section presents those differences.

5.4.1 Success predictors for teams batting first: ODI and T20I

From the study we can see that for both ODI and T20I teams batting first runs scored in the initial and end of the batting innings are both significant. Although the amount of overs studied in the initial and end of a batting innings differs between ODI and T20I, certain similarities can be seen in batting strategy. Batting strategy for teams batting first in ODI and T20I cricket is suggested to be very similar.

The amount of sixes scored during a cricket match is significant for both ODI and T20I cricket teams. The amount of fours scored is no-significant for both ODI and T20I cricket teams. This once again suggests that the strategy for winning a cricket match with regards to scoring sixes is similar for both ODI and T20I cricket teams.

Interestingly ODI and T20I results seem to be very similar. One result that is different is runs scored amongst different batting groups. For ODI cricket teams, runs scored amongst the top four batsmen is significant and correlates with winning the cricket

match. Runs scored amongst the top four batsmen in a T20I cricket match is not significant and has no correlation to winning.

5.4.2 Success predictors for teams batting second: ODI and T20I

This study shows that there is no significant correlation between winning an ODI cricket match and scoring runs in the initial and end part of a batting innings. For T20I cricket the study shows that although there is no correlation between winning and runs scored in the end part of the batting innings there is a correlation between runs scored in the initial part of the batting innings and winning a T20I match. This may suggest a possible difference in batting strategy between ODI and T20I teams batting second.

The amount of fours scored by both ODI and T20I teams is significant and correlates with winning a cricket match. Sixes scored for T20I teams is not significant. Sixes scored for ODI teams is significant but negatively correlates with winning a cricket match. We can then therefore see that scoring sixes in a cricket match for both T20I and ODI teams is perhaps of less importance than scoring fours.

The number of runs scored amongst the top four batsmen is significant for both ODI and T20I cricket teams. The amount of runs scored by the lower order batsmen is also significant for both ODI and T20I. This correlation is however a negative one. These findings show similarities between the ODI and T20I success predictors and perhaps batting strategy as well.

The fact that this study found different significant success predictors for winning an ODI and T20I cricket match suggest that further research into the differences between ODI and T20I success predictors is worthwhile. This study can be used as a foundation in an effort to investigate if ODI and T20I batting strategies should be different or similar.

5.5 Limitations and recommendations for future study

Time motion and notational systems that are used in sport today are very expensive and are also limited to the data that each can record. No club or franchise can currently afford the expenses of such a system. Therefore, quality data are very difficult to find or would be very expensive. Thanks to ESPN Cricinfo the statistics needed for this study are freely available to the public. A limitation of the study is that acquiring data from more cricket matches for the study would require resources out of reach for the individual researcher.

Unfortunately in order to keep the two articles of realistic length the researcher could not add more variables or include bowling and fielding as potential success predictors of cricket. Future studies can include more data, more variables, as well as include bowling and fielding as potential success predictors of cricket. The researcher suggests a deeper look into the probability graphs. More statistical analysis and calculation could be completed to provide not only a representation of increase or decrease in probability but also at what specific runs scored in the first and last over, fours and sixes scored, and runs scored between the different batting order increments at which probability tips over 50% win or lose.

5.6 References

Cannonier, C., Panda, B. & Sarangi, S. (2015). 20-Over Vesus 50-Over Cricket: Is There A Difference?. *Journal of Sports Economics,* 16(7), 76-783.

Appendix A: Author guidelines for *South African Journal for Research in Sport, Physical Education and Recreation*

INFORMATION FOR AUTHORS

The *South African Journal for Research in Sport, Physical Education and Recreation* is published by Stellenbosch University. Contributions from the fields of Sport Science, Physical Education, Recreation/Leisure Studies, Exercise Science and Dance Studies will be considered for publication. The articles submitted will be administered by the appropriate Subject Review Editor and evaluated by two or more referees. The decision as to whether a particular article is to be published or not, rests with the Editorial Board.

