



SPORTS PERFORMANCE
RESEARCH INSTITUTE, NEW ZEALAND
AN INSTITUTE OF AUT UNIVERSITY



Modelling Progression of Competitive Sport Performance

Rita Maria Malcata

A thesis submitted to Auckland University of Technology in fulfilment of the
requirements for the degree of Doctor of Philosophy (PhD)

2014

"All the mathematical sciences are founded on relations between physical laws and laws of numbers, so that the aim of exact science is to reduce the problems of nature to the determination of quantities by operations with numbers."

James Clerk Maxwell, On Faraday's Lines of Force, 1856

TABLE OF CONTENTS

LIST OF FIGURES	6
LIST OF TABLES	8
ATTESTATION OF AUTHORSHIP	10
CO-AUTHORED PAPERS	11
ACKNOWLEDGMENT	12
CONFIDENTIAL MATERIAL.....	13
ABSTRACT.....	14
CHAPTER 1	15
INTRODUCTION.....	15
1.1. Rationale	15
1.2. Objectives.....	16
1.3. Literature Review	16
1.4. Methodological Approach.....	18
1.5. Thesis Structure	19
1.6. Research Publications and Conference Presentations	22
CHAPTER 2	23
VARIABILITY OF COMPETITIVE ELITE PERFORMANCE	23
Overview	23
2.1. Introduction.....	24
2.2. Methods.....	25
<i>Data Search and Study Selection.....</i>	<i>25</i>
<i>Data Extraction</i>	<i>26</i>
2.3. Results	28
2.4. Discussion.....	34
2.5. Conclusion.....	38
CHAPTER 3	39
TRACKING CAREER TRAJECTORIES OF SUCCESSFUL TRIATHLETES	39
Overview	39
3.1. Introduction.....	39
3.2. Methods.....	41

<i>Data</i>	41
<i>Career trajectories</i>	41
<i>Conversion of Sprint to Olympic times</i>	42
<i>Age of peak performance and age-related performance change</i>	42
<i>Performance benchmark</i>	42
<i>Prediction of race outcomes</i>	43
3.3. Results	44
3.4. Discussion	49
3.5. Conclusion.....	52
CHAPTER 4	54
PERFORMANCE PROGRESSION IN ELITE OLYMPIC-DISTANCE TRIATHLON	54
Overview	54
4.1. Introduction.....	55
4.2. Methods.....	55
<i>Data</i>	55
<i>Year Trends</i>	56
<i>Gender Differences</i>	57
<i>Contribution of each race stage</i>	57
4.3. Results	57
4.4. Discussion	62
4.5. Conclusions.....	64
CHAPTER 5	65
MODELLING PROGRESSION OF COMPETITIVE PERFORMANCE OF AN ACADEMY'S SOCCER TEAMS	65
Overview	65
5.1. Introduction.....	66
5.2. Methods.....	66
<i>Data</i>	66
<i>Year and season trends</i>	67
5.3. Results	68
5.4. Discussion	69
5.5. Conclusion.....	71
CHAPTER 6	72
USING ATHLETES WORLD RANKINGS TO ASSESS PERFORMANCE OF COUNTRIES ...	72

Overview	72
6.1. Introduction.....	73
6.2. Methods.....	73
<i>Data</i>	73
<i>Weighting approaches</i>	73
<i>Validity</i>	74
<i>Reliability</i>	74
<i>Application to 2012 London Olympic Games</i>	75
6.3. Results	75
6.4. Discussion	82
6.5. Conclusion.....	83
CHAPTER 7	85
EVALUATING SPORTS PERFORMANCE OVER OLYMPIC CYCLES	85
Overview	Error! Bookmark not defined.
7.1. Introduction.....	Error! Bookmark not defined.
7.2. Methods.....	Error! Bookmark not defined.
<i>Data</i>	Error! Bookmark not defined.
<i>Performance progression</i>	Error! Bookmark not defined.
<i>Olympic effect</i>	Error! Bookmark not defined.
<i>Evaluation of sports and countries</i>	Error! Bookmark not defined.
7.3. Results	Error! Bookmark not defined.
7.4. Discussion	Error! Bookmark not defined.
7.5. Conclusion.....	Error! Bookmark not defined.
CHAPTER 8	86
DISCUSSION AND CONCLUSION.....	95
Modelling.....	96
Practical Applications.....	99
Limitations and future directions	100
Conclusion	107
REFERENCES.....	108
APPENDIX A	113
SAS LINEAR MODELS	113
Chapter 2- Estimation of variability of surf performance scores	113
Chapter 3- Mixed model for individual trajectories for running performance.....	113

Chapter 4- Mixed model for running performance trend.....	114
Chapter 5- Generalised linear mixed model for soccer scores.....	115
Chapter 6- Generalised linear model (logistic regression) for modelling proportion	117
Chapter 7- Linear regression of individual athletes placings at annual main competitions	117
APPENDIX B	118
APPENDIX C	120
APPENDIX D	122
APPENDIX E	123
TRIATHLON TRACKING TO 2016: RUN PERFORMANCE.....	123
Overview	123
Run performance progression (Individual).....	123
APPENDIX F	126
LONDON OLYMPICS 2012: NZ SWIM PERFORMANCE REPORT	126
Performance of New Zealand in Swimming.....	127
Appendix G	129
EXCEL APPLICATION TO ASSESS PROGRESSION OF PERFORMANCE OVER AN OLYMPIC CYCLE.....	129

LIST OF FIGURES

Figure 1. Thesis structure with brief description, levels of competitive performance, variables of performance and models to be used in each of the original-research studies (Chapter 3-7).	21
Figure 2. Schematic representation of study search and selection.	26
Figure 3. Observed (triangle) and corrected (circle) performances' times for an athlete, after times being adjusted to a mean race. The black curve is an athlete's trajectory, illustrating the fitting of a quadratic age trend to the corrected times.....	45
Figure 4. Performance trajectories as functions of age for top men and women for swimming, cycling, running and total performance time. Top 3 athletes' trajectories are displayed in black while Top 16 are presented in grey. (See Methods for athletes' classifications.)	47
Figure 5. Comparison of an athlete's season performance and his trajectory against the typical age-related changes in running performance (mean and 90% performance range, see methods).	48
Figure 6. Relationship between observed performances in the 2012 London Olympics and predicted chances of winning. Chances of winning were predicted using athletes' career trajectories derived including races up to the Olympics and assuming similar race conditions (environmental and race-course profile) as 2011 London test event.	48
Figure 7. Correlations (with 90% confidence limits) of observed Olympic placings with predictions using trajectories and with placings at the last race before the Olympics. The mean correlation between Olympic placings and placings in each of the 2012 World Triathlon Series races before the Olympics is also shown.	49
Figure 8. Progression of mean swimming, cycling, running and overall performance times for the top 16 male and female athletes in every key race (World Championships and Olympic Games) between 2000 and 2012. Estimated linear year trends and 90% confidence limits are presented as bold and dashed lines, respectively.....	59
Figure 9. Differences between mean times for the top-16 men and top-16 women in all races (World Cup, World Triathlon Series, World Championships and Olympic Games races; 180 men's and 178 women's races) between 2000 and 2012 for swimming, cycling, running and overall performance. Estimated linear year trends and 90% confidence limits are presented as bold and dashed lines, respectively.....	60
Figure 10. Correlations of performance times in each stage with overall placings in all races (World Cup, World Triathlon Series, World Championships and Olympic Games), in elite and in junior World Championships, for men and women between 2000 and 2012. Regression lines for each stage are shown.	61

Figure 11. Raw mean number of goals scored in each season for the three Aspire cohort and their Opposition teams. Lines show adjusted means and 90% confident limits for the Aspire and Opposition teams, for the six seasons between 04/05 though to 09/10.	69
Figure 12. Observed proportion of athletes on a given world rank who ever achieved top rank in the 22-year period. The curve derived by logistic regression is $p = \frac{3.49 \cdot \text{Rank} - 1.46}{1 + 3.49 \cdot \text{Rank} - 1.46}$ (blue line), while the line is the simpler inverse-rank function (green line).	76
Figure 13. Four different weightings used to derive country scores from athletes' ranks based on: Inverse Rank ($w=1/\text{Rank}$), Top Rank (weightings given by the above proportion p , for $\text{Rank}>1$), Top-3 Rank (based on modelled proportion of athletes on a given world rank who ever achieved any top-three rank, $w= \frac{11.39 \cdot \text{Rank} - 1.48}{1 + 11.39 \cdot \text{Rank} - 1.48}$, for $\text{Rank}>3$) and Gold Medal (based on the modelled proportion of athletes on a given world rank who ever achieved a gold medal at major competitions, $w= \frac{3.07 \cdot \text{Rank} - 1.14}{1 + 3.07 \cdot \text{Rank} - 1.14}$, for $\text{Rank}>1$).	76
Figure 14. Progression of country performance scores and rankings from 1990 to 2011 for the top-three ranked countries in 2011, USA, China and Japan, by summing the inverse of the world ranks for all the swimmers represented in the top 150 for all of the 16 female and 16 male pool swimming events.	78
Figure 15. Chapter is under embargo.	88

LIST OF TABLES

Table 1. Green and Oakley's characteristics of elite sport development systems (Green & Oakley, 2001)	15
Table 2. List of conference and publications originated by Chapter 2-6 of this PhD thesis	22
Table 3. Estimates of within-athlete variability from competition to competition within a season separately by event and sport (mean and 90% confidence interval). Information regarding number of races included in the analysis, type of competition and type of competitors is also shown. Season has different duration dependent on event and sport.	29
Table 4. Mean estimates and confidence intervals of within-athlete variability for competitions within and between seasons (expressed as coefficient of variation and 90% confidence intervals), and mean intraclass correlation coefficients (ICC) and 90% confidence intervals for competition within and between seasons for the sports identified in Table 3.	33
Table 5. Number of competitors (mean \pm SD), race times (mean \pm SD, min) and typical spread of performances in a race (mean of standard deviations on the performance times in each race) in 224 men's and 222 women's Sprint and Olympic distance triathlon races between 2000 and 2012.....	44
Table 6. Predicted age of peak performance and 5-year improvement for swim, bike, run and total performance for the "top" and all athletes (see Methods for athletes' classification). Only athletes who had already reached their predicted age of best performance were included in these estimates (30 for Top men, 30 for Top women, 152 for Other men and 79 for Other women).	46
Table 7. Mean race time (min) and typical difference (min) between athletes within a race for top 16 athletes in every World Cup, World Triathlon Series, World Championships and Olympic Games race from 2000 through 2012. In total, data consisted of 362 individual male and 342 female athletes competing in 180 and 178 races, respectively.	58
Table 8. Predicted mean top-16 times and 90% confidence limits (min) for swim, cycle, run and overall in 2016. Estimates were predicted by extrapolating year progressions towards 2016, using data from key races (World Championships and Olympic Games) between 2000 and 2012 shown in Figure 8.	58
Table 9. Performance scores, rankings and medal counts for countries ranked in the top 20 in 2011 by summing the inverse of the world ranks for all the swimmers represented in the top 150 for all of the 16 female and 16 male pool swimming events. The scores and resulting country rankings based on three other approaches for deriving weightings are also shown: the modelled proportion of all athletes on the rank who ever achieved top rank (p top 1) and similar modelled proportions for achieving top three (p top 3) and gold medal at Olympics or World Championships (p gold).	79

Table 10. Rank-based performance scores, medal count and rankings of the top-20 countries for pool-swimming events at the 2012 London Olympic Games. The rank-based score was obtained by summing the inverse of the ranks of each country's athletes in the 32 events and can be interpreted as gold-medal equivalents.	81
Table 11. Chapter is under embargo	90
Table 12. Chapter is under embargo	91
Table 13. Chapter is under embargo	92
Table 14. Chapter is under embargo	92
Table 15. Summary of the performance metrics, models, useful findings, limitations and future directions investigated, developed and discussed in this thesis.....	103

ATTESTATION OF AUTHORSHIP

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any degree or diploma of a university or other institution of higher learning.

Chapter 2 to 7 of this thesis represent six separate papers that have either been published or have been submitted to peer-reviewed journals for consideration for publication. My contribution and contribution of the various co-authors to each of these papers are outlined at the beginning of this thesis (see Co-authored papers section). All co-authors have approved the inclusion of the joint work in this doctoral thesis.

Rita Malcata		June 2014
Will G Hopkins		June 2014
Simon Pearson		June 2014
Tom J Vandenbergae		June 2014

CO-AUTHORED PAPERS

Chapter	Publication reference	Contribution
Chapter 2	Malcata, R. M., & Hopkins, W. G. (submitted). Variability of competitive elite performance. <i>Sports Medicine</i> .	RM 80% WH 20%
<i>Contribution</i>	RM- Search and review of literature, quantification of effects and writing manuscript. WH- Clarification of concepts in the literature, guidance in the statistical analysis and review of manuscript.	
Chapter 3	Malcata, R. M., Hopkins, W. G., & Pearson, S. (2014, in press). Tracking career performance of successful triathletes. <i>Medicine and Science in Sports and Exercise</i> .	RM 80% WH 15% SP 5%
<i>Contribution</i>	RM- Data collection, statistical modelling and writing manuscript. WH- Advice on statistical modelling and review of manuscript. SP- Advice for discussion and interpretation of the findings.	
Chapter 4	Malcata, R. M., Hopkins, W. G., & Pearson, S. (submitted). Performance progression in elite Olympic-distance triathlon. <i>Scandinavian Journal of Medicine and Science in Sports</i> .	RM 80% WH 10% SP 10%
<i>Contribution</i>	RM- Data collection, statistical modelling and writing manuscript. WH- Advice on statistical modelling and review of manuscript. SP- Advice for discussion and interpretation of the findings.	
Chapter 5	Malcata, R. M., Hopkins, W. G., & Richardson, S. (2012). Modelling the progression of competitive performance of an academy's soccer teams. <i>Journal of Sports Science and Medicine</i> , 11, 533-536.	RM 80% WH 17% SR 3%
<i>Contribution</i>	RM- Statistical modelling and writing manuscript. WH- Advice on statistical modelling and review of manuscript. SR- Data collection and research question.	
Chapter 6	Malcata, R. M., Vandenberghe, T. V., & Hopkins, W. G. (2014, in press). Using athletes' world rankings for assessing country performance. <i>International Journal of Sport Physiology and Performance</i> .	RM 80% TV 10% WH 10%
<i>Contribution</i>	RM- Data collection, statistical modelling and writing manuscript. TV- Research question and advice for discussion and interpretation of the findings. WH- Advice on statistical modelling and review of manuscript.	
Chapter 7	Malcata, R. M., Hopkins, W.G. & Young, R. (embargo). Evaluating sports performance over Olympic cycles.	RM 82% WH 15% RY 3%
<i>Contribution</i>	RM- Data collection, statistical modelling and writing manuscript. WH- Advice on statistical modelling and review of manuscript. RY-Research question.	

ACKNOWLEDGMENT

First I would like to thank my supervisors, Will, Simon and Tom, who have been a constant source of advice, support and inspiration, and who have been role models for my development as a sports scientist. I must state that throughout this thesis when I have used the pronoun “I”, it often means “we”. This project has been a joint work and I would not have achieved it without any of them. Will Hopkins has been a real mentor guiding my learning of statistics and analysis of sport performance, engaging me always with his new research challenges and subsequent solutions, but also stimulating knowledge and discussion in other areas of life. Thank you for all the patience over the last three years and the countless one-to-one sessions required to improve my understanding of statistical concepts and modelling features within SAS, and not forgetting the “tremendous” work required to translate the written Portuguese English to academic English. Simon Pearson, whose knowledge and experience in applied sport science helped drive this project as a whole and never forgetting the sport context. Thanks for the particular contribution on the triathlon research, and the insights that provided useful understanding of the “dynamics” within the sport. Thank you to Tom Vandenberghe whose working dedication has inspired me, and whose expertise in swimming science made a great contribution particularly to one of the studies.

Scott Richardson (Aspire, Qatar) and Richard Young (HPSNZ, New Zealand) for providing the research question for studies in Chapter 5 and 7 and, Jason Healy (old-New Zealand Academy of Sport North Island) for imagining this PhD scholar position and bringing me over to New Zealand.

I would like to acknowledge High Performance Sport New Zealand (HPSNZ) and Auckland University of Technology (AUT) for providing the funding for this thesis.

Thanks also to all the friends that I have met in the last three and half years, since my arrival in this country. You have been the support and an away family; without you I would not have enjoyed as much this incredible journey (although I probably would have finished this thesis a few months earlier).

Finally to mum, dad and Francisco, who have always been the support and safety net which enabled me to take the risk and move to the other side of the world to embrace what has been a fantastic journey.

CONFIDENTIAL MATERIAL

Chapter 7 and Appendix G of this thesis are currently under embargo. High Performance Sports New Zealand believes that the method and subsequent findings provide competitive advantage to New Zealand sport by helping planning and driving the strategic view for the 2016 Rio Olympic campaign.

ABSTRACT

Athletes, coaches, sport scientists and managers need objective assessment of changes in competitive performance to provide evidence for guiding athletes' development, for assessing programme effectiveness, and for supporting decisions regarding allocation of funds in sports campaigns. This PhD is focused on the development of analytical tools using the Statistical Analysis System (SAS) software for assessing changes in competitive performance.

The topic of variability of competitive performance is reviewed first, because estimates of variability provide thresholds of magnitude for assessing important changes in performance. Five original-research studies are then presented for assessing performance changes in five levels of performance: athlete, sport, team, country squad and all Olympic sports of a country. First, mixed linear modelling was used to develop individual career trajectories of triathletes while accounting for environmental and other external factors. This analytical tool allows evaluation and comparison of athletes against the typical performance progression of successful elite triathletes. Secondly, linear performance trends with calendar year were evaluated using a mixed modelling approach to investigate progression of mean performance times for the sport of triathlon providing coaches and support staff with the current state of the sport. Thirdly, improvement of a football team's performance was quantified using generalised mixed linear modelling to assess the effectiveness of a youth-talent development programme. The focus of the fourth investigation was the development of a country score to provide a more comprehensive measure of performance than measures based on medal counts. These scores were derived by properly combining each country's athletes' world rankings. Finally, performance progression of individual athletes and teams over an Olympic quadrennium was assessed using linear regression of athletes' placings at annual main competitions. The analysis also provided a measure to evaluate under- and over-achievement at Olympics.