SUBMISSION

Manuscripts that do not comply with the following requirements regarding process, style and format will not be handled. Manuscripts should be typed with **one and a half spacing** in 12-point Times New Roman letter size. The original manuscript can be submitted by e-mail. The length may not exceed 20 pages (tables, figures, references, etc. included). The **page setup** (cm) must be in the following format:

MARGINS PAPER SIZE

Top: 3.56 cm *Width*: 17.5 cm
Bottom: 1.78 cm *Height*: 24.5 cm
Left: 2.11 cm
Right: 2.11 cm
Gutter: 0.00 cm
Header: 2.03 cm
Footer: 0.89 cm

Original manuscripts may be submitted in English or Afrikaans and should be sent to:

The Editor **Editorial Office**
South African Journal for Research in Sport, Tel.: +27 (0)21-808 4915 / 4715
Physical Education and Recreation Fax.: +27 (0)21-808 4817
Department of Sport Science E-mail: sajrsper@sun.ac.za
Private Bag X1
7602 Matieland, STELLENBOSCH
Republic of South Africa

CONDITIONS

Each manuscript must be accompanied by a cover letter in which the following is declared:

(1) that the manuscript contains original research; (2) that the manuscript or parts of the manuscript has not been published elsewhere previously; (3) that the manuscript is not currently being presented elsewhere for publication; and (4) that all the authors have read and approved the manuscript. This signed declaration regarding the originality must accompany each manuscript.

Authors are also requested to name three (3) potential referees, of which one (1) **must** be an international referee (the journal is not bound to use these referees). Complete information regarding the referees (name, surname, e-mail address and telephone numbers) must be *South African Journal for Research in Sport, Physical Education and Recreation* provided in the cover letter.

We discourage the practice of parts of one study in different journals. Authors who submit a manuscript from a study of which some data has been or will be published elsewhere, must provide a strong justification in an accompanying letter to the Editor. The justification for not publishing all the data together in one paper must also be motivated in the covering letter. Should the article be taken from a master's

dissertation or doctoral dissertation, academic ethic requires that the student will be the first author.

The author should also ensure that the ***LANGUAGE*** of the manuscript has been thoroughly edited at the time of submission (English [UK]). The name, address and telephone number of the person who has done the language editing must be provided. Any expenses incurred by the journal dealing with language editing will be added to the author's page fees.

The manuscript must have an ***ETHICAL CLEARANCE NUMBER*** that was supplied by authentic ethical committee of a specific institution. The process that was followed to obtain ethical clearance must be described in the manuscript under the heading **Ethical clearance**.
No manuscript can be published without this declaration.

Any uncertainty regarding the ***STATISTICAL PROCEDURES*** that arise during the assessment of the manuscript will be referred to a local statistician. Any expenses incurred by the journal dealing with statistical procedures will be added to the author's page fees.

PREPARATION OF MANUSCRIPT

Manuscripts must be presented in a format that is compatible with *Microsoft Word for Windows* (PC). Tables, all figures (illustrations, diagrams, etc.) and graphs are also regarded as text and must be presented in a format that is compatible with *Word*. Photographs must be presented in *jpg* format.

Original manuscripts must contain the following sections in the following sequence: Title page, Abstract, Introduction, Methodology, Results, Discussion, Practical application,
Conclusions, Acknowledgements and References.

Title page

The first page of each manuscript should indicate the *title* in English and Afrikaans (will be translated for foreign authors), the *names* (**title, first name in full and other initials, surname**) of the author(s), the *telephone* numbers (work & home [Mobile – for local authors]), *facsimile* number, *e-mail* address (if available) and the *field of study*. The **complete mailing address** of the first named author and the institution where the work was conducted should be provided in full. When more than one author and/or authors from various departments are involved the **1author(s)** must be numbered according to their **1department(s)**. If any of the above-mentioned information should change during the review process, please inform the subject editor. A **short title,** of not more than **45 characters** including the spaces, should be provided for use as a running head.

Abstract

Each manuscript must be accompanied by an abstract of approximately 150-200 words in

English and should be set on a *separate page* as a SINGLE paragraph (one and a half

spacing). A list of three to seven **key words** in *English* is required for indexing purposes and should be typed below the abstract. Articles in Afrikaans must include an ***additional*** extended summary (500-1000 words) in English. This summary must start on a new page (just before the reference list) and also provide the English title of the article at the beginning.

Text

Start the text on a new page with the title of the article (centred and *without* the names of the authors). Follow the style of the most recent issue of the journal regarding the use of headings and subheadings. Use only **one space** after a paragraph. Only make use of ***section breaks*** and not ***page breaks***. The text, as well as the tables and figures, may under **no circumstances** be in any other format than **normal**. Thus, no ***style sheets*** may be used.

Tables and figures

Tables and figures should be numbered in *Arabic* numerals (1, 2, etc.). Tables require the heading at the *top* while figures a legend *below* and both separate from the table/figure. **Note:**
Use the decimal POINT (**not** the decimal comma). The site where the table or figure must be placed in the text must be indicated clearly in the manuscript. ***Tables and figures are placed after the reference list with each on a separate page***.

References

In the *text* the Harvard method must be adopted by providing the author's surname and the date placed in parentheses. *For example:* Daly (1970); King and Loathes (1985); (Botha & Sonn, 2002); McGuines *et al.* (1986) or (Daly, 1970:80) when Daly is not part of the sentence. More than one reference must be arranged **chronologically** (Daly, 1970; King & Loathes, 1985). Note that *et al.* (italics) is used in the body of the text when there are **more than two authors**, but never in the list of references.

List of references

Only the references cited in the text should be listed alphabetically according to surname (last name) of authors (capitals) after the body of text under the heading, **References** (capitals) starting on a new page. In the case where the TITLE of an article, book, etc., is in any other language than English, the author must also provide an English translation of the title in parentheses.

In the case of articles published in ***JOURNALS,*** references listed should include the surnames and initials (capitals) of all authors, the date of the publication in parentheses, the full title of the article, the full title of the journal (italics), the volume number, the serial number in parentheses (omitted **only** if the said journal does not use issue numbers), followed by a colon and a space with the first and last page numbers separated by a hyphen.

Example:

VAN WYK, G.J. & AMOORE, J.N. (1995). A practical solution for calculating instantaneous values of tension in the extensor muscles of the knee joint during

extension and flexion. *South African Journal for Research in Sport, Physical Education and Recreation,* 18(1): 77-97.

If the reference is a ***BOOK****,* the surname (last name) and initials of the author or editor (Ed.) must be given, followed by the date of publication in parentheses, the title of the book (italics) as given on the title page, the number of the edition (ed.) in parentheses, the city (and abbreviation for the state in the case of the USA OR the country [in curved parentheses]) where published, followed by a colon, a space and the name of the publisher.

South African Journal for Research in Sport, Physical Education and Recreation

Example:

JEWETT, A.E.; BAIN, L.L. & ENNIS, C.E. (1995). *The curriculum process in Physical Education* (2nded.). Madison, WI: WCB Brown & Benchmark.

For a ***CHAPTER*** from a book, the page numbers of the chapter cited must be provided in parentheses (not italics) after the title of the book. For further details, authors should consult the most recent publication of this Journal for other examples.

Example:

DE RIDDER, J.H. (1999). Kinanthropometry in exercise and sport. In L.O. Amusa, A.L. Toriola, I.U. Onyewadume (Eds.), *Physical Education and sport in Africa* (pp.235-263). Ibadan (Nigeria): LAP Publications.

If the reference is a DISSERTATION (master's level) or DISSERTATION (doctoral level), italics is **not** used in the title as it is an unpublished work.

Example:

CRAVEN, D.H. (1978). The evolution of major games. Unpublished PhD-dissertation. Stellenbosch: Stellenbosch
University.

For ***ELECTRONIC SOURCES*****,** all references start with the same information that would be provided for a printed source (if available). The web page information follows the reference.

It will usually contain the name of the author(s) (if known), year of publication or last revision, title of complete work in inverted commas, title of web page in italics, Uniform Resource Locater (URL) or access path in text brackets (do not end the path statement with a full stop) and date of access. See *"How to cite information from the Internet and the Worldwide Web"* at http://www.apa.org/journals/webref.htmlfor specific examples. When citing a web site in the text, merely give the author and date (in this case: Ackermann, 1996).