In this thesis, general and generalized linear and mixed linear models proved to be appropriate for modelling changes in sport competitive performance. Further investigation is required to extend the models presented here to other sports and to explore non-linear models for analysis of competitive performance.

CHAPTER 1

INTRODUCTION

1.1. Rationale

Monitoring performance is an integral component of many domains of human endeavour. It is often associated with business or engineering software as an important tool to generate objective data to judge success or failure of investments and applications. In both cases, only when properly informed, management can adopt strategies to further improve performance. A similar scenario is present in sport.

Green and Oakley (2001) identified ten determinants of success of a “western” elite sport system (Table 1). Items 3, 8 and 9 highlight the importance of monitoring different levels of sport performance for evidence-based decisions. In common with many other national sports organizations, High Performance Sport New Zealand has identified on-going need of such evidence-based decisions, and is therefore promoting the development and implementation of monitoring system and evaluation of sport performance (HPSNZ, 2012). Establishment of athlete development pathways, identification of effects of training strategies and monitoring sports campaigns are examples of areas requiring such analysis of performance. The intended outcomes would provide a better understanding of athlete development, quantification of factors affecting sport performance and optimisation of investment for performance achievements.

Table 1. Green and Oakley’s characteristics of elite sport development systems (Green & Oakley, 2001)

-
1. Role clarity for the agencies involved and effective communication
 2. Simplicity of administration through common sporting and political boundaries
 3. Statistical identification and monitoring of the progress of talented and elite athletes
 4. Provision of sports services to create a culture of excellence in which all members of the team (athletes, coaches, managers and scientists) can interact
 5. Well-structured competitive programmes with on-going international exposure
 6. Well-developed facilities with priority for elite athletes
 7. Targeting resources on focus sports, with real chance of success at world level
 8. Comprehensive planning for each sports needs
 9. Recognition that excellence costs, funding appropriately people and infrastructure
 10. Lifestyle support during and post elite athlete’s career
-

Performance of a sport system can be evaluated in different ways, but HPSNZ believes on a “performance-based system that has accountable outcomes” (HPSNZ, 2012). Consequently, performance of athletes, coaches, managers and support staff are judged ultimately by athletes’ successes or failures at competitions.

1.2. Objectives

My aim in this PhD has been to develop analytical models for tracking changes of sport competitive performance over a period of years, building on previous research and addressing their limitations. Models for tracking changes in competitive performance will provide coaches, support staff and high performance directors with objective information to support evidence with regards of athlete's development, defining benchmarks and consequently identifying the gap between their athletes and their competitors, and identifying successful sport campaigns.

The research question of modelling changes in competitive performance has been answered by designing models for tracking changes for the different levels of sport competitive performance: athlete, sport, team, country squad and all Olympic sports of a country. I focused only on athletic competitive performance, because such performances are the ultimate measure of success.

The analytical models have been developed based on linear models, by assuming a (generalised) linear relationship of competitive performance, the dependent variable, with the time predictors, age or calendar year. Furthermore, in order to establish the link between research and applied sport, the developed models for analysing competitive performance have been integrated into practical applications and reports which have provided coaches and high performance managers with objective assessment of performance.

1.3. Literature Review

Scientific research on monitoring (or tracking) changes in competitive performance has been limited and has been focused mainly in changes in the mean performance of a group of athletes and in changes in world records over time. Tracking mean performance of a group of athletes over a number of years provides information on trends and developmental stage of sports, reflecting the impact of new training strategies, advance in sport science and impact of technology and equipment (El Helou et al., 2010; Haake, Foster, & James, 2013). While progression of world records has been investigated mainly in relation to evolution of human capacity and potential limits to human performance (Berthelot et al., 2008; Nevill & Whyte, 2005).

Limited research has been performed on assessment and quantification of individual athletes' careers. Balyi (2001) proposed a five-stage model to explain the long-term development of an athlete (LTDA) in a sport such as track-and-field athletics. The five stages included an initial development of *fundamental skills* up to age of 10 years; followed by a period of *training to train* when athletes learn the basic skills of the specific sport; after the of age ~14 years athlete developed their competitive skills, during the stage called *train to compete*; followed by the elite career period where athletes *train to win* and peak their performance at competitions. The final stage is the *retirement*. Research on tracking changes of individual athletes' performances at competitions has been developed for baseball (Albert, 2002), cross-country skiing (Alam, Carling, Chen, & Liang, 2008), skeleton (Bullock & Hopkins, 2009), swimming (Allen, Hopkins, & Vandenberg, 2012; Berthelot et al., 2012; Pike, Hopkins, & Nottle, 2010), tennis (Guillaume

et al., 2011) and track-and-field athletics . Depending on the sport the modelled measures of performance were time, distance, speed, baseball player performance score and annual percentage of tennis matches won. The different authors used several models to describe changes in performance. Alam et al. (2008), focused on stage 2 and 3 of LTDA model, used a sigmoid function of age to describe the effects of puberty in speed cross-country skiing. A double exponential function of age was used to track performance changes up to retirement for top tennis players (Guillaume, et al., 2011) and from junior ages to master ages in track-and-field athletics and swimming (Berthelot, et al., 2012). In the other studies, performance was explained by a quadratic function of age, describing progression of athletes' performance mainly in stage 3 and 4 of LTDA model. The majority of the studies, with exception of skeleton (Bullock & Hopkins, 2009) and cross-country skiing (Alam, et al., 2008), involved sport performances with relatively stable competition environments, therefore with no consideration for effects of environmental and other external factors in performance. Furthermore, in some studies repeated measurements on athlete were not taken into account, because authors either did not specify it in the model (Alam, et al., 2008) or produced an individual fit for each athlete (Berthelot, et al., 2012; Guillaume, et al., 2011). There is therefore a need to develop a model to describe the relationship of competition performance with age over an athlete's career with adjustment for environmental and other course-related factors and accounting for repeated measurement.

The relationship of performance changes with age has been investigated commonly in cross-sectional studies by tracking mean performance of top athletes in each age group, for a large range of sports (Bernard, Sultana, Lepers, Hausswirth, & Brisswalter, 2010; Bird, Balmer, Olds, & Davison, 2001; Etter et al., 2013; Guillaume, et al., 2011; Lepers & Maffiuletti, 2011; Rüst, Knechtle, & Rosemann, 2012; Stevenson, Song, & Cooper, 2012). This resulting relationship between performance and age represents the relationship for an ideal athlete, always consistently among the top performers for all ages and not the typical relationship of a top individual athlete.

Tracking mean performance of a group of athletes over the years has also been applied for evaluating calendar year trends in a sport and investigate gender differences, often with performances at *world championships* (El Helou, et al., 2010; Knechtle, Knechtle, & Lepers, 2010; Lepers, 2008; Wainer, Njue, & Palmer, 2000), *Olympic Games* (Radicchi, 2012) or the *best performances within a year* (Haake, et al., 2013). These studies were performed in a cross-sectional fashion, regardless of the identity of athletes and therefore ignoring the fact that the same athletes may be represented in different years. There has been a similar disregard of athlete identity in the investigation of progression of *world records* (Desgorces et al., 2008; Kuper & Sterken, 2003; Lippi, Banfi, Favaloro, Rittweger, & Maffulli, 2008; Nevill & Whyte, 2005; Wainer, et al., 2000). The majority of these studies has been on track-and-field athletics and swimming; not only because of the availability of the data but also because the reliability of the performance measures. Statistical models to represent the relationship between performance and time varied depending on the study. While in some studies, linear model were deemed appropriate to represent the trend of progression of performance (Tatem, Guerra, Atkinson, & Hay, 2004; Wainer, et al., 2000), for other time windows more complex models, such as exponential relationships (Haake, et al., 2013; Kuper & Sterken, 2003; Nevill & Whyte, 2005) were needed to

represent the “natural limit” of performance. Haake and colleagues (Haake, et al., 2013) provided a summary table of the several mathematical models used to describe the relationship of progression of performance and time. It is still unclear what defines the best model and there appear to be several issues in the choice of the model. First, the relationship of performance and time depends on the time window of analysis. Secondly linear models may show the global trend of progression for performance and time by providing, for example, estimates of rate of improvement (or decline). It is however unrealistic to extrapolated these linear trends to derive performances in the distant future (Reinboud, 2004). Thirdly, more complex models may produce better fitting of the data, but the resulting relationships may be difficult to derive information useful for a practical setting.

Progression of individual athletes and changes in the overall sport are important, but managers, media and public are also interested in progression of countries, particular after each Olympic Games. Research on long-term assessment of countries performance in a sport and/or in a big sport event, such as the Olympic Games, has been limited. These assessments are often based on the number of medals, and many ways have been suggested to combine the different medal colours (Blight & Rogers, 2012; Seiler, 2013; Steiler, 2010; Wood, 2012) or accounting for other socio-economic factors, such as total country population (Putt, 2013). Although important, these measures provide only an assessment of podium performances. There is a need to provide more comprehensive measures of competitive performance, rewarding also non-medal performances and improvement over an Olympic quadrennium. These assessments will provide coaches and managers with more information about a country’s talent base, enabling a more strategic view for planning future athletes’ careers.

1.4. Methodological Approach

This thesis is centred on progression or tracking changes of competitive performance. Before quantifying changes, there is a need to identify the measure of performance. I have limited my research to competitive performance (performance outcomes at competitions), because these are the performance outcomes that an athlete is ultimately judged on. Measures of performances used as dependent variables were triathlon times, soccer scores, world ranking of swimmers and athletes’ placings at annual main competition.

A measurement of any quantity is only ever an estimate of its true value, because the measurement includes an error term. A measurement of performance is determined by the true value of ability with additional contribution of external factors affecting performance and a random error term. Errors and effects of external factors in performance are inherent to any measurement and determine the magnitude of the uncertainty of estimates around the true value. It is therefore important to understand that in a sport, the measure of performance is only a surrogate of the true athlete, team or country’s ability.

For the different studies on performance progression, I have used different modalities of linear models to accommodate the different types of performance outcomes and to account for

the different sources of errors and effects of external factors on sport performance. For this purpose, I have explored the different linear models available in the Statistical Analysis System (SAS) software for modelling the progression of competitive performance in my original research articles.

Linear models are the first choice in data analysis because of the availability of well-developed analytical procedures in statistical packages. Linear models also allow a direct interpretation and quantification of the effect of a predictor on a dependent variable from the coefficients (parameters) in the equation representing the model (W.G. Hopkins, 2010). Furthermore, the additive nature of linear models allows quantification of effect of predictors on a dependent variable while adjusting for effects of other predictors (moderators and mediators) (W.G. Hopkins, 2010); for example, adjusting for the effect of altitude on performance in track and field athletics (Hollings, Hopkins, & Hume, 2012). Mixed models are a particular modality of linear models that I have used frequently in my research. In mixed models the dependent variable is explained by *fixed effects* and *random effects* (Truxillo et al., 2012). Fixed effects are predictors that affect the entire population (or subgroup of the population) in the same way, and whose levels in the model are all the possible levels in the population. Random effects are predictors whose levels are a random sample of all the levels in the population. Random effects are used only when there is a need to account for repeated measurements or clustering within each level of the random effect (e.g., repeated measurements for each athlete). The solution for the random effects represents the deviation of such clusters (e.g., performances of an individual athlete) after adjustment for all other effects in the model. These complex models account, therefore for different sources of error and effects of factors affecting sport performance within athletes (W.G. Hopkins, 2003), providing estimates of individual responses and quantification of performance change for individual athletes and individual teams. Additionally, mixed models allow for analysis of dataset with missing values, where conventional analysis of variance would fail (W.G. Hopkins, 2003).

1.5. Thesis Structure

This thesis is organized in eight chapters (Figure 1). The first and present chapter provides the rationale, objectives, methodological approach and overview of the research on tracking performance changes that positions the reader for what follows. This chapter is not intended to be a full literature review of this thesis, because each chapter contains a review of the relevant literature for the performance measure and model analysed.

The most important chapters consist of a systematic review and five original-research studies. Each chapter starts with an overview and it consists of the publishable manuscript (either in print, Chapter 3, 5 and 6; submitted for publication, Chapter 2 and 4; or ready for publication but currently under embargo, Chapter 7). All the original-research studies were non-interventional, and data for the analyses were available online, therefore no ethics approvals were required.

Each of the five original studies represents an application of mixed linear models for tracking different levels of competitive performance: athlete, sport, team, a national squad and all of a country's Olympic sports. The sports investigated in this thesis (triathlon, football, swimming and all the country's Olympic sports) were chosen mainly because support staff of such sports (in HPSNZ and ASPIRE) had identified the need for such models and believed that data currently available would provide useful information towards monitoring and planning athletic performance.

In Chapter 2, the topic of variability of competitive performance of elite athlete is reviewed. I have done a systematic review of this topic, because estimates of variability provide the thresholds for assessing magnitude of performance changes and estimation of variability is one application of mixed modelling. This literature review is intended to be a stand-alone publishable article, in a define topic related but not necessarily encompassing the entire thesis.

Chapter 3 is the first of my original-research studies on progression of competitive performance. Here, I have presented a model for deriving career trajectories of elite triathletes, by tracking triathletes' performance times while accounting for effects of environmental and other course-related factors. Using a similar modelling approach to account for effects of environmental and other course-related factors and repeated measurements for athletes, I present in Chapter 4 a model of progression of mean performance times of top triathletes over the years. I have used this model to investigate developmental phase of the sport of triathlon and to enquire about the contribution of the three race stages, *swim, bike and run*, to overall performance. The focus of Chapter 5 is evaluation of a team performance over a period of five years by modelling performance scores of a football academy team. This study was performed in collaboration with Aspire Academy for Sport Excellence (Doha, Qatar) for answering the question of success (or not) of the youth development programme over those five years. In Chapter 6, progression of performance of a country's swimming squad is investigated by combining athletes' world rankings. There was a need for more comprehensive approaches to assess country's performance than tracking the number of medals, because many countries compete internationally without winning medals regularly. Swimming was chosen for designing such approach because of the availability of world rankings for multiple strokes and distances. In Chapter 7, I have derived performance metrics to assess country's Olympic sports by investigating progression of athletes' placings at the annual main competition over an Olympic cycle. The idea for this final study was raised by the Knowledge Edge for Rio programme (HPSNZ), because up until now success was measured only by number of medals, and there was a need to identify development athletes who progressed extremely well during their Olympic campaign but did not win a medal. Finally, in Chapter 8, overall conclusions are drawn and several perspectives for future research are described.

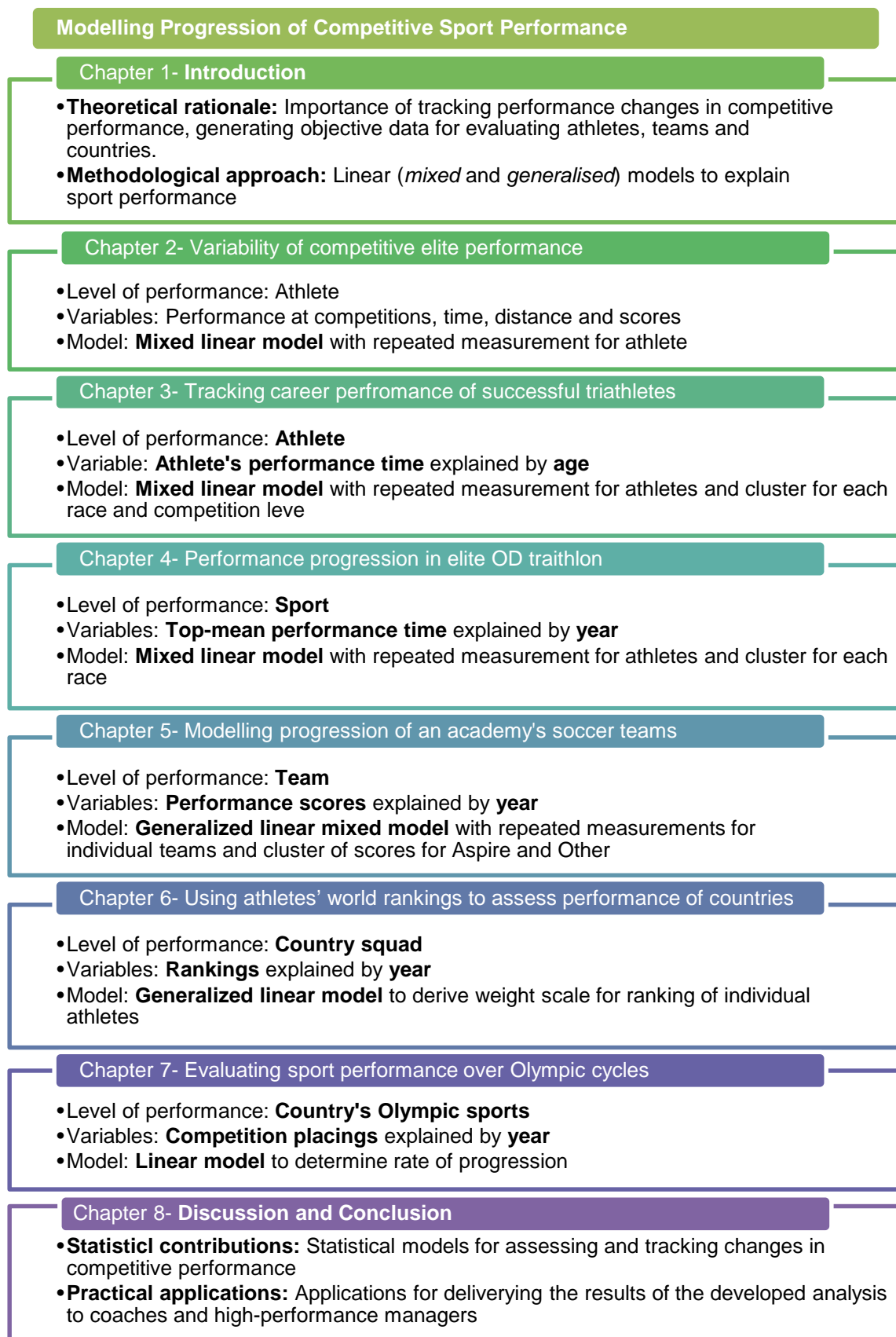


Figure 1. Thesis structure with brief description, levels of competitive performance, variables of performance and models to be used in each of the original-research studies (Chapter 3-7).