Example of Web Page:

ACKERMANN, E. (1996). "Writing your own Web Pages."*Creating Web Pages*. Hyperlink: [http://www.mwc. edu/ernie/writeweb/writeweb.html]. Retrieved on 22 October 1999.

When referencing an article in a ***NEWSPAPER*** the key word of the newspaper is typed in capitals, as this is how **it will appear in the alphabetical references**, namely *The CAPE*

ARGUS will appear under "C" or *Die BURGER* will appear under "B".

Example*:*

CAPE ARGUS (The) (1980). 10 January, p.4.

INTERVIEWS

Example:

POTGIETER, J.R. (2003). Personal interview with the Chairperson of the Department of Sport Science. 31 January.

Stellenbosch: Stellenbosch University.

CORRESPONDENCE

Example:

POTGIETER, J.J.J. (2003). Personal correspondence of the Director of the Sport Bureau, 5 February. Stellenbosch:

Stellenbosch University.

CONGRESS PROCEEDINGS

Example:

RENSON, R. (Ed.) (1976). *The history, the evolution and diffusion of sport and games in different cultures*.
Proceedings of the 4th International HISPA Congress, Leuven, Belgium, 1-5 April 1975. Brussels (Belgium):
B.L.O.S.O.
South African Journal for Research in Sport, Physical Education and Recreation
When referring to a paper presented in the above-mentioned proceedings, it must be presented as follows:
REES, R. (1976). Organisation of sport in nineteenth century Liverpool. In R. Renson (Ed.), *The history, the evolution and diffusion of sport and games in different cultures* (pp.237-247). Proceedings of the 4th
International HISPA Congress, Leuven, Belgium, 1-5 April 1975. Brussels (Belgium): B.L.O.S.O.

ADMINISTRATION

If authors honour the rules and specifications for the submission of manuscripts, unnecessary delays would be avoided. A manuscript that does not meet the requirements, as set out above, will be returned to the author without being evaluated. Requesting copying right concerning figures or photographs is the responsibility of the authors and should be indicated.
The corresponding author will receive a complimentary copy of the journal and five reprints of the article. The original manuscripts and illustrations will be discarded one month after publication unless a request is received to return the original to the first-named author. A page fee of **R300** per page is payable on receipt of an account issued by the Editor.

Appendix B: Author Guidelines for *The African Journal for Physical Activity and Health Sciences (AJPHES)*

AIM

The African Journal for Physical Activity and Health Sciences (AJPHES) is a peer-reviewed journal established to:

i) Provide a forum for health specialists, researchers in physical activity, professionals in human movement studies as well as other sport-related professionals in Africa, the opportunity to report their research findings based on African settings and experiences, and also to exchange ideas among themselves. Research-related contributions by specialists in physical activity and health sciences from other continents are also welcome.

ii) Afford the professionals and other interested individuals in these disciplines the opportunity to learn more about the practice of the disciplines in different parts of the continent.

iii) Create an awareness in the rest of the world about the professional practice in the disciplines in Africa.

GENERAL POLICY

AJPHES publishes research papers that contribute to knowledge and practice, and also develops theory either as new information, reviews, confirmation of previous findings, application of new teaching/coaching techniques and research notes. Letters to the editor relating to the materials previously published in AJPHES could be submitted within 3 months after publication of the article in question. Such letter will be referred to the corresponding author and both the letter and response will be published concurrently in a subsequent issue of the journal. Manuscripts are considered for publication in AJPHES based on the understanding that they have not been published or submitted for publication in any other journal. In submitting papers for publication, corresponding authors should make such declarations. Where part of a paper has been published or presented at congresses, seminars or symposia, reference to that publication should be made in the acknowledgement section of the manuscript. AJPHES is published quarterly, i.e. in March, June, September and December.

Supplements/Special editions are also published periodically.

SUBMISSION OF MANUSCRIPT

Original manuscript and all correspondence should be addressed to the Editor-In-Chief:

Professor L. O. Amusa Tel: +27 15 9628076/+27729883817
Centre for Biokinetics, Recreation Fax: +27 15 9628861 and
Sport Science, University of Venda,
E-mail: amusalbw@yahoo.com
P. Bag X5050,
Thohoyandou 0950
Republic of South Africa

Articles should be submitted electronically, i.e. via e-mail attachment. However, the corresponding author should ensure that such articles are virus free. AJPHES reviewing process normally takes 4-6 weeks and authors will be advised about the decision on submitted manuscripts within 60 days. In order to ensure anonymity during the reviewing process authors are requested to avoid self-referencing or keep it to the barest minimum.

PREPARATION OF MANUSCRIPT

Manuscripts should be type written in fluent English (using 12-point Times New Roman font and 1½ line-spacing) on one side of whiteA4-sized paper justified fully with 3cm margin on all sides. In preparing manuscripts, MS-Word, Office 2007 for Windows should be used. Length of manuscripts should not normally exceed 12 printed pages (including tables, figures, references, etc.). For articles exceeding 12 typed pages US$ 10.0 is charged per every extra page. Authors will be requested to pay a publication fee to defray the very high cost of publication. The pages of manuscripts must be numbered sequentially beginning with the title page. The presentation format should be consistent with the guidelines in the publication format of the American Psychological Association (APA) (6th edition).

Title page

The title page of the manuscript should contain the following information:

Concise and informative title.

Author(s') name(s) with first and middle initials. Authors' highest qualifications and main area of research specialisation should be provided.

Author(s') institutional addresses, including telephone and fax numbers.

Corresponding author's contact details, including e-mail address.

A short running title of not more than 6 words.

Abstract

An abstract of 200-250 words is required with up to a maximum of 5 keywords provided below the abstract. Abstract must be typed on a separate page using single line spacing, with the purpose of the study, methods, major results and conclusions concisely presented. Abbreviations should either be defined or excluded.

Text

Text should carry the following designated headings also using single line spacing: Introduction, materials and methods, results, discussion, acknowledgement, references and appendices (if appropriate).

Introduction

The introduction should start on a new page and in addition to comprehensively giving the background of the study it should clearly state the problem and purpose of the study. Authors should cite relevant references to support the basis of the study. A concise but informative and critical literature review is required.

Methodology

This section should provide sufficient and relevant information regarding study participants, ethics/informed consent, instrumentation, research design, validity and reliability estimates, data collection procedures, statistical methods and data analysis techniques used. Qualitative research techniques are also acceptable.

Results

Findings should be presented precisely and clearly. Tables and figures must be presented separately or at the end of the manuscript and their appropriate locations in the text indicated. The results section should not contain materials that are appropriate for presentation under the discussion section. Formulas, units and quantities should be expressed in the *systeme internationale (SI)* units. Colour printing of figures and tables is expensive and could be done upon request at authors' expense.

Discussion

The discussion section should reflect only important aspects of the study and its major conclusions. Information presented in the results section should not be repeated under the discussion. Relevant references should be cited in order to justify the findings of the study. Overall, the discussion should be critical and tactfully written.

References

The American Psychological Association (APA) format should be used for referencing. Only references cited in the text should be alphabetically listed in the reference section at the end of the article. References should not be numbered either in the text or in the reference list. Authors are advised to consider the following examples in referencing:

Examples of citations in body of the text:-

For one or two authors; Kruger (2003) and Travill and Lloyd (1998). These references should be cited as follows when indicated at the end of a statement: (Kruger, 2003); (Travill & Lloyd, 1998).

For three or more authors cited for the first time in the text; Monyeki, Brits, Mantsena and Toriola (2002) or when cited at the end of a statement as in the preceding example; (Monyeki, Brits, Mantsena & Toriola, 2002).

For subsequent citations of the same reference it suffices to cite this particular reference as: Monyeki et al. (2002).

Multiple references when cited in the body of the text should be listed chronologically in ascending order, i.e. starting with the oldest reference. These should be separated with semi colons. For example, (Tom, 1982; McDaniels & Jooste, 1990; van Heerden, 2001; de Ridder at al., 2003).

References

In compiling the reference list at the end of the text the following examples for journal references, chapter from a book, book publication and electronic citations should be considered:

Examples of journal references:

Journal references should include the surname and initials of the author(s), year of publication, title of paper, name of the journal in which the paper has been published, volume and number of journal issue and page numbers.