1.6. Research Publications and Conference Presentations

The research studies (Chapter 2-7) from this doctoral thesis have resulted in conferences presentations and in either publications or articles submitted for publication. Chapter 7 has not been submitted to any journal or conference because it is currently under embargo, as HPSNZ believes findings may provide competitive advantage for New Zealand sport.

Table 2. List of conference and publications originated by Chapter 2-6 of this PhD thesis

Chapter	Publication or Conference
Chapter 2	Malcata, R. M., & Hopkins, W. G. (2014, in press). Variability of competitive elite performance. <i>Sports Medicine</i> .
Chapter 3	Malcata, R. M., Hopkins, W. G., & Pearson, S. (2014). Tracking career performance of successful triathletes. <i>Medicine and Science in Sports and Exercise</i> , 46 (6), 1227-34. Malcata, R. M., Pearson, S., & Hopkins, W. G. (2013). <i>Tracking career performance of successful triathletes</i> . Presented at the meeting of the European College of Sport Science, Barcelona, Spain.
Chapter 4	Malcata, R. M., Hopkins, W. G., & Pearson, S. (submitted). Performance progression in elite Olympic-distance triathlon. <i>Scandinavian Journal of Medicine and Science in Sports</i> .
Chapter 5	Malcata, R. M., Hopkins, W. G., & Richardson, S. (2012). Modelling the progression of competitive performance of an academy's soccer teams. <i>Journal of Sports Science and Medicine</i> , 11, 533-536. Malcata, R. M., Hopkins, W. G., & Richardson, S. (2011). <i>Modelling the progression of competitive performance of an academy's soccer teams</i> . Presented at the meeting of the European College of Sport Science, Liverpool, United Kingdom.
Chapter 6	Malcata, R. M., Vandenbogaerde, T. V., & Hopkins, W. G. (2014). Using athletes' world rankings for assessing country performance. <i>International Journal of Sport Physiology and Performance</i> , 9, 133-138. Malcata, R. M., Vandenbogaerde, T. V., & Hopkins, W. G. (2012). <i>Using athletes' world rankings for assessing country performance</i> . Presented at the meeting of the European College of Sport Science, Bruges, Belgium.

CHAPTER 2

VARIABILITY OF COMPETITIVE ELITE PERFORMANCE

This chapter comprises the following paper submitted to Sports Medicine:

Malcata, R. M., & Hopkins, W. G. (submitted). Variability of competitive elite performance. *Sports Medicine*.

Overview

Thresholds for important changes in competitive performance of elite athletes are needed to assess potential performance-enhancing strategies and other factors that could affect an athlete's chances of winning. These thresholds are obtained from estimates of within-athlete variability, the random variation shown by athletes from competition to competition. **Purpose:** To systematically review estimates of within-athlete variability of competitive performance in various sports. **Methods:** We searched SPORTDiscus and Google Scholar for studies providing estimates of within-athlete variability between competitions as coefficients of variation. Estimates are reported here only for the best athletes. Some studies also combined within-athlete variability with between-athlete differences into a measure of predictability expressed as an intraclass correlation coefficient, reported here for the full field of competition. **Results:** Estimates of within-athlete variability ranged from 0.15% for speed-skating times to 39% for surfing scores. For sprint and endurance sports, such as running and cycling, estimates ranged from 0.6% to 1.4%, while for sports requiring explosive power in a single effort, such as field events and weightlifting, estimates ranged from 1.4% to 3.3%. Greater contributions of skill, chance or environment presumably accounted for relatively higher within-athlete variability in canoe slalom, mountain biking and surfing. Speed skating and skeleton had the lowest variability, because performance times are determined mainly by performance in the initial phase of the race. Predictability correlations ranged from 0.18 (surfing) to 0.93 (cross-country skiing). There was little difference in variability or predictability between men and women. **Conclusion:** The wide range in within-athlete variability of performance between sports must be due to different factors affecting performance, including physiology, skill, competition dynamics, environment, measure performance (time, distance, score), and chance.

2.1. Introduction

Athletes show natural variability in their performance from one competition to the next. Understanding and quantifying this variability is important. In an article published in 1999, the variability in an athlete's performance from competition to competition was shown to provide an estimate of the smallest important change in performance, a crucial item of information for sport scientists monitoring individual athletes or designing and analysing studies of factors affecting athletic performance (W.G. Hopkins, Hawley, & Burke, 1999). Since then, investigation of variability of competitive performance has become a focus of research interest, not only to provide estimates of the smallest important effect in various sports but also to gain understanding of the factors responsible for performance variability.

In the 1999 article, the smallest important change in performance was defined as the change in performance that results in one extra medal in every ten competitions for athletes who are already winning medals frequently. For sports in which athletes compete as individuals for the best time, distance or performance score, simulations showed this change to be 0.3 of the typical random variation that an athlete shows from competition to competition (W.G. Hopkins, et al., 1999). Thresholds for moderate, large, very large and extremely large changes in performance between competitions, representing three, five, seven and nine extra medals in every ten competitions respectively, were shown subsequently to be 0.9, 1.6, 2.5 and 4.0 of this variability (W.G. Hopkins, Marshall, Batterham, & Hanin, 2009). Estimates of the within-athlete variability between competitions are therefore crucial for identifying important performance changes in practical and research settings.

When an athlete's performance changes between competitions, part of the change is experienced by all the athletes in the competition; for example, all triathletes will have an increase in performance time if the race course is longer than usual. Such changes in performance should be excluded from the analysis to derive estimates of the within-athlete variability, because factors that affect performance of all athletes do not alter their chances of winning. In almost every published study, the appropriate analysis has been performed with mixed linear modelling, with fixed and random effects to explain performance. *Fixed effects* are used to adjust for factors affecting mean performance (such as competition identity, race distance, environmental conditions and so on); *random effects* include athlete identity to estimate differences between athletes, athlete identity interacted with year to estimate within-athlete between years variability, athlete identity interacted with run to estimate within-athletes between trails variability (for competitions in which total performance is result of the sum of the performances in the different trials, such as in skeleton (Bullock, Hopkins, Martin, & Marino, 2009)); and residual to estimate the typical variation shown by athletes from competition to competition. This typical random variation provides a measure of consistency of athletes' performances after accounting for effects of factors affecting mean performance. The typical random variation is also known as within-athlete variability and has been expressed in all studies as a coefficient of variation, a standard deviation expressed as percentage of a mean. Coefficients of variations are derived when log-transformed performances are analysed with linear models, because errors in a linear model are

derived as an additive effects which become percent effects after back-transformation. The coefficient of variation in competitive performance is equivalent to the typical percent error for a performance test in a reliability study (W.G. Hopkins, 2000).

The within-athlete variability can be combined with differences between athletes (spread of performances between athletes) to derive a measure of predictability of performance in a sport: the smaller the within-athlete variability relative to the between-athlete differences, the more predictable the sport. One such measure is the intraclass correlation coefficient (ICC), defined as the proportion of the differences between athletes that are not due to the within-athlete variability, or the “pure” athlete differences expressed as proportion of the observed athlete differences. This definition of ICC for competitive performance is equivalent to a test-retest correlation with competitions treated as tests in a reliability study (W.G. Hopkins, 2000).

There has been more than a decade of research aimed at estimating within-athlete variability for competitive performance in different sports. Estimates of the within-athlete variability provide the thresholds for assessing performance changes, but the comparison of such estimates between sports also provides insights into factors affecting performance in competition. Measures of predictability have also been reported in some of these studies recently. The present study is the first systematic review of this research.

2.2. Methods

Data Search and Study Selection

We searched the literature for estimates of the random variability in competitive performance of elite athletes from competition to competition for different sports. The search was limited to elite and senior performances, considering only athletes whose performances are at an international level and excluding performances of juniors and master athletes (e.g. triathlon performances included only athletes competing internationally in the professional circuit). For articles where athletes were ranked as upper-ranked and lower-ranked athletes, we chose to present only the estimates for the upper-ranked athletes. SPORTDiscus and Google Scholar were used as the databases, and search was limited to articles published up to October 2013. Initial searches for terms anywhere in the article produced an impractically large number of citations (13291 in SPORTDiscus or 38900 in Google Scholar), with too many citations not relevant for this literature search. In SPORTDiscus we restricted the search to the title, abstract and key words using the search terms *variability and (sport or athlete* or elite* competit*) and (performance or time)*. In Google Scholar two searches were performed, the first using the key term *smallest worthwhile change*, and the second using the option of searching related articles for one of the relevant citations (this option is limited to 101 citations). Relevant studies were selected first on the base of title and abstract, and then, after an investigation of the complete article. To be included the published article had to be written in English and report an estimate of within-athlete variability between competitions for a measure of competitive performance (time, distance, scores). We excluded three articles in which estimates of within-athlete variability between competitions were not presented. We excluded two studies in which estimates of

variability were derived for performances of non-elite athletes (juniors, masters and club-level athletes). One study was excluded because estimates of variability were for a surrogate of a performance (power output) as opposed to a measure of competitive performance. Seven studies were also excluded, because estimates of within-athlete variability or smallest worthwhile change were reported from previous published studies. Figure 1 summarises the search and selection process. No studies were excluded for reasons of poor quality. One study was initially excluded because of unrealistic large estimates of within-athlete variability for surfing performance scores (Mendez-Villanueva, Mujika, & Bishop, 2010); we subsequently contacted the authors, who provided the original data for further analysis.

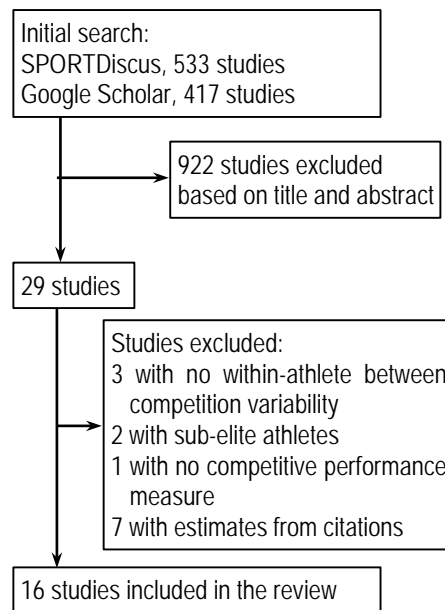


Figure 2. Schematic representation of study search and selection.

Data Extraction

Estimates of within-athlete variability between competitions within and between seasons for upper-ranked athletes are presented as coefficients of variation, with uncertainty expressed as 90% confidence intervals. Season is defined as a period of time with consecutive competitions. Duration of the season was reported in only four articles (W.G. Hopkins, 2005; McGuigan & Kane, 2004; Paton & Hopkins, 2005; Trewin, Hopkins, & Pyne, 2004), so this measure is not included in the table of results. In four studies (Bullock, et al., 2009; W.G. Hopkins, 2005; Muehlbauer, Schindler, & Panzer, 2010; Paton & Hopkins, 2005) confidence intervals were not reported, authors of three of these studies have provided the missing data. For the research of surfing performance scores, we estimated within-athlete variability and intraclass correlation coefficient using a straightforward reliability model in which competition identity was specified as fixed effect (adjusting for the each competition mean performance) and athlete identity as random effect (e.g. (Paton & Hopkins, 2005)). For estimates of CV of three studies (McGuigan & Kane, 2004; Paton & Hopkins, 2006; Pyne, Trewin, & Hopkins, 2004), 95% confidence limits were converted to 90%, using a published spreadsheet (W.G. Hopkins, 2006).

We estimated mean coefficients of variation for each sport or event type (e.g., track events and field events for athletics). Mean estimates were calculated first for the two genders separately (when data were available); but with the exception of weightlifting, differences between genders were trivial, so men's and women's estimates were combined into a mean for the sport. Mean coefficients of variation were estimated as the square root of the mean of the squares of coefficients of variation weighted by the degrees of freedom. If degrees of freedom were not reported (in the majority of the articles), they were estimated from the confidence limits of the CV. The degrees of freedom were estimated based on the sampling distribution of a variance being twice the square of the standard error divided by the degrees of freedom. Rearranging this relationship and assuming normality of sampling distribution, separate estimates for degrees of freedom were obtained for the upper and lower confidence limits. It was found empirically that adding one to the geometric mean of these two estimates gave an accurate estimate of the degrees of freedom for the coefficient of variation of each event. The 90% confidence limits of the mean coefficient of variation were obtained using a published spreadsheet (W.G. Hopkins, 2007a).

Estimates of within-athlete variability between competitions between seasons were reported only in four studies. In two of these studies (Bullock, et al., 2009; Smith & Hopkins, 2011), only the pure between-season variability was reported, which is the variability additional to the within-season variability. The observed between-season variability was derived as the square root of the sum of the squares of the within-season variability and the pure between-season variability. Only the means for each sport are shown for the between-season variability.

Predictability of performance from competition to competition is presented as intraclass correlation coefficient (ICC) with uncertainty as 90% confidence limits. Owing to paucity of estimates of ICC for the different events and sports, we have chosen to present only mean ICC for each sport, calculated using a published spreadsheet (W.G. Hopkins, 2006). ICCs are shown for all athletes in competitions (opposed to the within-athlete variability, which are presented for the upper-ranked athletes only).

A magnitude-based inferences approach was used to assess differences for CVs and ICC (Smith & Hopkins, 2011). Briefly, the usual thresholds for evaluating a difference in mean performance (0.3, 0.9, 1.6, 2.5 and 4.0 of the within-athlete competition-to-competition CV for small, moderate, large, very large and extremely large, respectively) are halved when comparing differences in CVs, because doubling the CV (or halving the threshold) is consistent with evaluating 2 SD of a virtual covariate that fully explains the CV (W.G. Hopkins, et al., 2009). For example, if one CV is 3%, the threshold for a small difference is 0.45% ($0.3/2 \times 3\%$), and the CV that represents such a difference is 3.45% ($3\% + 0.45\%$) or a ratio of 1.15 ($3.45\%/3.0\%$). Expressed as ratios, the thresholds are independent of the absolute values of CV and are 1.15, 1.45, 1.8, 2.25 and 3, for small, moderate, large, very large and extremely large differences.

Thresholds for evaluating differences in ICCs were based on the assumption that a 2-SD difference in performance between athletes in one competition should predict a small, moderate, large,... difference between those athletes in another competition, if the correlation is small, moderate, large,... (W.G. Hopkins, et al., 2009). The thresholds were then derived via the well-

known equation, $\Delta Y/SD_Y = r \cdot \Delta X/SD_X$, where Y is performance in the second competition, X is performance in the first competition, r is the correlation (here the ICC), ΔY and ΔX are the changes in performance in the competitions, and SD_Y and SD_X are the standard deviations in the competitions. The ΔX is $2SD_X$ and the ΔY is $f \cdot SD_W$, where SD_W is the within-athlete variability and f is 0.3, 0.9, 1.6,... for a small, moderate, large,... correlations. Making use of two further relationships, $SD_Y = \sqrt{(SD_B^2 + SD_W^2)}$ and $ICC = SD_B^2 / (SD_B^2 + SD_W^2)$, where SD_B is the true SD between athletes, the equation can be rearranged into a quadratic equation of ICC in terms of f, $(4/f^2)ICC^2 + ICC - 1 = 0$, with the positive solution being given by $f^2(\sqrt{(1+16/f^2)} - 1)/8$. The final thresholds for magnitudes of ICCs and differences between ICCs were 0.14, 0.36, 0.54, 0.69 and 0.83 for low, moderate, high, very high and nearly perfect correlations.

2.3. Results

Table 3 shows the retrieved estimates of the random variability shown by upper-ranked athletes during a competitive season. The lowest within-athlete variability was 0.15% for speed-skating, while the majority of estimates ranged from 0.6% (swimming and men's lightweight four) to 3.4% (women's javelin). Surfing scores with a variability of 39% were a clear outlier.

In Table 4, mean estimates for within-athlete variability and intraclass correlation coefficients for competitions within season and between seasons are presented for each sport, summarizing their typical uncertainties. Mean estimates for the sport were derived although CVs of the different events within a sport could be different (e.g., men's long jump and men's pole vault).. Estimates of variability within-season ranged from 0.6% to 1.4% for sprint and endurance sports, such as running and cycling, while for sports requiring explosive power in a single effort, such as field events and weightlifting, estimates ranged from 1.4% to 3.3%. Relatively higher variability was observed for the sports of canoe slalom, mountain biking and surfing. Surfing was the only sport with subjective scoring for which published estimates of within-athlete variability were available. To investigate whether other sports with subjective scoring have high CVs, we have analysed data from international performances for half-pipe freestyle skiing and snowboarding between 2010 and 2014. The within-season CVs were 34% (90% confidence interval 30-39%) and 53% (49-58%), and the ICC were 0.45 (0.32-0.57) and 0.17 (0.08-0.26), respectively.

Between-season variability was higher by a factor of ~1.1-1.2 (10%-20%) than variability in competitions within-season. Sports showed a wide range of predictability, and, with the exception of weightlifting, differences between mean estimates of variability and predictability for men and women were trivial.

Table 3. Estimates of within-athlete variability from competition to competition within a season separately by event and sport (mean and 90% confidence interval). Information regarding number of races included in the analysis, type of competition and type of competitors is also shown. Season has different duration dependent on event and sport.