For one author: McDonald, A.K. (1999). Youth sports in Africa: A review of programmes in selected countries. *International Journal of Youth Sports*, 1(4), 102-117. For two authors: Johnson, A.G. & O'Kefee, L.M. (2003). Analysis of performance factors in provincial table tennis players. *Journal of Sport Performance*, 2(3), 12-31.

For multiple authors: Kemper, G.A., McPherson, A.B., Toledo, I. & Abdullah, I.I. (1996). Kinematic analysis of forehand smash in badminton. *Science of Racket Sports*, 24(2), 99-112.

Examples of book references:

Book references should specify the surname and initials of the author(s), year of publication of the book, title, edition, page numbers written in brackets, city where book was published and name of publishers. Chapter references should include the name(s) of the editor(s) and other specific information provided in the third example

below: For authored references: Amusa, L.O. & Toriola, A.L. (2003). *Foundations of Sport Science* (2nd ed.) (pp. 39-45). Makhado, South Africa: Leach Printers.

For edited references: Amusa, L.O. & Toriola, A.L. (Eds.) (2003). *Contemporary Issues in Physical Education and Sports* (2nd ed.) (pp. 20-24). Makhado, South Africa: Leach Printers.

For chapter references in a book: Adams, L.L. & Neveling, I.A. (2004). Body fat characteristics of sumo wrestlers. In J.K. Manny & F.O. Boyd (Eds.), *Advances in Kinanthropometry* (pp. 21-29). Johannesburg,
South Africa: The Publishers Company Ltd.

Example of electronic references:
Electronic sources should be easily accessible. Details of Internet website links should also be provided fully. Consider the following example:
Wilson, G.A. (1997). Does sport sponsorship have a direct effect on product sales? *The Cyber- Journal of Sport Marketing (online*), October, 1(4), at http://www.cad.gu.au/cjsm/wilson.html. February 1997.

PROOFREADING
Manuscript accepted for publication may be returned to the author(s) for final correction and proofreading. Corrected proofs should be returned to the Editor-In-Chief electronically within one week of receipt. It will be assumed that the publication should go ahead if no response is received from the corresponding author within one week. Minor editorial corrections are handled by AJPHES.

COPYRIGHT AGREEMENT
LAM Publications Ltd. holds the copyright for AJPHES. In keeping with copyright laws, authors will be required to assign copyright of accepted manuscripts to LAM Publications Ltd. This ensures that both the publishers and the authors are protected from misuse of copyright information. Requests for permission to use copyright materials should be addressed to the Editor-in-Chief.

COMPLIMENTARY COPY OF AJPHES AND ORDER INFORMATION

Principal authors will receive three (3) complimentary copies of the journal. In case of two or more joint authors the principal author distributes the copies to the co-authors. Electronic copies can be ordered from SABINET at www.sabinet.co.za or AJOL at www.ajol.info or EBSCO at www.ebscohost.com

Appendix C: Correspondence from Ethics Committee

School for Allied Health Professions (SAHP)/ Skool vir Aanvullende Gesondheidsberoepe (SAGB), UV/UFS

University of the Free State
205 Nelson Mandela Avenue
Exercise and Sport science center
10 July 2017

The Chair: Ethics committee (Dr SM le Grange)
Faculty of Health Sciences
University of the Free State

Dear Dr le Grange

Project Title: HSREC 73/2017 (UFS-HSD2017/0677) MATCH STATISTICS THAT DISCRIMINATE BETWEEN WINNING AND LOSING TEAMS IN ODI AND T20I CRICKET

Please find enclosed the modifications required for approval of the above research protocol.

1.1. Change data collection and ethical approval to July 2017 instead of June 2017: This was changed accordingly in the protocol. – *Page 15 of proposal*

1.2. Ethics clearance form: This project does not involve human subjects, please update the application form accordingly. This was changed accordingly in the application form.

1.3. Pilot study- Will you use 2 of the teams included in the study or other teams? Will you be able to include the data in the results or not? - Changed in the Protocol: Teams that will form part of the original study will be used in the pilot study, although none of the data and findings from the pilot study will be used in the main study. – *Page 14 of Proposal*

All modifications are highlighted in yellow in the Proposal.