Study	Sport	Number races	Competition	Type of competitors	Gender	Event	Perf. Measure	Variability within season (%)	
								Mean	90% CI
Hopkins (2005)	Track-and- field athletics	≤17	1997 Grand Prix Series	Top half in each competition	Men	Running <3km ^b	Time	0.8	0.7-0.9
						Running 3-10km ^c		1.1	1.0-1.3
						Triple jump	Distance	3.1	1.8-13.6
						Long jump		2.3	1.8-3.1
						Pole vault		1.9	1.5-2.6
						Discus		1.0	0.7-1.8
					Women	Javelin	Time	1.7	3.2-9.0
						Running <3km ^b		1.0	0.9-1.0
						Running 3-10km ^c	Distance	1.1	1.0-1.2
						High jump		1.6	1.1-2.7
						Triple jump		1.8	1.3-2.9
						Javelin		3.4	2.6-5.2
Nibali (2011)	Canoe- Slalom, Non- penalised	?	2000-2007 World Cup, World Champs and Olympics	Top half of finalists	Men	C1 ^d	Time	1.2	1.1-1.3
						C2 ^d (double canoe)		1.7	1.5-1.9
						K1 ^d		1.0	0.9-1.1
					Women	K1 ^d		1.5	1.4-1.7
	Canoe- Slalom, Penalised	?	2000-2007 World Cup, World Champs and Olympics	Top half of finalists	Men	C1 ^d		1.6	1.5-1.8
						C2 ^d (double canoe)		2.1	1.9-2.3
						K1 ^d		1.2	1.1-1.3
					Women	K1 ^d		1.9	1.7-2.1
Bonneti (2010)	Canoe- Sprint	?	2003/07 World Cup, World Champs and Olympics	"A" finalists	Men	C1-200 ^d	Time	0.9	0.7-1.2
						C1-500 ^d		1.1	0.9-1.3
						C1-1000 ^d		0.7	0.4-1.1
						K1-200 ^d		1.1	0.9-1.3

Bonneti (2010)	Canoe-Sprint		2003/07 World Cup, World Champs and Olympics	“A” finalists	Women	K1-500 ^d	Time	1.0	0.8-1.2
						K1-1000 ^d		0.8	0.6-1.1
						K1-200 ^d		1.2	0.9-1.6
						K1-500 ^d		0.7	0.6-0.8
						K1-1000 ^d		1.3	1.0-1.7
Flyger (2009)	Cycling Track	25	2005-2006 World Cups World Champs Commonwealth	?	Women	200-m Flying sprint	Time	1.3	1.1-1.6
		15				Individual pursuit		1.2	1.0-1.6
		12				500-m time trial		0.7	0.6-1.0
		33			Men	200-m Flying sprint		1.3	1.1-1.5
		13				Individual pursuit		1.0	0.8-1.3
		14				Team pursuit		1.0	0.8-1.3
		12				1-km Time trial		0.7	0.6-1.0
Paton (2006)		6	US national	Top half in each competition		1-km Time trial		1.2	0.8-1.8
Paton (2006)	Cycling Road	10	2000 World Cups	Top 8 in the each competition	Men	Road race	Time	0.4	0.3-0.5
		18	Tour de France	Top 8 in each competition				0.7	0.7-0.7
		3	World Champs and Olympics	Top half in each competition		Time trials		1.3	0.9-2.0
		2	Tour de France	Top 8 in each competition				1.7	1.2-2.4
		7	World Cups	Top quarter in each competition	Men			2.4	2.1-2.7
Paton (2006)	Cycling Mountain Bike	8			Women			2.5	2.2-2.9
Spencer (2014)	Cross- Country Skiing ^e	28	2002-2011 World Cup, World Champs and Olympics	Top 10 over the entire season	Men	15-km Classical	Time	1.2	1.1-1.3
						15-km Free		1.1	1.0-1.2
						Sprint Classical		1.2	1.1-1.3
						Sprint Free		1.2	1.0-1.2
					Women	10-km Classical		1.4	1.3-1.5
						10-km Free		1.2	1.1-1.3
						Sprint Classical		1.3	1.2-1.4
						Sprint Free		1.3	1.2-1.4

Smith (2011)	Rowing	46	1999- 2009	“A” Finalists	Men	Single scull	Time	1.2	1.0-1.4
						Lightweight double sculls		0.7	0.6-0.8
						Double sculls		1	0.9-1.2
						Pair		0.9	0.8-1.0
						Lightweight four		0.6	0.5-0.7
						Four		0.7	0.6-0.8
						Quadruple sculls		1.1	1.0-1.3
					Women	Eight		0.8	0.7-0.9
						Single		1	0.9-1.2
						Lightweight double sculls		1	0.9-1.2
						Double sculls		0.9	0.8-1.0
						Pair		0.9	0.8-1.0
						Quadruple sculls		0.9	0.8-1.0
						Eight		0.8	0.7-0.9
Bullock (2009)	Skeleton	22	2000-2006 World Cup	Top 10 in each competition	Men		Time	0.4	0.3-0.5
		25			Women			0.4	0.4-0.5
Muehlbauer (2010)	Speed Skating	5	2007-2008 World Cup	Upper ranked athletes over the entire season	Men	1000-m Sprint	Time	0.15	
					Women			0.33	
Mendez (2010)	Surfing	12	2002 World Champs Tour	Upper ranked athletes over the entire season			Scores	39	35-43
Trewin (2004)	Swimming	~13	2000 Olympics	Top 50 world-rank athletes	Men	Average for all events	Time	0.8	0.7-0.8
					Women			0.8	0.8-0.9
Pyne (2004)	Swimming	~52 ^f	1999/2000 Olympic Pan Pacific Champs,	26 US and 25 AUS Olympic Swimmers	Both	Freestyle	Time	0.6	0.4-0.9
						Breaststroke		0.8	0.5-1.3
						Butterfly		1.0	0.7-1.4
						Backstroke		0.6	0.5-0.8
						Individual medley		0.7	0.4-1.1

Pyne (2004)	Swimming	~52 ^f	1999/2000 Olympic Pan Pacific Champs	26 US and 25 AUS Olympic Swimmers	Both	50 + 100 m	Time	0.7	0.6-0.8
						200 + 400 m		0.7	0.6-0.8
						800 + 1500 m		1.0	0.7-1.3
Fulton (2009)	Paralympic Swimming 100-m freestyle	15	2004-2006 National Trials, Paralympics and World Champs	Finalists (top 8 in each competition)	Men	S2-S4 ^g	Time	3.7	2.9-5.3
						S5-S7 ^g		1.2	0.9-1.9
						S8-S10 ^g		1.3	1.1-1.7
						S11-S13 ^g		2.4	1.7-4.2
					Women	S2-S4 ^g		2.9	2.0-5.0
						S5-S7 ^g		2.6	2.1-3.4
						S8-S10 ^g		1.7	1.4-2.1
Paton (2005)	Triathlon	9	1997/1998 International Triathlon Races	Top 10% athletes over an entire season	Men	S11-S13 ^g	Time	1.5	1.2-2.1
						Total		1.1	0.9-1.4
						Swim		1.2	1.0-1.6
						Cycling		1.3	1.1-1.6
McGuigan (2004)	Weightlifting	?	1998/2000 National and International Competitions	Top 5 ranked athletes at 2000 Olympics	Men	Run	Weights (Scores)	2.5	2.1-3.1
						Total		1.7	1.4-2.1
						Snatch		1.9	1.6-2.3
					Women	Clean and jerk		2.0	1.7-2.4
						Total		3.3	2.7-4.1
						Snatch		3.6	2.9-4.5
						Clean and Jerk		3.7	3.0-4.6

^aCI Confidence Interval.

^b100-m to 1500-m runs; 100- to 400-m hurdles.

^c1500- to 10000-m runs; male 3000msteeplechase.

^dC for canoe events and K for Kayak events; 200, 500 and 1000 represent the event distance in metres.

^eonly interval start races.

^f13events*4 competitions.

^gS2-S0 most through least physically impaired and S11-S13 most through least visually impaired.

Table 4. Mean estimates and confidence intervals of within-athlete variability for competitions within and between seasons (expressed as coefficient of variation and 90% confidence intervals), and mean intraclass correlation coefficients (ICC) and 90% confidence intervals for competition within and between seasons for the sports identified in Table 3.

Sport	Class of event	Variability (%)				Intraclass correlation coefficient			
		within season		between season		within season		between season	
		Mean	90% CI	Mean	90% CI	Mean	±90% CI	Mean	±90% CI
Athletics	Track	1.1	1.0-1.1						
	Field	1.4	1.3-1.6						
Canoe slalom, non-penalised		1.4	1.3-1.4						
Canoe slalom, penalised		1.7	1.7-1.8			0.31	0.16	0.27	0.16
Canoe sprint		1.0	0.9-1.1						
Cycling track	Sprint/Pursuit	1.2	1.1-1.3						
	Time trial	0.8	0.7-0.9						
Cycling road ^b	Road race	0.6	0.5-0.6						
	Time trials	1.5	1.2-2.1						
Cycling mountain bike		2.4	2.2-2.7						
Cross-country skiing		1.2	1.2-1.3	1.3	1.2-1.3	0.93	0.00	0.90	0.00
Rowing		0.9	0.9-0.9	1.0	0.9-1.1	0.65	0.04	0.56	0.04
Skeleton		0.4	0.4-0.5	0.5	0.5-0.6	0.36	0.19	0.15	0.21
Speed skating	1000-m sprint	0.2				0.74			
Surfing ^b		39	45-43			0.18	0.07		
Swimming		0.8	0.7-0.9						
Paralympic swimming	100-m freestyle	2.0	1.9-2.2	~2.7		0.88			
Triathlon ^b		1.1	0.9-1.4						
Weightlifting	Men	1.7	1.4-2.1						
	Women	3.3	2.7-4.1						

^aICC is Intraclass correlation coefficient and CI is confidence interval.

^bonly men's data was used

2.4. Discussion

In this systematic review we have investigated data from 16 original studies, reporting estimates of within-athlete variability and predictability of competitive performance for different events and sports. Expressing variability as a coefficient of variation (CV) enables comparison of variability of events within sports and comparison of mean variability between sports, which would not be possible using performance scores in their original units.

Variability in performance must arise from variability in the factors that affect performance, including power output, environment, race dynamics, skill, and subjective scoring. Sports that differ in their CVs must differ in the relative contributions of these factors.

For most of the sports in the tables, performance is determined mainly by the athlete's ability to output power to achieve a shortest performance time. For elite athletes, one would expect similar physiological variability in power output from competition to competition, regardless of the sport. However, CVs for power output do not translate directly into CVs for performance time in many sports. The contribution of physiological variability in power output to variability in performance time is determined by the relationship between power (P) and velocity (V) for the sport. This relationship can be expressed as $P=kV^x$, where k and x are constants for the sport (Will G Hopkins, Schabort, & Hawley, 2001). For example, x has values of 3 for rowing and canoe-sprint, ~2.3 and ~2.0 for cycling and swimming, and 1.0 for running (Will G Hopkins, et al., 2001). Cross-country skiing, mountain biking, and any other sport similar to running, where most of the work is done against gravity, will probably all have values of ~1.0. For the small changes in velocity represented by the CVs in performance times, application of calculus to the power-velocity relationship shows that the CV for power is equal to the CV for velocity or time multiplied by the constant x. A 1% change in power output would therefore lead to a 1% change in performance time for running and similar sports but a ~0.3% to 0.5% (1/3 to 1/2) change in performance times for rowing, canoe-sprint, cycling and swimming. However, CVs of performance times for these sports were similar to those for running, implying that factors other than physiological production of power must contribute to their variability. The most likely factors are environmental, whereby changes in the environment differ between athletes from one race to the next. For example, in an endurance cycling time trial, application of calculus to the power-velocity relationship shows that a net increase in wind speed of only 1 km.h⁻¹ would increase a cyclist's performance time by ~1.5%.

Race dynamics may explain the remarkably low values of CV for speed skating and skeleton times compared with those of other sprint sports. Performance in the initial sprint phase largely determines the differences in total performance time (Zanoletti, La Torre, Merati, Rampinini, & Impellizzeri, 2006). The percent variability in the initial phase, which is presumably similar to that in other sprint events, will manifest as a lower percent variability when expressed as a percent of the total performance time (Bullock, et al., 2009).

The higher values of CV for sports requiring explosive power in a single effort (athletics field events and weightlifting) may reflect the importance of skill in these performances. In a review of

movement variability of javelin and discus throws, Bartlett and colleagues (Bartlett, Wheat, & Robins, 2007) reported that even elite athletes show variability in their patterns of movement, which could lead to variability in performance. Skill in addition to environment could also be responsible for the even larger CV of canoe-slalom and mountain biking. Skill and environment are also likely factors affecting variability in performance in the sports with the largest CV (surfing, half-pipe freestyle skiing and snowboarding), but subjectivity in scoring appears to be the greatest contributor here.

Estimates of within-athlete variability for competitions between seasons represent the random variation shown by athletes from one season to the next and are therefore important for assessing magnitude of long-term performance changes in competitive performance. As expected athletes showed a larger variability in performance for competitions between seasons than for competitions within season, reflecting greater differences in performances over a longer time interval. Skeleton had the lowest CV and Paralympic swimming had the highest CV, presumably reflecting greater skill demands in some Paralympic classes.

In some studies estimates of predictability of the sport were also reported as intraclass correlation coefficients (ICC), with values ranging from low to nearly perfect predictability according to the scale of magnitudes defined in (Smith & Hopkins, 2011). Surfing and other sports with low predictability require a large number of competition scores to establish a precise ranking of athletes' abilities. For highly predictable sports, such as cross-country skiing, placings in any single competition are a good representation of a ranking of athletes' abilities.

Within-athlete variability and predictability were combined for men's and women's events, with exception of weightlifting, the only sport/event with a large difference between CVs. For the particular case of weightlifting, the greater within-athlete variability of female athletes is likely due to the lower level of competitiveness of women's field, as this event was added later to the Olympic program (McGuigan & Kane, 2004).

Upper-ranked athletes often have less variability compared to that of the other athletes (Bonetti & Hopkins, 2010; Bullock, et al., 2009; W.G. Hopkins, 2005; McGuigan & Kane, 2004; Nibali, et al., 2011; Paton & Hopkins, 2005, 2006; Smith & Hopkins, 2011; Spencer, et al., 2014), who may be inconsistent in their training and in their effort in competitions if a medal is unlikely (McGuigan & Kane, 2004). We have reported estimates of within-athlete variability for upper-ranked athletes only, because we are interested in estimates of variability that provide the smallest important change for winning medals. However, the definition of *upper-ranked athlete* varied across studies. In five studies (Bonetti & Hopkins, 2010; Bullock, et al., 2009; Fulton, et al., 2009; Nibali, et al., 2011; Paton & Hopkins, 2006) those athletes were identified from top placings in each competition (e.g., the top 10 in skeleton competitions (Bullock, et al., 2009)); consequently, poorer performances of these athletes in other competitions were not included in the estimation of variability. This approach of selection of best performances underestimates the variability for deriving the smallest important change in performance. In other studies, upper-ranked athletes were selected either by using a seasonal ranking provided by the sport (e.g. FINA world-ranking for swimming (Trewin, et al., 2004)), or by producing a ranking based on athletes' average performances over the season (W.G. Hopkins, 2005; Paton & Hopkins, 2005; Spencer,

et al., 2014). These averages were estimated using a simple mixed model in which performance was explained by competition identity as a fixed effect and athlete identity as a random effect. The random-effect solution for athlete identity provided athletes' averages properly adjusted for any competitions that the athletes did not enter. The analysis of such seasonally identified upper-ranked athletes provides accurate estimates of the variability for defining smallest important change in performance. Comparison of estimates of variability derived using the two approaches is needed; for example, variability of top 10 performances in each cross-country skiing competition and of the 10 upper-ranked skiers over the entire season differed by a factor of ~1.2 to 1.7 (WG Hopkins, unpublished observations). There is also a need for better ways to estimate variability of upper-ranked athletes than with an arbitrary subset (e.g., the top 10); for example, use of a random effect to specify a gradual increase in within-athlete variability for athletes who are less well ranked.

The usual mixed modelling approach for deriving estimates of variability as in previous studies (Bonetti & Hopkins, 2010; Bullock, et al., 2009; Fulton, et al., 2009; W.G. Hopkins, 2005; Nibali, et al., 2011; Paton & Hopkins, 2005, 2006; Pyne, et al., 2004; Smith & Hopkins, 2011; Spencer, et al., 2014; Trewin, et al., 2004) overestimates variability between consecutive competitions, because variability is calculated for all competitions within a season regardless of how far apart in time the competitions were. Typical variability from one competition to the next could be estimated either by analysing separately each consecutive pair of competitions and then averaging the variances, or by specifying a spatial distribution for the variance in the mixed model, assuming that performances at competitions further apart are less related in some systematic manner. Future research on variability of competitive performance should utilise one or other of these approaches to properly estimate variability between consecutive competitions.

Estimation of thresholds for smallest and other important changes in performance as fractions and multiples of within-athlete variability between competitions is an important practical application of the research reviewed here. The thresholds apply directly to any studies where performance is investigated in competitions, but almost all studies of factors affecting performance are conducted with laboratory or field tests, under the assumption that effects on performance in such tests will translate into effects of similar magnitude in competitions. Conditions extrinsic and intrinsic to the athlete (such as environment, tactics, motivation, and anxiety) inevitably differ between a competition and a test, and such differences could modify the effects. For example, caffeine might have little or no effect in competitions compared with that in a laboratory test, if the only effect of the caffeine is to increase motor drive in the test towards the higher level experienced in competitions. Nevertheless, researchers can legitimately investigate effects in the laboratory or field, and in the absence of evidence to the contrary, they have to assume that these effects will translate into effects of equal magnitude in competitions. With this assumption, the thresholds for assessing changes in competitive performance can be used directly to assess performance of athletes in a laboratory or field test, provided the test reproduces the demands of the competition and the measure of performance in the test is in the same units as those of performance in the competition. When the protocol or the units of

measurement in the test differ from those of the competition, transformation of the competition thresholds may be necessary.

If performance in the test represents the athlete's ability to output power, the best approach to transformation is to convert the performance thresholds in the competition and the effects on performance in the test into percentages of mean power. This approach accommodates sports in which environmental, skill or other factors make performance in the competition more variable than in the test, because it addresses the question of the enhancement of power output needed to overcome the variability in the competition and thereby increase an athlete's chances of winning a medal. The conversion of the performance thresholds depends on the power-velocity relationship for the sport as described above ($P = kV^x$). If the test is performed on an ergometer and the outcome can be expressed as a percent effect on mean power, no further transformation is needed, but otherwise the test protocol may have to be taken into account. Percent effects on time for a fixed distance or distance for a fixed time will need to be converted to percent effects on power using the power-velocity relationship for the ergometer, which may differ from that for the competition. Effects on performance in incremental tests need careful treatment when they are converted to effects on power. If the outcome is measured on an ergometer as peak power, it is reasonable to assume percent effects in the test apply directly to mean power in the competition (Will G Hopkins, et al., 2001). Percent effects for time in incremental tests that start at zero speed or power and increase linearly to maximum effort are equivalent to percent effects on power, but percent effects on time in incremental tests that start at a proportion (p) of a ramp or series of linear steps to peak power or speed have to be reduced by a factor of $1-p$. Percent effects on time to exhaustion in a constant-load test are also problematic; these have to be converted to percent change in power output in a time trial of similar duration via the power-duration relationship for sub-VO₂maximal exercise, typically by dividing by ~15 (Will G Hopkins, et al., 2001) and by smaller factors derived from the critical-power or log-log relationship for supramaximal exercise (Hinckson & Hopkins, 2005).

This review of literature revealed some areas requiring further research. The contributions of the factors affecting variability in competitive performance in various sports needs further investigation, because understanding these contributions may provide avenues for performance enhancement. There is a need to replicate the studies of variability for athletics field events and Paralympic swimming with larger samples to achieve estimates with acceptable uncertainty. For purposes of properly assessing changes in performances of individual athletes between seasons, further research is needed to provide estimates of between-season variability and predictability of competitive performance. The majority of the published articles were for sports with objective measures of performance (time, distance or score), but there is a need for research on variability of performance in judgement-based sports, such as diving, gymnastics and combat sports. Finally, variability of competitive performance needs investigation in team sports.

2.5. Conclusion

There was a wide range of within-athlete variability in performance for the different events and sports. We argued for various factors affecting the variability: physiological and skill demands of the sport, dynamics of competition, environmental and other external factors affecting performance, and the measure of competitive performance itself (time, distance, scores). In this review we focused on the within-athlete variability as a statistic to derive the smallest and other important magnitude thresholds for changes in competitive performance. These thresholds can be transformed into thresholds for research on factors affecting performance in laboratory and field tests.

CHAPTER 3

TRACKING CAREER TRAJECTORIES OF SUCCESSFUL TRIATHLETES

This chapter comprises the following paper accepted for publication at *Medicine and Science in Sports and Exercise*:

Malcata, R. M., Hopkins, W. G., & Pearson, S. (2014, in press). Tracking career performance of successful triathletes. *Medicine and Science in Sports and Exercise*.

Overview

Purpose: Tracking athletes' performances over time is important but problematic for sports with large environmental effects. Here we have developed career performance trajectories for elite triathletes, investigating changes in swim, cycle, run stages and total performance times while accounting for environmental and other external factors. **Methods:** Performance times of 337 female and 427 male triathletes competing in 419 international races between 2000 and 2012 were obtained from triathlon.org. Athletes were categorized according to any top-16 placing at World Championships or Olympics between 2008 and 2012. A mixed linear model accounting for race distance (Sprint, Olympic), level of competition, calendar-year trend, athlete's category and clustering of times within athletes and races, was used to derive athletes' individual quadratic performance trajectories. These trajectories provided estimates of age of peak performance and predictions for the 2012 London Olympic Games. **Results:** By markedly reducing the scatter of individual race times, the model produced well-fitting trajectories suitable for comparison of triathletes. Trajectories for top-16 triathletes showed different patterns for race stages and differed more among women than among men, but ages of peak total performance were similar for men and women (28 ± 3 y, mean \pm SD). Correlations between observed and predicted placings at Olympics were slightly higher than those provided by placings in races prior to the Olympics. **Conclusion:** Athletes' trajectories will help identify talented athletes and their weakest and strongest stages. The wider range of trajectories among women should be taken into account when setting talent-identification criteria. Trajectories offer a small advantage over usual race placings for predicting men's performance. Further refinements, such as accounting for individual responses to race conditions, may improve utility of performance trajectories.

3.1. Introduction

In the elite sport environment, monitoring of athletes' competition performances provide valuable information for guiding training programmes, setting performance goals, and selecting talented athletes. For sports with consistent race environments, such as swimming, competition

results offer a relatively reliable measure of athlete's ability. However, in many sports competition results are affected by external factors, and an athlete's ability should be estimated using appropriate models to account for the extra variation on performances arising from such factors. Triathlon, consisting of swim, bike and run race stages, is a particular example of a sport where a variety of environmental factors (temperature, wind, race course profile) influence race conditions and consequently performance outcomes.

Triathlon performance is affected by environmental conditions. Windy or wet conditions result in slower times across all three triathlon stages (swim, cycle and run), but other conditions can affect one stage in particular. For example, temperature determines whether or not it is a wetsuit swim, and the wetsuits typically results in better swim performances (Chatard, Senegas, Selles, Dreanot, & Geyssant, 1995; G. Millet & Vleck, 2011). Race courses are another source of variation of performance times. Swim courses vary in geometry (buoy distance and configuration), in water current, and in whether the water is fresh or salty; cycle and run courses vary in elevation and number of sharp curves, and even the distance of the cycle stage can vary up to 10% of the standard 40 km (ITU, 2013). These environmental effects should be taken into account when assessing triathletes' performances.

Studies in age-related performance changes in triathlon have been done mainly by Lepers and his colleagues. For Olympic distance triathlon, effects of age and gender were analysed using mean times of top triathletes (Bernard, et al., 2010; Etter, et al., 2013; Lepers, Sultana, Bernard, Hausswirth, & Brisswalter, 2009; Stevenson, et al., 2012; Young & Starkes, 2005). While this approach provided an important general understanding of age and gender differences in triathlon performance, a different approach is needed to monitor individual athletes over time.

The analysis of long-term career changes in performance of individual athletes using competition results has been developed for track-and-field athletics (Berthelot, et al., 2012; Hollings, et al., 2012; Young & Starkes, 2005), swimming (Allen, et al., 2012; Berthelot, et al., 2012; Pike, et al., 2010), skeleton (Bullock & Hopkins, 2009), and cross-country skiing (Alam, et al., 2008). Bullock et al. (Bullock & Hopkins, 2009) developed individual performance trajectories for a 4-year Olympic cycle, using a linear model to adjust for widely different race times arising from weather conditions and course profiles. Hollings et al. (Hollings, et al., 2012) estimated specific environmental and other venue-related effects directly, using a model that included age-related performance changes over competitive career of track-and-field athletes. By building on these two studies, the purpose of our study was to develop an analytical tool for tracking the progression of performance of individual triathletes over their competitive career. Specifically, we have combined the approach of Bullock et al. (Bullock & Hopkins, 2009) to account for the large environmental factors in triathlon with that of Hollings et al. (Hollings, et al., 2012) to describe the age-related changes in performance. In addition, the analysis has provided benchmarks guides for talent selection policies, using profiles of successful triathletes, and allowed prediction of future race performances.

3.2. Methods

Data

Official times of international triathlon races from World Cup, European Championships, World Triathlon Series (named previously World Championship Series), World Championships (including Junior and Under-23) and Olympic Games were downloaded from triathlon.org. Race dates were obtained also from triathlon.org. We searched the Internet for each athlete's date of birth, using as a primary source infostradasports.com. Athletes were excluded when date of birth was not found (150 male and 250 female athletes, mainly competing during the early 2000s and with a low number of performances) and performances were not included when athletes were disqualified or did not finish. Overall, 446 international competitions (224 men's and 222 women's) were raced over the period from 2000 to 2012, with a total of 427 male and 337 female athletes competing in at least two of those competitions. Athletes were categorized as top athletes if they finished 16th or better at any World Championship or Olympic Games between 2008 and 2012. This classification was used to create performance benchmarks. For purpose of developing career trajectories these athletes were further separated into Top 3 (10 male and 15 female), having ever finished 1st-3rd, and Top 16 (43 male and 38 female), best finish of 4th-16th. The remaining athletes were grouped as Others.

Career trajectories

Individual performance trajectories were generated using the high-performance mixed linear model procedure (Proc Hpmixed) in the Statistical Analysis System (Version 9.3, SAS Institute, Cary, NC), (see APPENDIX A). The fixed-model included a mean quadratic trend for age, a linear trend for calendar year, and an intercept to adjust Sprint-distance into Olympic-distance times (providing two overall means race times for Sprint and Olympic-distance races); all of these factors were interacted with athlete grouping factor with three levels (*Top 3*, *Top 16*, *Others*). Random effects were included to derive individual age quadratic trends and year-adjustments for each athlete, the latter representing consistent deviation from the quadratic fitting in a particular year (due to injury, new training, etc.). A random effect for race accounted for mean effect of environmental and other course-related factors (e.g., varying weather and water conditions, course distances, profiles of cycle and run courses) on performance times¹. An unstructured covariance matrix was specified for the random effects representing the individual

¹Triathlon performance need to be normalised because triathlon races are not performed in a standard environment. The normalization is achieved in mixed modelling by including identity of the athletes (random effect for athlete), which estimates mean abilities for athletes, and by clustering performances within a race (random effect for race), which identifies changes in performance that are the same for all the athletes that competed in that particular, identifying slower and faster races compared to a mean. The slow and fast races (after accounting for athlete identity) arise from effects of environmental and other course-related factors in performance times, because these changes are computed as all the athletes competed in every race.

quadratic trajectories to allow for correlation between the three parameters defining the trajectories. The residual random effect, representing race-to-race athlete variability, was specified differently according to the type of competition (three levels: *Elite World Championships, Olympic Games and World Triathlon Series; World Cup, European Championships and U23 World Championships; Junior World Championships*). This model was applied to each stage (swim, cycle, run and total time) and to each gender separately. Race times were log-transformed to yield the effects and errors in percent changes from the mean. Observations were considered outliers and excluded from the analysis if standardized residuals were greater than 4 standard deviations from the predicted value; 54 swimming, 88 cycling, 80 running and 63 total performances were thereby excluded from the analyses. The appropriateness of the model was investigated by analysis of residuals: plots for residuals vs predicted performances were inspected to ensure there was no unacceptable non-uniformity, and residuals were plotted against age (centred on age of peak performance) to ensure no substantial systematic trend in the residuals on either side of age of peak performance (the minimum of each quadratic curve). The quadratic model was deemed appropriate to represent performance changes with age.

Conversion of Sprint to Olympic times

A conversion factor for adjusting Sprint times to Olympic-distance times is implicit in the trajectories model as the difference between the two levels of the corresponding fixed effect (the two intercepts used in the model). There were little differences between the conversion factors of the three groups (Top 3, Top 16 and Others), so factors were averaged. To generate a conversion factor of the most use to the sport, we repeated the analysis excluding junior and Under-23 performances, although the resulting conversion factor was very similar to that for all athletes.

Age of peak performance and age-related performance change

Athletes' age of peak performance was determined as the minimum of the individual quadratic age trend. For any individual trajectory that did not show the expected quadratic behaviour, age of peak performance was not determined and their values did not contribute to the mean. For the included athletes, age of their best placing at a World Championships or Olympic Games (age of best performance) was also identified. Age-related performance change was calculated for each athlete as the performance change over the 5-year period leading to age of peak total performance. Age of peak performance and age-related performance change for swimming, cycling, running stages and total performance are presented as means and standard deviations. These statistics were compared and magnitudes of standardized differences were assessed using thresholds of 0.2, 0.6 and 1.2 for small, moderate and large, respectively (W.G. Hopkins, et al., 2009). Uncertainty was calculated as 90% confidence limits.

Performance benchmark

A benchmark range for performance changes with age, representing the typical age-related performance changes among successful athletes, was obtained by combining the career trajectories of Top athletes. The mean performance for each year of age was calculated using a

meta-analytic model, where age-estimates from individual athletes were combined using the inverse of the square of the standard error of the estimate as a weighting factor. The upper and lower limits of the benchmark range were calculated as the 90% reference range, assuming a normal distribution with mean (as just described) and standard deviation given by the square root of the sum of the square of the standard error of estimated mean and the between-athletes variance, both provided by the meta-analytic model.

Prediction of race outcomes

Performances at 2012 London Olympic Games were predicted from the last race before the London Olympics (2012 Hamburg World Triathlon Series race). The mixed model for developing trajectories was modified by including venue as a predictor to characterize the effect of specific venues. Athletes' performances were predicted by extrapolating each athlete's quadratic trajectory to the dates of Olympic races and assuming the same conditions as in the 2011 London test event (which was held on a very similar race course and with a similar field of competition, as many countries used these performances for Olympic selection). We simulated 5000 (a large number of races) individual races taking into account athletes' race-to-race variability as follows. In each race, each athlete's performance was given by the sum of their predicted value plus a random unit normal deviate multiplied by the standard error of the predicted value and an extra component, derived by randomly selecting from the residuals obtained with the original mixed model for trajectories. Chance of winning (the proportion of the 5000 races won by each athlete) and ranking of the chances were then determined for athletes competing at the London Olympics. The predictability of performance was assessed by correlating the log-transformed rankings derived from chance of winning with the log-transformed observed race placings. The performance of the female athlete Paula Findlay was not included as it was a clear outlier: she had been injured, was not physically fully prepared and had not been competing internationally for more than a year (Yahoo, 2012). Correlations between log-transformed placings in Olympics and in each of 2012 World Triathlon Series races were also calculated and averaged using the Fisher transformation. The log-transformation of placings was used to give equal importance to percent or factor differences in placings rather than absolute differences. For example, the difference between second and first is equivalent to the difference between tenth and fifth with log transformation and to the difference between tenth and ninth without transformation. Correlations were assessed using 0.1, 0.3, 0.5, 0.7, 0.9 as thresholds for small, moderate, large, very large and extremely large (W.G. Hopkins, et al., 2009) and uncertainty expressed as 90% confidence limits. Predictability of race outcomes is affected by the variability of athletes' performances. Reliability analyses were performed for each calendar year for men and women to estimate typical differences between athletes and typical variation of athletes' total performance time from one race to the next. The typical differences between athletes and the typical variation in men's and women's performances (expressed as SD) were used to explain differences in correlations.

3.3. Results

Raw mean performance times (expressed as h:min:s) for Sprint and Olympic distance races for both gender are shown in Table 5. Variations between athletes within a race, representing the typical spread of performances in a race, are also displayed in Table 5.

Derived from the mixed model, mean times improved as a function of calendar year, at a rate of 0.25% (90% confidence limits, $\pm 0.13\%$), 0.13% ($\pm 0.12\%$) and 0.09% ($\pm 0.12\%$) for men and 0.35% ($\pm 0.17\%$), 0.24% ($\pm 0.12\%$) and 0.15% ($\pm 0.12\%$) for women, for the Top 3, Top 16 and Other athletes, respectively. The conversion factor for mean time between elite Sprint and Olympic distance races was 1.98 (± 0.04) for both genders. In addition to the uncertainties shown in Table 5, athletes' performances varied from one race to the next typically by 1.5% of the race time for men and 1.5% of the race time for women (at high level races: World Triathlon Series, Elite World Championships and Olympic Games). In swimming, cycling and running the typical race-to-race variabilities were 1.1%, 1.7% and 3.2% for men and 1.5%, 1.7% and 2.8% for women, respectively. Uncertainties in these estimates of variability were negligible. A simple reliability analysis allowing for within-athlete within- and between-year variability for the 55 athletes competing at the 2012 Olympic Games (with an average of 37 performances per male athlete and 32 performances per female athlete) showed a 1.6% and 1.5% for the race-to-race within year variability over the 13-year period for men and women, differences between athletes within year of 1.6% for men and 1.4% for women and additional differences between athletes from year to year 0.8% for men and 0.9% for women, resulting in intraclass correlation coefficient (ICC) between years of 0.45 (90% confidence limits, ± 0.09) for men and 0.38 (± 0.09) for women and within-year ICC of 0.56 (± 0.09) for men and women.

Table 5. Number of competitors (mean \pm SD), race times (mean \pm SD, min) and typical spread of performances in a race (mean of standard deviations on the performance times in each race) in 224 men's and 222 women's Sprint and Olympic distance triathlon races between 2000 and 2012.

	Men	Women
Sprint		
Competitors per race	60 \pm 16	46 \pm 13
Mean race time	59.5 \pm 4.9	66.7 \pm 5.6
SD of race time	\pm 2.7	\pm 3.1
Olympic		
Competitors per race	48 \pm 13	35 \pm 13
Mean race time	114.2 \pm 5.1	127.0 \pm 5.7
SD of race time	\pm 3.1	\pm 3.9

Figure 3 illustrates the difference between observed and corrected performances' times, after observed times being adjusted to a mean race. An athlete performance trajectory, displayed as the black curve, represents the fitting of a quadratic age trend to corrected times.

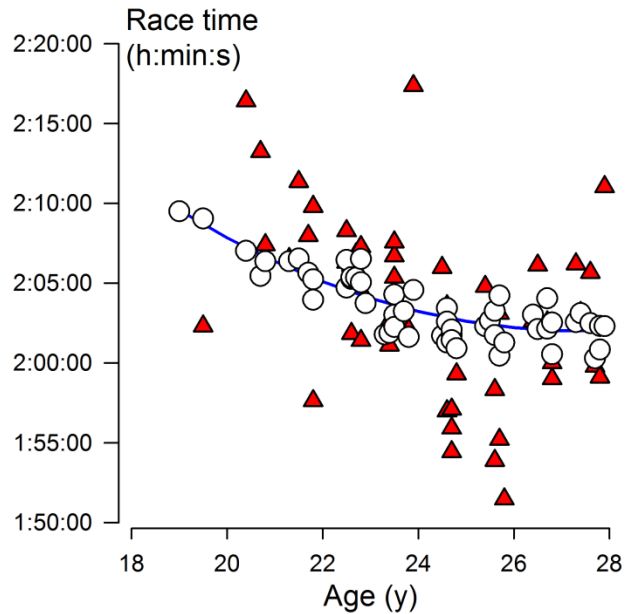


Figure 3. Observed (triangle) and corrected (circle) performances' times for an athlete, after times being adjusted to a mean race. The black curve is an athlete's trajectory, illustrating the fitting of a quadratic age trend to the corrected times.

Figure 4 shows the performance trajectories as functions of age for Top male and female athletes on the three stages individually and total race time. In this figure, Top 3 athletes' performance trajectories are displayed in black while Top 16 athletes' performances are presented in grey. Different patterns of progression were observed for swimming, cycling, running and total performances and between genders.