Yours faithfully

……

M.C. Schaefer

mcschaefer26@gmail.com

0735069384

Appendix D: Ethics Approval

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA

UFS·UV
HEALTH SCIENCES
GESONDHEIDSWETENSKAPPE

IRB nr 00006240
REC Reference nr 230408-011
IORG0005187
FWA00012784

26 July 2017

MR MARK C SCHAEFER
DEPT OF EXERCISE AND SPORT SCIENCES
FACULTY OF HEALTH SCIENCES
UFS

Dear Mr Mark C Schaefer

HSREC 73/2017 (UFS-HSD2017/0677)
PRINCIPAL INVESTIGATOR: MR MARK C SCHAEFER
SUPERVISOR: R SCHOEMAN
PROJECT TITLE: MATCH STATISTICS THAT DISCRIMINATE BETWEEN WINNING AND LOSING TEAMS IN ODI AND T20I CRICKET

APPROVED

1. You are hereby kindly informed that the Health Sciences Research Ethics Committee (HSREC) approved this protocol after all conditions were met at the meeting held on 25 July 2017.

2. The Committee must be informed of any serious adverse event and/or termination of the study.

3. Any amendment, extension or other modifications to the protocol must be submitted to the HSREC for approval.

4. A progress report should be submitted within one year of approval and annually for long term studies.

5. A final report should be submitted at the completion of the study.

6. Kindly use the **HSREC NR** as reference in correspondence to the HSREC Secretariat.

7. The HSREC functions in compliance with, but not limited to, the following documents and guidelines: The SA National Health Act. No. 61 of 2003; Ethics in Health Research: Principles, Structures and Processes (2015); SA GCP(2006); Declaration of Helsinki; The Belmont Report; The US Office of Human Research Protections 45 CFR 461 (for non-exempt research with human participants conducted or supported by the US Department of Health and Human Services- (HHS), 21 CFR 50, 21 CFR 56; CIOMS; ICH-GCP-E6 Sections 1-4; The International Conference on Harmonization and Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH Tripartite), Guidelines of the SA Medicines Control Council as well as Laws and Regulations with regard to the Control of Medicines, Constitution of the HSREC of the Faculty of Health Sciences.

Yours faithfully

.....
DR SM LE GRANGE
CHAIR: HEALTH SCIENCES RESEARCH ETHICS COMMITTEE

Health Sciences Research Ethics Committee
Office of the Dean: Health Sciences
T: +27 (0)51 401 7795/7794 | E: ethicsfhs@ufs.ac.za
Block D, Dean's Division, Room D104 | P.O. Box/Posbus 339 (Internal Post Box G40) | Bloemfontein 9300 | South Africa
www.ufs.ac.za

Appendix E: Turn it in Digital Receipt

turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Mark Schaef er
Assignment title:	Turnitin Submission
Submission title:	Masters
File name:	Masters_Mark_Schaef er_2_1.pdf
File size:	1.93M
Page count:	83
Word count:	18,143
Character count:	100,374
Submission date:	29-Jan-2018 09:07AM (UTC+0200)
Submission ID:	908109695

Match statistics that discriminate between winning and losing teams in ODI and T20I cricket

MARK CHRISTOPHER SCHAEFER

In fulfillment of the degree

MAGISTER ARTIUM
(SPORT SCIENCE)

In the

Faculty of Humanities

(Department of Exercise and Sport Sciences)

At the

University of the Free State

Study Leader:

Dr R. Schoeman

Masters

ORIGINALITY REPORT

9%	5%	6%	3%
SIMILARIT Y INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

G Guthrie. "Cricket interruptus: fairness and incentive in limited overs cricket matches", Journal of the Operational Research Society,

1%

08/2004

Publication

Submitted to University of Sheffield

1%

Student Paper

Submitted to University of Bedfordshire

1%

Student Paper

www.jssm.org

1%

Internet Source

Tim B. Swartz, Paramjit S. Gill, David Beaudoin, Basil M. deSilva. "Optimal batting

orders in one-day cricket", Computers & Operations Research, 2006

1%

Publication

eprints.utm.my

1%

Internet Source

etd.uovs.ac.za