Performance trajectories were used to estimate age of peak performance and age-related performance change (Table 6). Differences between ages of peak performance for genders and stages are trivial to moderate in magnitude, and most of the substantial differences are clear. There was a trivial difference between age of peak performance and age of best performance (age of best placing in a race; data not shown) at World Championship and/or Olympic Games. Age-related performance changes were evaluated over the 5-year period leading to athletes' peak performance; with a negative change representing an improvement (decrease on performance times).

Table 6. Predicted age of peak performance and 5-year improvement for swim, bike, run and total performance for the “top” and all athletes (see Methods for athletes’ classification). Only athletes who had already reached their predicted age of best performance were included in these estimates (30 for Top men, 30 for Top women, 152 for Other men and 79 for Other women).

	Swim	Bike	Run	Total
Age of Peak Performance				
Top				
Men	25 ± 4	29 ± 3	28 ± 2	28 ± 2
Women	28 ± 2	28 ± 3	26 ± 5	27 ± 4
Other				
Men	26 ± 3	29 ± 2	29 ± 2	29 ± 2
Women	26 ± 5	28 ± 4	30 ± 3	28 ± 3
5-year improvement to peak performance (%)				
Top				
Men	-0.8 ± 0.5	0.7 ± 0.3	-3.5 ± 1.1	-0.9 ± 0.3
Women	-1.1 ± 1.0	-0.2 ± 0.4	-5.1 ± 2.6	-1.9 ± 0.9
Other				
Men	-0.9 ± 0.5	0.5 ± 0.2	-2.1 ± 0.8	-0.7 ± 0.1
Women	-0.5 ± 0.8	0.0 ± 0.3	-2.4 ± 0.8	-0.7 ± 0.2

Data are mean ± SD.

Uncertainties (90% confidence limits) for the means are ~0.3SD for top men and women and ~0.15SD for other men and ~0.2SD for other women.

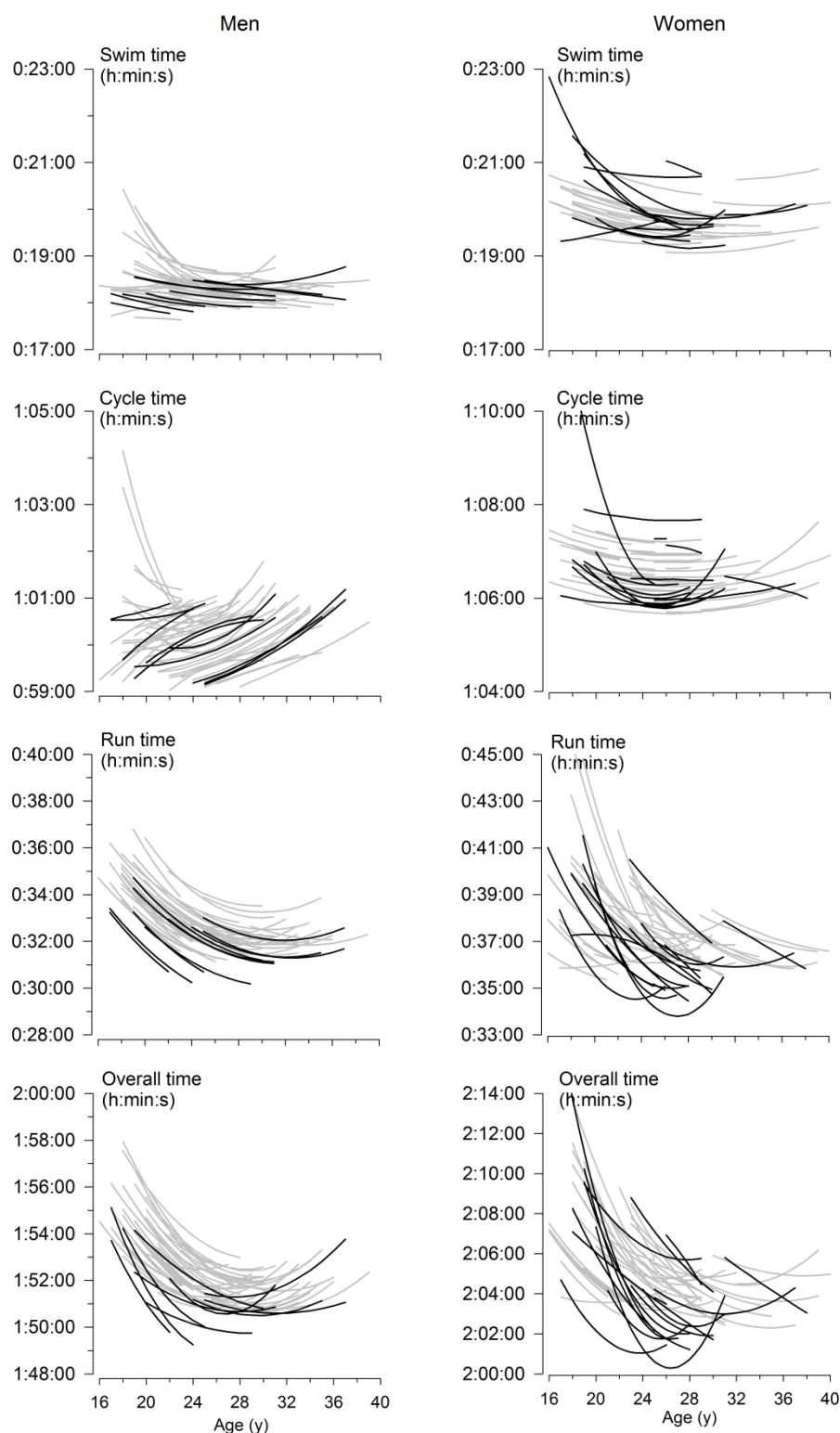


Figure 4. Performance trajectories as functions of age for top men and women for swimming, cycling, running and total performance time. Top 3 athletes' trajectories are displayed in black while Top 16 are presented in grey. (See Methods for athletes' classifications.)

Top athletes' performances were then combined to produce a performance benchmark range, representing the typical pattern of performance changes with age. As it shown in Figure 5, an athlete's performance trajectory and year performance can then be assessed against the benchmark range.

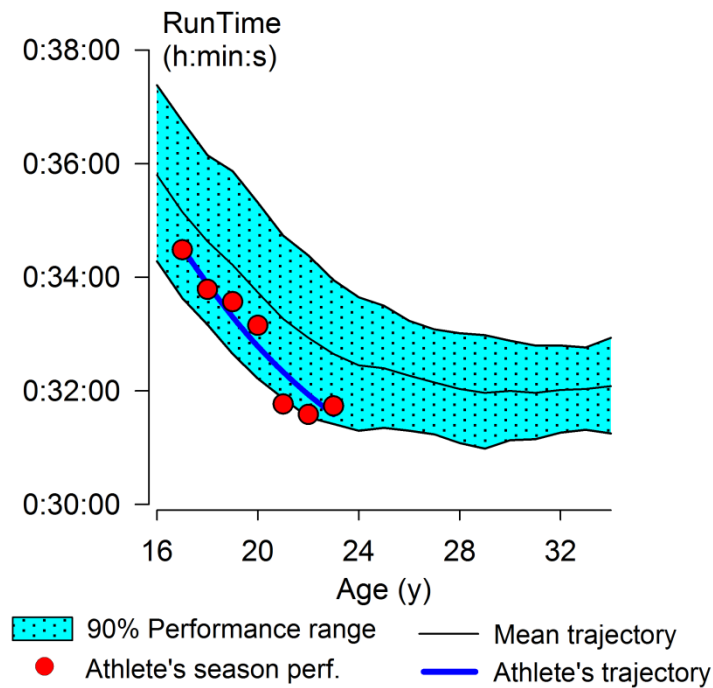


Figure 5. Comparison of an athlete's season performance and his trajectory against the typical age-related changes in running performance (mean and 90% performance range, see methods).

The probabilities (chances) of winning the 2012 London Olympic Games were plotted against the observed race time at the Olympics (Figure 6). Correlations of Olympic placings with predictions using trajectories and with placings at the last race before the Olympics are shown in Figure 7. The mean correlations between Olympic placings and placings in each of the 2012 World Triathlon Series races before the Olympics are also shown. Predictions using trajectories produced the highest correlation for men, but for women there was little difference between the three approaches.

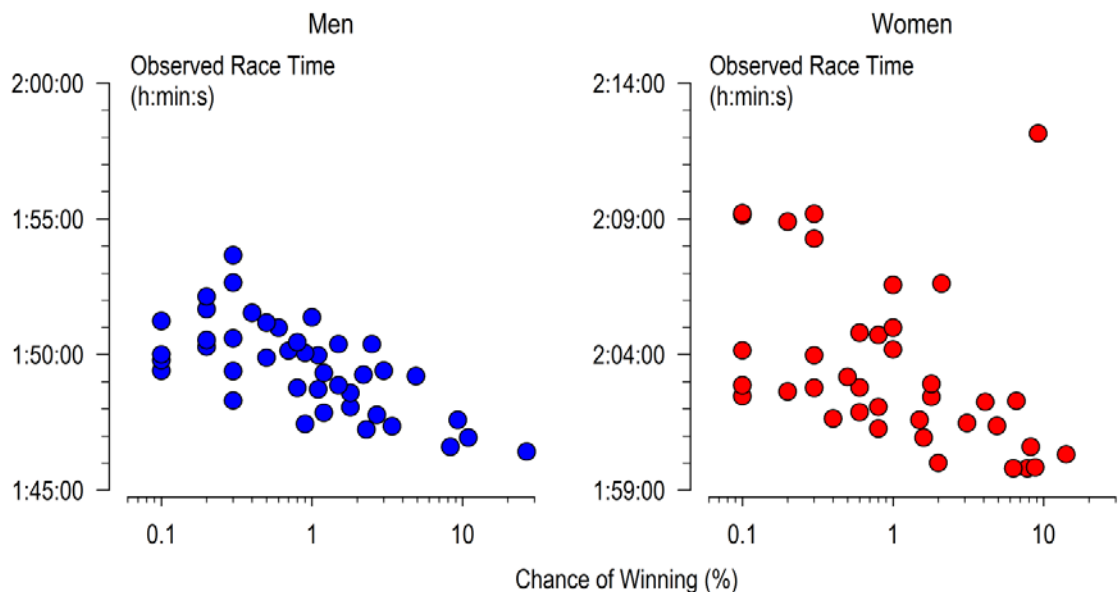


Figure 6. Relationship between observed performances in the 2012 London Olympics and predicted chances of winning. Chances of winning were predicted using athletes' career trajectories derived including races up to the Olympics and assuming similar race conditions (environmental and race-course profile) as 2011 London test event.

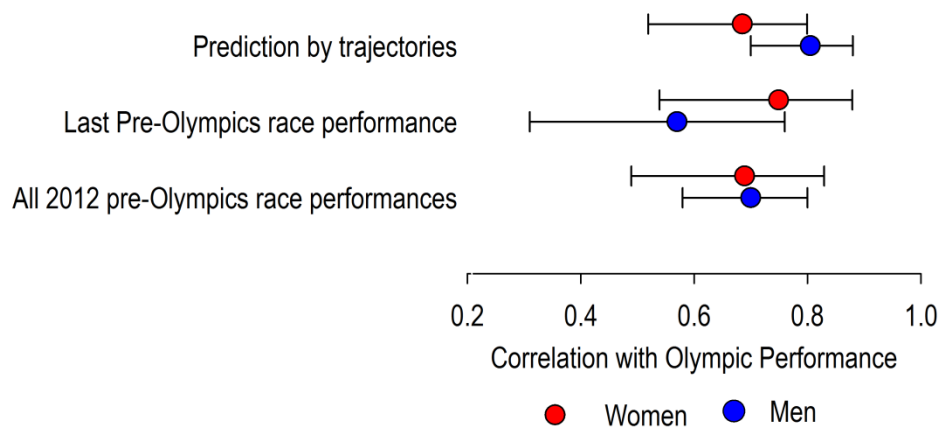


Figure 7. Correlations (with 90% confidence limits) of observed Olympic placings with predictions using trajectories and with placings at the last race before the Olympics. The mean correlation between Olympic placings and placings in each of the 2012 World Triathlon Series races before the Olympics is also shown.

3.4. Discussion

In this study we developed individual career trajectories of elite triathletes, investigating performance changes for the swim, cycle and run race stages and total times. Performance change was modelled as a quadratic function of age and a linear function of calendar year in a mixed model that accounted for difference in mean race times arising from environmental and other course-related factors. By markedly reducing the scatter of race times, the model produced well-fitting individual quadratic trajectories that were suitable for assessment of athletes' performance changes during their competitive career.

Analysis of the residuals from the mixed model provided good evidence that the underlying quadratic function of age with clusters for race and athlete was appropriate for tracking athletes' performance changes over an age span of 15 to 41 years. The other models used previously to describe individual performance changes in other sports would not have worked well with triathlon data. Alam et al. (Alam, et al., 2008) modelled the performance of boys and girls in cross-country skiing as a sigmoid function of age to describe effects of puberty. Their analysis was limited to an age span of 10 to 18 years and accounted for environmental factors by standardizing race times using the median of each race, which did not account properly for repeated measurement and athletes' abilities. For track-and-field athletics and swimming, Berthelot et al. (Berthelot, et al., 2012) modelled athletes' annual best performances as a double exponential function of age by assuming performances from junior through elite ages improved exponentially to a plateau and then declined exponentially from elite through master ages. This complex non-linear model was fitted to each athlete separately, so there was no correction for environmental and other race-related factors. Our model used race clusters to allow adjustment to an overall mean race time, thereby accounting for effect of environmental conditions and other race-course related factors that introduce extra variation in athletes' performances. The reduction of the scatter between

observed and corrected performances is evident in Figure 3. Furthermore, our model included clusterings for athletes and athletes within year. The repeated measurements for athletes produced individual career trajectories, highlighting athletes' differences resulting from different physiology, training history and nutrition regimes. The repeated measurement for athletes within a year was included to identify consistent changes in performance arising for short-term (one-year) changes in training programmes, nutrition strategies, injuries, and so on.

Career trajectories allow a direct and visually clear evaluation of athletes' performance changes (Figure 4). For top men and women, performance changes for total time are closer to those for running than those for swimming and cycling. Running is also the triathlon stage where individual trajectories have the widest spread in time, indicating that differences among athletes in total performance arise mostly from differences in athletes' running performance. These findings are not unexpected, as running performance is the most important stage for success in triathlon (Fröhlich, Klein, Pieter, Emrich, & Gießing, 2008; Knechtle & Kohler, 2009; Vleck, Bürgi, & Bently, 2006).

Our career trajectories are consistent with the findings of Landers et al. (2000), who analysed senior performance times in the 1997 World Championships and found the smallest percent difference for swim times (2.2% for men and 4.4% for women) and the biggest percent difference for run times (5.0% for men and 5.9% for women). There are several explanations for the differences between trajectories for swimming, cycling and running. Performance improvements within the athlete can result from physical and physiological maturation, training adaptation, improvement in skills and biomechanics, and increase in knowledge of race tactics (Mikulic, 2011; Sowell & Mounts Jr, 2005; Stevenson, et al., 2012). Differences in the development of physiology, biomechanics and skill may explain the fact that running had the largest 5-year improvement (Table 6), in particular the improvement in economy arising from the increase in muscle and tendon stiffness with age and training (Dumke, Pfaffenroth, McBride, & McCauley, 2010; Saunders, Pyne, Telford, & Hawley, 2004). Triathletes are also likely to focus their training more on running (Bürgi, 2013), which would make improvement in aerobic power more evident in this stage. Furthermore, during the swim and cycle stages of a race, athletes must be strategic: they need a good position for the subsequent stage, but by drafting they can reduce energy costs up to 30% (Hauswirth & Brisswalter, 2008) and thereby save energy for the running stage. Therefore age-related changes in athletes' endurance abilities may not be manifested fully in the swimming and cycling trajectories. It should also be noted that the majority of men's cycling trajectories show an unexpected increase in duration of the cycling stage with age. Two possible explanations for this effect are a gradual increase in difficulty of more recent cycling courses compared with those in the early 2000s (mean cycling times for the top 16 athletes are increasing with calendar year, Chapter 4) and a potential race strategy, where athletes may not fully manifest their cycling ability in order to not compromise their running performances. The effect is not evident in women's cycling trajectories, presumably because the greater performance improvement for women (shown across all the stages, Table 6) offsets the slowing of cycling times. Furthermore, pack riding must reduce the effect of age on cycling performance time, because the time of an individual athlete while riding in a pack will reflect the

average of the performances of that pack. If there are smaller packs with women (Vleck, Bentley, Millet, & Bürgi, 2008), there will be a bigger effect of age on the women's trajectories.

Differences in performance changes for total time between top men and women are also evident in the results: men show a smaller spread of performances and trajectories of a similar shape (Figure 2). This difference probably reflects a different developmental phase of the sport between the genders. Men's uniformity presumably indicates a group of athletes who have been training since early ages as triathletes, who show similar levels of abilities for swimming and cycling and for whom running is the most important stage for total performance. On the other hand, the heterogeneity of women's trajectories for total time and their greater performance improvements reflect less depth in women's competitiveness. Similar phenomenon has occurred in other triathlon modalities (Lepers, Knechtle, & Stapley, 2013), where bigger differences between winner and tenth-placed competitors have been reported for women compared with those for men (although women's differences have been decreasing faster than men's). The smaller depth of women's competitiveness results in a wider range in abilities, training and physiology among athletes and therefore a wider variety of "ways" for women to succeed in triathlon.

Our study is not the first to address age-related changes in triathlon performance. Previous researchers have used cross-sectional studies with age-group athletes. Performances were deemed similar between ages 20-35 and significant performance declines were reported after ages 40 to 55, depending on stage and triathlon distance (Bernard, et al., 2010; Etter, et al., 2013; Lepers, et al., 2013; Lepers, et al., 2009). By developing a quadratic model to track performance progression of individuals within the period of their elite career, we have revealed performance differences across ages (Figure 2) that were not evident in these cross-sectional studies.

Estimates of age of peak performance in triathlon were approximately 26-28 years, with trivial to moderate differences between genders and across stages. Our estimates of age of peak performance align well with age of athletes' best performance at a World Championships or Olympic Games. They were also consistent with previous findings for age of the best triathletes in an Olympic distance race: 27 ± 6 y for men 28 ± 6 y for women (Etter, et al., 2013). In longer distance triathlon, age of best Ironman triathletes was 33-34 y (Gallmann, Knechtle, Rüst, Rosemann, & Lepers, 2013; Rüst, Knechtle, Knechtle, Rosemann, & Lepers, 2012a). Physiological characteristics are similar for Olympic and long-distance triathletes (Grégoire P Millet, Dréano, & Bentley, 2003), so the difference between ages of peak performance are likely to be due to the fact that many triathletes compete in Ironman after retiring from Olympic competition, and that longer events may require more years of race experience.

Analyses of performance trajectories provide evidence of the typical pattern of progression of successful triathletes, which should assist with the setting of benchmarks for talent identification and development programmes. The analysis of athletes' trajectories and their season performances ought also to help detect successful (and unsuccessful) performance improvement strategies. In Figure 5, we gave an example of these applications. The athlete shown is progressing within the successful range, and at age 21 he improved his running performance substantially. A subsequent analysis of this athlete's training history may reveal

whether a new training approach, coach, nutrition strategy and so on contributed to the performance improvement. Furthermore, the comparison of swim, cycle and run trajectories of the same athlete against benchmarks will highlight the athlete's strongest and weakest stages, providing additional guidance for the athlete's career.

Career trajectories can also be used to predict performance for men. We found strong associations between observed and predicted performances at the 2012 London Olympic Games, although these associations were not much higher than those obtained with the simpler approach of correlating typical race placings at previous competition(s) with Olympic placings. Additionally, all the correlations were higher for the men than for the women. Analysis of the residuals over the 13-year period for the athletes competing at London Olympics showed that men and women have similar race-to-race variability, but male triathletes showed larger differences between competitors compared with that for women. The relationship of difference between athletes and within-athlete variability lead to the greater intraclass correlation coefficient (within year) for men. The particular lower predictability of women performance using trajectories can also be presumably explained by the wider diversity of trajectories shown by women (Figure 4), leading to a poorer fitting of trajectories for women.

Practical applications for tracking triathletes' performances with our method could be developed further. First, tables showing annual percent improvements for each year of age, as means and standard deviations, will make expected performance goals clear and easy to communicate to coaches and athletes. Secondly, career trajectories were developed giving equal importance to each performance; however predictions for future performance might be more accurate if more weight is assigned to more recent performances. Thirdly, performance predictions may also be improved by including random effects to specify individual responses to environmental and other course-related factors (e.g., temperature, wetsuit swim, race-course profiles) and importance of the race (e.g., Olympic Games, World Triathlon Series, World Cup races). Fourthly, data were limited to international races performed at a professional level performances at younger ages in lower level competitions should be included in the analysis to make the method more useful for talent identification. For this purpose data of substantial numbers of the same athletes competing at international and at lower level competitions would be needed. Finally, an athlete's season performance appearing as an unexpectedly large deviation from the quadratic trajectory could provide evidence of the use of a banned performance-enhancing substance.

3.5. Conclusion

We have presented a new method for tracking the development of elite triathletes performing internationally. The resulting career trajectories and reference ranges represent objective measures of performance that should provide useful information for funding talented athletes and identifying successful (or unsuccessful) performance enhancement strategies. Furthermore, the comparison of an athlete's three trajectories (swim, cycle and run) with corresponding reference ranges could also help apportion appropriate training focus to the

specific race stages. The full use of the method presented here to address relative strength and weakness of a given athlete in the different race stages will need to consider the contribution of each stage to total time. Trajectories also offer a small advantage over usual race placings for predicting men's performance, but further refinements of the model, by including athletes' individual responses to race conditions, may allow more accurate projection of triathletes' trajectories into the future.

CHAPTER 4

PERFORMANCE PROGRESSION IN ELITE OLYMPIC-DISTANCE TRIATHLON

This chapter comprises the following paper submitted to Scandinavian Journal of Medicine and Science in Sports:

Malcata, R. M., Hopkins, W. G., & Pearson, S. Performance progression in elite Olympic-distance triathlon. *Scandinavian Journal of Medicine and Science in Sports*.

Overview

Determining year-to-year performance trends is problematic in sports with large environmental effects. **Purpose:** To determine progressions in elite Olympic-distance triathlon times and trends in contribution of swim, cycle and run race stages to overall performances. **Methods:** Performance times of athletes competing in World Championships, World Triathlon Series, Olympics and World Cup races from 2000 through 2012 were obtained from triathlon.org. The top 16 performances in each race were analysed using a mixed linear model to derive year trends by adjusting for repeated measurements of athletes, race identity and race type. Contribution of each stage was assessed by correlating overall placing's with performance times in the individual stage. **Results:** Typical race-to-race differences, which reflect variation arising from environmental and other course-related factors, were ~4% for overall times. From 2000 to 2012 changes in performance showed different year trends for race stages and genders, with running showing the largest improvement and women progressing faster than men. Running made the highest contribution to overall performance and there was an apparent gradual increase in this contribution over the years. **Conclusion:** Our data suggests similarity between athletes' abilities during swim and cycle and that differentiation among top triathletes is made mainly during running.

4.1. Introduction

Sport performance continues evolving as result of innovations with training, nutrition and equipment. It is important to monitor performance to provide coaches and support staff with an evidence based approach for guiding athletes' preparation. In many sports, effects of external factors, such as temperature, wind and course profiles, produce extra variation in performance that needs to be taken into account when assessing performance. Triathlon, which consists of a sequence of a swimming, cycling and running, is one such sport.

Performance progression in triathlon has been studied mainly by Lepers and colleagues. Participation rates, progression of mean top times and gender differences were investigated for the various triathlon modalities: Olympic (Bernard, et al., 2010; Etter, et al., 2013; Lepers, et al., 2009), Ironman (Baker & Tang, 2010; Lepers, 2008; Lepers & Maffiuletti, 2011; Rüst, Knechtle, et al., 2012a; Sowell & Mounts Jr, 2005; Stiefel, Knechtle, & Lepers, 2012), off-road (Lepers & Stapley, 2010, 2011) and ultra-endurance triathlons (Knechtle, et al., 2010; Lepers, Knechtle, Knechtle, & Rosemann, 2011; Rüst, Knechtle, Knechtle, Rosemann, & Lepers, 2012b). In previous studies, performance progression has been assessed using mean time of top athletes at the same race over multiple years. Although this approach reduces effects of course profile on performance times, other factors can have substantial effect of the estimates of progression. Depending on the study, authors have not accounted for repeated measurements from the same athlete competing over multiple years, calibre of athletes competing, and difference in environmental conditions. Mixed modelling provides a statistical approach to overcome such limitations. The first aim of this study was therefore to determine year trends for the swimming, cycling, running and overall performance times in elite Olympic triathlon using mixed modelling.

A unique aspect of performance progression in triathlon is the contribution of swim, cycle and run stages to overall performance. Although researchers are agreed on the importance of running phase, there has been more debate on the importance of swimming and cycling (Fröhlich, et al., 2008; Landers, et al., 2000; G.P. Millet & Bentley, 2004; Vleck, et al., 2006). Furthermore, no one has examined how the contributions of the three stages have changed over the years. A second aim of this study was therefore to address the contributions of the three stages over the years, and examine potential differences between the three development groups: junior, under-23 and elite athletes.

4.2. Methods

Data

Official times of international triathlon races from World Triathlon Series (titled World Championship Series between 2009-2011), World Cup, World Championships (Junior, Under-23 and Elite) and Olympic Games between 2000 and 2012 were downloaded from triathlon.org, together with race dates. Only top-16 performances from each race were included. Overall, 358 international competitions (180 men and 178 women) were raced over this period, with a total of 362 individual male and 342 individual female athletes who ever finished within the top 16 of a

race. Olympic Games and World Championships races for elite athletes were categorized as key races.

Year Trends

Year trends were generated using mixed linear model procedure in the Statistical Analysis System (Version 9.2, SAS Institute, Cary, NC), (see APPENDIX A). The fixed effect model included an intercept and a linear trend for calendar year. Random effects included adjustment for athlete's identity, accounting for repeated measurements of the same athlete over multiple races; annual adjustment for athlete, accounting for an annual consistent variation of athlete's performances (consequence of a new training programme, different nutrition strategy, injury); and adjustment for race identity, the clustering of all performance in each race accounts for the effect of environmental and other course-related factors on performances. The residual random effect representing the within athlete race-to-race variability was specified according to the level of competition, assuming two different levels: *World Triathlon Series*, *World Championships* and *Olympic Games*; and *World Cup* (often a weaker field of competition and potentially more variable). Race times were log-transformed to yield the effects and errors as percentage changes of the mean. Observations were considered outliers and excluded from the analysis if residuals were greater than four standard deviations (15 swimming, 28 cycling, 17 running and 20 overall performances were excluded). The appropriateness of the model was investigated by analysis of distribution of all random effects to ensure no skewed distributions, and analysis of residuals: plots for residuals vs predicted performances were inspected to ensure there was no unacceptable non-uniformity, and residuals were plotted against year to ensure no substantial systematic trend in the residuals. The mixed linear model was deemed appropriate to represent mean performance changes with calendar year.

Year trends were estimated for each stage (swim, cycle, run and overall performance) and each gender separately. The same model was also applied using performances at key races (Olympics Games and World Championships) only, to estimate performance progression at main events. Future mean performance times for men and women, in each stage at all races and at key races only were predicted by extrapolating year trends up to four years out (coincidentally with 2016, year of the Rio Olympic Games). Magnitudes of changes in mean performance times were assessed using a modified scale for standardized differences in means (W.G. Hopkins, et al., 2009): thresholds for small, moderate, large, very large and extremely large were 0.2, 0.6, 1.2, 2.0 and 4.0 of the typical race-to-race variation at key races for each stage and gender. The usual thresholds based on the race-to-race variability of individual athletes' performance times are not appropriate for assessing trends in mean race times, because those thresholds were defined to assess change of individual athletes as opposed to assess changes in mean. Typical race-to-race variations were determined as the standard deviation of the random effect for race identity. Uncertainty on the estimates of performance change and predicted mean times were derived as 90% confidence limits.

Performances of athletes were categorized according to placing in each race as *Top 3*, for athletes finishing first to third; *Top 8*, for athletes finishing between fourth to eighth; and *Top 16*, for athletes finishing between ninth and 16th. Year trends were then calculated for each stage and gender for these three groups. Trivial differences were found between year progressions of these three groups; we then decided to combine all data and present only the analysis using all the top 16 performances in each race.

Gender Differences

Gender differences were calculated for swim, cycle, run and overall performance as the difference between mean performance times for men and women, with means estimated by the year progression model (as described previously). Gender differences were calculated using a meta-analytic model, where the difference between annual mean performance estimates for men and women was weighted by the inverse of the square of the standard error of the estimates.

Contribution of each race stage

Relative contribution of swim, cycle and run was evaluated by calculating correlation between log-transformed overall placings and performance rank in each stage. The log-transformation of placings was used to give equal importance to percent or factor differences in placings rather than absolute differences (as discussed in Chapter 3). For example, the difference between second and first is equivalent to the difference between tenth and fifth with log transformation and to the difference between tenth and ninth without transformation. Correlations between each stage and overall performance times were not performed as results would be biased by the proportional time spent in each of the stage. Annual correlations were derived using a Fisher-Z transformation to each race correlation and were then plotted against calendar year. A linear calendar year trend was determined for the Fisher-Z transformed correlations and after back transformation, the relationship between correlation and performance was non-linear. To compare the development groups (*juniors*, *under-23* and *elite*), we calculated similar correlations using performance at World Championships races only. These data were chosen as World Championships are disputed in the same venue, over the same weekend, with junior athletes competing in the sprint distance. It should be noted that in 2000 and 2001, junior races were raced as Olympic distance and there were no Under-23 competitions. Correlations were assessed using 0.1, 0.3, 0.5, 0.7, 0.9 as thresholds for small, moderate, large, very large and extremely large (W.G. Hopkins, et al., 2009), and difference in the year trends analysed accordingly to the overlap of the 90% confidence limits of each trend.

4.3. Results

Table 7 presents race times (mean \pm standard deviation) and typical differences between athletes' performances within a race for the international races from 2000 through 2012. When analysed as percentage of mean stage times, running was the stage with the largest differences between athletes, whereas swimming and cycling stages had the highest race-to-race variation.

Typical differences between athletes within a race and typical differences between mean race times derived by mixed modelling were similar to these raw values (data not shown).

The analysis of key races only (World Championships and Olympic Games) produced the following race-to-race variability for swim, cycle, run and overall: 3.5%, 8.5%, 2.8%, 4.6% for men and 4.9%, 8.4%, 1.8%, 4.5% for women. These standard-deviations were used to set thresholds for assessing important performance changes (see Methods).

Table 7. Mean race time (min) and typical difference (min) between athletes within a race for top 16 athletes in every World Cup, World Triathlon Series, World Championships and Olympic Games race from 2000 through 2012. In total, data consisted of 362 individual male and 342 female athletes competing in 180 and 178 races, respectively.

	Swim	Cycle	Run	Overall
Men				
Mean race time (mean \pm SD)	18.3 \pm 1.2	59.9 \pm 4.1	32.2 \pm 1.3	111.2 \pm 4.7
Within-race SD (mean)	\pm 0.3	\pm 0.5	\pm 0.8	\pm 0.8
Women				
Mean race time (mean \pm SD)	19.8 \pm 1.5	66.1 \pm 4.4	36.9 \pm 1.9	123.8 \pm 5.4
Within-race SD (mean)	\pm 0.5	\pm 0.7	\pm 1.2	\pm 1.6

Sprint-distance races were not included in the calculation of times.

Figure 8 illustrates raw means and modelled year trends of swim, cycle, run and overall performance times for men and women at key races between 2000 and 2012. Changes in mean performance times were evaluated over four years, with a negative change representing an improvement (decrease in performance time). The mean change (\pm 90% confidence limits) in performance at key races for swimming, cycling, running and overall were: -1.7% (\pm 1.3%), 0.6% (\pm 2.9%), -1.8% (\pm 1.1%) and 0.0% (\pm 1.7%) for men and -0.6% (\pm 2.0%), 0.4% (\pm 3.1%), -2.1% (\pm 1.1%) and -0.4% (\pm 1.8%) for women per four years. Performance changes (and confidence limits) were proportional larger for longer periods of time. The majority of performance progressions for all races were similar to those for key races, with differences being observed for men's swimming and both men's and women's cycling (data not shown). In men's swimming, the clear improvement (reduction) in times during key races contrasted with the potential increase of 0.3% (\pm 1.0%) of mean swimming times for all races, while in cycling, the unclear changes in performance at key races contrasted with a decline in mean performance of 1.4% (\pm 0.9%) for men and 0.8% (\pm 0.9%) for women per four years when taking into account all races.

Predictions of mean times for 2016 in key races are presented in Table 8. Predictions were determined by extrapolating year performance trends towards the Olympic year of 2016.

Table 8. Predicted mean top-16 times and 90% confidence limits (min) for swim, cycle, run and overall in 2016. Estimates were predicted by extrapolating year progressions towards 2016, using data from key races (World Championships and Olympic Games) between 2000 and 2012 shown in Figure 8.

	Predicted time in 2016 (min)			
	Swim	Cycle	Run	Overall
Men	17.4; \pm 0.6	60.4; \pm 4.6	30.3; \pm 0.9	109.9; \pm 4.9
Women	19.3; \pm 1.0	65.5; \pm 5.1	34.3; \pm 0.9	120.6; \pm 5.5

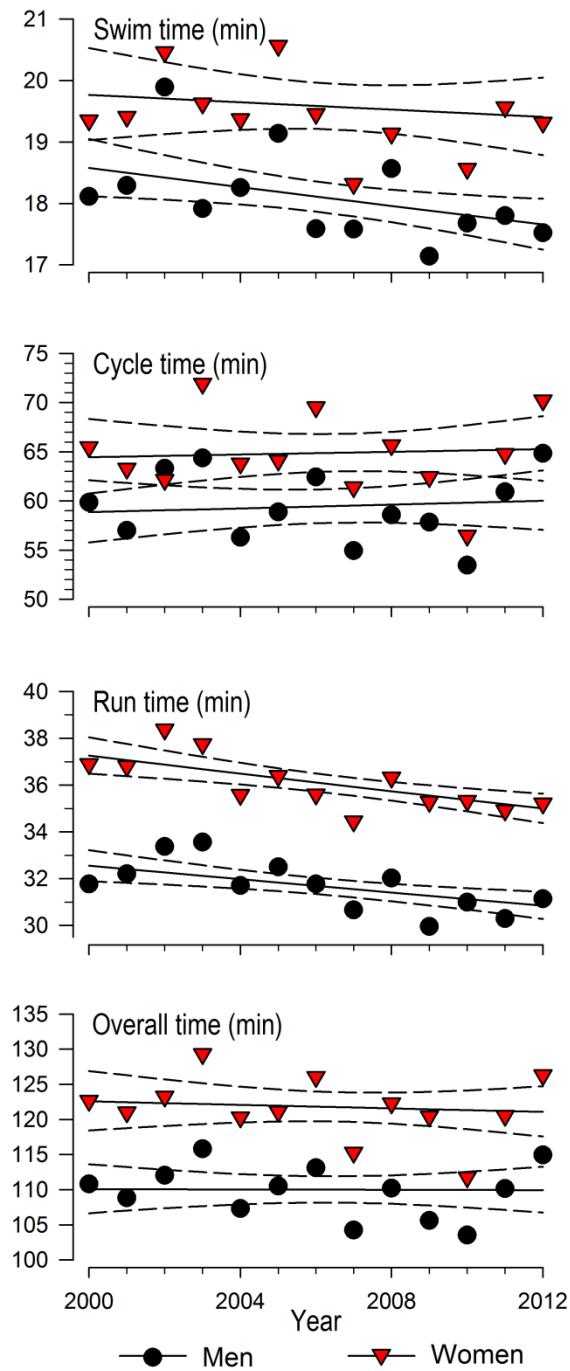


Figure 8. Progression of mean swimming, cycling, running and overall performance times for the top 16 male and female athletes in every key race (World Championships and Olympic Games) between 2000 and 2012. Estimated linear year trends and 90% confidence limits are presented as bold and dashed lines, respectively.

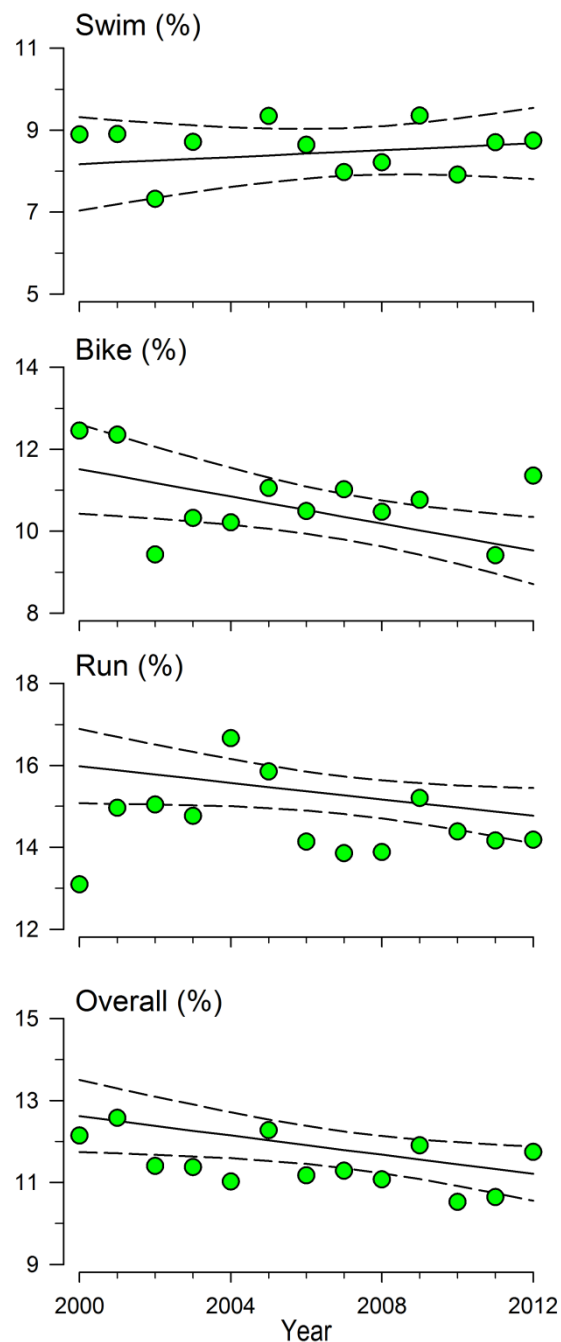


Figure 9. Differences between mean times for the top-16 men and top-16 women in all races (World Cup, World Triathlon Series, World Championships and Olympic Games races; 180 men's and 178 women's races) between 2000 and 2012 for swimming, cycling, running and overall performance. Estimated linear year trends and 90% confidence limits are presented as bold and dashed lines, respectively.

Gender differences are presented in Figure 9. Between 2000 and 2012 swimming was the stage with the smallest gender differences and was the only stage in which gender differences have increased a little. For cycling, running and overall performance gender differences have decreased, with cycling difference changing with the fastest rate.

In Figure 10, annual-mean correlations of performance times in each stage with overall placings for all elite races are presented as a function of calendar year. Correlations for running stage were the highest, with women showing in 2012 the same correlation as men. Although men and women showed a similar increase in running and decline on cycling correlations over the

years, men's correlations for swimming were stable while those for women declined. Figure 10 also shows correlations for performances at World Championships only, for elite and junior athletes. Correlations for the elite athletes at World Championships showed a pattern of progression comparable with those for all races. Differences in the progressions are evident between junior and elite men: cycling and running have become equally important for performance of juniors, while elites show the opposite trend. Women's correlation progressed similarly for the juniors and elites among the three development groups, with running on average having the highest contribution. The patterns of progression for the under-23 athletes were intermediate between those of juniors and elites (data not shown).

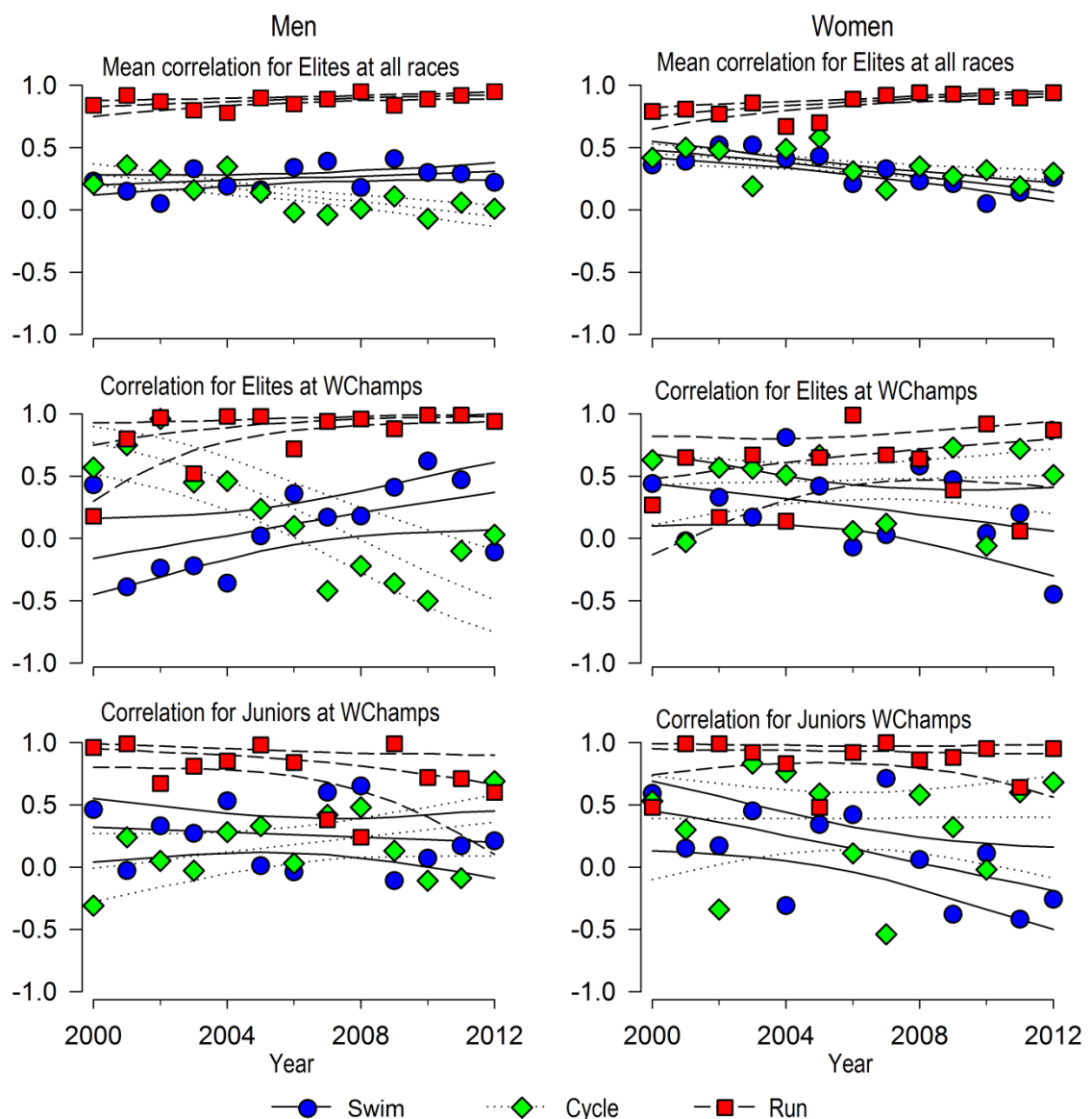


Figure 10. Correlations of performance times in each stage with overall placings in all races (World Cup, World Triathlon Series, World Championships and Olympic Games), in elite and in junior World Championships, for men and women between 2000 and 2012. Regression lines for each stage are shown.

4.4. Discussion

This study was the first where performance progressions for any triathlon modality were investigated accounting properly for repeated measurements for athletes and extra variation arising from environmental and other course-related factors. Although there was little change in overall performance time in Olympic-distance triathlon races between 2000 and 2012, there were different progressions for swimming, cycling and running stages. Effects of individual stages on overall performance times have also changed, reflecting an increasing importance for running in triathlon.

A novel aspect of this study was the use of mixed modelling to properly account for the repeated measurements and estimate the effect of external sources of variation. The analysis of residuals and distributions of random effects provided evidence that the underlying model was appropriate to estimate performance progressions. Our model included clusters for athletes, which accounted for individual differences arising from differences in training and competition experience, and clusters for races, which allowed for consistent deviations of race-performance times from the year trends arising from environmental and other course-related factors. These consistent deviations reflect the impact of environmental and other course-related factors in the mean performance time in a race. Windy or wet conditions can result in slower times across all three stages, whereas other factors can affect mainly one stage. For example, temperature determines whether or not it is a wetsuit swim, which can lead to faster times (Chatard, et al., 1995). Race courses also contribute to the variation of performance times. Swim courses vary in geometry (buoy distance and configuration), in water current, and whether the water is fresh or salty; cycle and run courses vary in elevation and number of sharp curves, and even the distance of the cycle stage can vary up to 10 % of the standard 40 km (ITU, 2013). Cycling was the race stage with the highest difference in mean performance times between races (after accounting for differences in mean performance times arising for the different abilities of the athletes competing in each race).

We would expect gradual improvement in performance over the years owing to innovations in training methods, nutrition strategies and technology. Such improvements were clear but small for men's swimming at key races and moderate for men's and women's running at key races and all races. The unexpected finding was the possibly small slowing down of mean cycling performances for all races. This apparent decline in cycling performance may consequence of two factors: a consistent bias towards more technical cycle courses, with more elevations and sharp curves, justifying why triathletes are taking longer times to complete the cycle stage, and a more important role of race tactics (drafting), where athletes are being more strategic to guarantee the best position for a run performance. Cycling performance changes for key races were unclear presumably because of high race-to-race variability in cycling. Studies in performance progression in Ironman (Lepers, 2008), off-road (Lepers & Stapley, 2010) and ultra-endurance triathlons (Knechtle, et al., 2010) have found plateau in overall performance time similar to those observed here. In these other triathlon modalities, different progressions for swimming, cycling and running performances were also observed, although they are not all comparable to ours. In the only other study of performance progression in Olympic-distance

triathlon, Etter et al. (2013) found progression at the Zürich Triathlon for overall and swim times comparable to ours but trends for cycling and running opposite to ours. This race is a non-drafting race and it was not part of the world triathlon circuits, so their findings may not reflect progression of elite performance.

By extrapolating performance progressions, we have predicted mean performances in key races for the 2016 Olympic year. With the possible exception for running, the uncertainty in these predictions are too large for using predicted times as futures targets. The uncertainty arises from differences environmental and course-related factors from race to race, so it should be possible to reduce the uncertainty of the estimates by using a model that estimates the specific effects of such factors (e.g., temperature, wetsuit swim, race-course profiles).

Although men and women have been exposed to the same innovations in training methods and nutrition (Lepers, 2008), the gender difference for overall performance has been decreasing. This decrease may be explained by an increase of female participation and consequent improvement of female competition depth, similar to that in other triathlon races (Knechtle, et al., 2010; Lepers, et al., 2013; Stiefel, et al., 2012). For comparison between stages, differences in mean performance times between genders were smaller for swimming than for cycling and running. Similar findings were reported in studies of Ironman and off-road triathlon (Lepers, 2008; Lepers, et al., 2013; Lepers & Stapley, 2010). These authors explained gender differences by men's greater muscle mass and higher aerobic capacity but offset to some extent in swimming by women's anthropometry and swimming technique. In our study, cycling is the stage where gender differences have been decreasing at the fastest rate, potentially as result of the recent bias towards more technical cycling courses which would make athletes' performances more reliant on skill than on physiology. Running, as in previous studies (Etter, et al., 2013; Lepers, 2008; Lepers, et al., 2013; Lepers & Stapley, 2010), was the stage with the largest gender difference comparable to that for Ironman (13%), Olympic-distance triathlon (17%) and off-road (18%).

We have investigated the importance of each stage by correlating performance ranking in swimming, cycling and running stages with overall placing. The highest contribution of running to overall performance and the apparent increase in the contribution over the years emphasise the importance of running ability for success in triathlon. The highest correlation for running was previously reported by Landers et al. (2008). The comparative lower values found for the swimming relative to the correlation for cycling can be explained by the fact that such correlations were calculated with performance times, and therefore correlations were biased toward the cycle (stage with the largest percent of total time). In addition, the lower contribution of swimming and cycling to overall performance and the smaller differences between athletes in these stages reflect less heterogeneity among swimming and cycling abilities. Surprisingly, cycling and running were equally important for junior male athletes, a finding that may be explained by the wider range of cycling and running abilities among these athletes (as reported by Landers et al. (2000); although Junior and elite athletes raced in distinct conditions). Furthermore, the apparent decline of running and increase of cycling contributions to overall performance for these junior athletes

may reflect a race strategy different from that of the elites; presumably junior athletes have a greater tendency for successful cycling breakaways.

4.5. Conclusions

Between 2000 and 2012 running performance of elite Olympic-distance triathletes became faster, and the contribution of running to overall performance increased. Over these years, little changes were observed for swimming, while cycling showed greatest effects of environmental and course-related factors. Contributions of swimming and cycling to overall performance have been generally negligible, indicating that athletes gain no real advantage in swim or cycle stages and that the race is won in the run. On the other hand all the three stages have become important for the male junior triathletes. Further studies are required for a better understanding of race strategies and differences between junior and elite athletes to develop the talent of the future successful elite triathletes

CHAPTER 5

MODELLING PROGRESSION OF COMPETITIVE PERFORMANCE OF AN ACADEMY'S SOCCER TEAMS

This chapter comprises the article published at Journal of Sports Science and Medicine:

Malcata, R. M., Hopkins, W. G., & Richardson, S. (2012). Modelling the progression of competitive performance of an academy's soccer teams. *Journal of Sports Science and Medicine*, 11, 533-536.

Overview

Progression of a team's performance is a key issue in competitive sport, but there appears to have been no published research on team progression for periods longer than a season. **Purpose:** Report the game-score progression of three teams of a youth talent-development academy over five seasons using a novel analytic approach based on generalised mixed modelling. **Methods:** The teams consisted of players born in 1991, 1992 and 1993; they played totals of 115, 107 and 122 games in Asia and Europe between 2005 and 2010 against teams differing in age by up to 3 years. Game scores predicted by the mixed model were assumed to have an over-dispersed Poisson distribution. The fixed effects in the model estimated an annual linear progression for Aspire and for the other teams (grouped as a single opponent) with adjustment for home-ground advantage and for a linear effect of age difference between competing teams. A random effect allowed for different mean scores for Aspire and opposition teams. All effects were estimated as factors via log-transformation and presented as percent differences in scores. Inferences were based on the span of 90% confidence intervals in relation to thresholds for small factor effects of $\times/\div 1.10$ (+10%/-9%). **Results:** Most effects were clear only when data for the three teams were combined. Older teams showed a small 27% increase in goals scored per year of age difference (90% confidence interval 13 to 42%). Aspire experienced a small home-ground advantage of 16% (-5 to 41%), whereas opposition teams experienced 31% (7 to 60%) on their own ground. After adjustment for these effects, the Aspire teams scored on average 1.5 goals per match, with little change in the five years of their existence, whereas their opponents' scores fell from 1.4 in their first year to 1.0 in their last. The difference in progression was trivial over one year (7%, -4 to 20%), small over two years (15%, -8 to 44%), but unclear over >2 years. **Conclusion:** The generalized mixed model has marginal utility for estimating progression of soccer scores, owing to the uncertainty arising from low game scores. The estimates are likely to be more precise and useful in sports with higher game scores.

5.1. Introduction

“Has your team improved?” is an important question for coaches and support staff that needs to be addressed with appropriate measures of performance in competitions. Match analysis can provide measures of various aspects of performance, but the game score itself is the criterion for assessing overall progression. Surprisingly, there has been no published research using game scores to track progression of team performances over periods longer than a year. In previous studies of association football (soccer), game scores have been analysed mainly to predict individual game outcomes and probability of a team winning a national league (Karlis & Ntzoufras, 2003, 2009; Lee, 1997; Maher, 1982; Rue & Salvesen, 2000) or a knock-out tournament (Dyte & Clarke, 2000; Koning, Koolhaas, Renes, & Ridder, 2003). In these analyses game scores were modelled assuming a distribution appropriate for count data, the Poisson or over-dispersed Poisson distribution. Important predictors included in previous models were parameters describing relative quality of teams. In national leagues, where all teams play each other the same number of times, the parameters described each team’s attacking and defensive ability (Karlis & Ntzoufras, 2003; Lee, 1997; Maher, 1982; Rue & Salvesen, 2000). For analyses of tournaments at World Cups, differences in teams’ abilities were addressed using the FIFA ranking system (Dyte & Clarke, 2000). All previous models included a game location effect addressing whether a team was playing at home or away.

The models used in previous studies cannot be applied directly to develop the performance progression of soccer teams of youth talent-development academies, for the following reasons. First, progression implies tracking the performance in different years, therefore a time variable is required in the analyses. Secondly, quality of competitors cannot be addressed using attacking/defensive parameters or FIFA world rankings, which are derived from series of games between most or all possible pairings of teams. Finally, the models need to include an effect for age difference between playing teams, which at an academy level is likely to impact performance.

In the present study we have applied a generalised mixed linear model to game scores with effects accounting for an annual trend of performance, quality of teams, age of competitors and home-ground advantage. We investigated the progression of three youth soccer teams from the Aspire Academy for Sport Excellence (Doha, Qatar) for the years 2005 to 2010, comparing their performance against that of their opponents.

5.2. Methods

Data

The data were official game scores of three Aspire teams and their respective opponents over the period 2005 to 2010. Informed consent was not required for approval by our institutional ethics committee, because game scores are in the public domain. The three Aspire cohorts consisted of players born in 1991, 1992 and 1993. Over the five years of their development programme, these cohorts played 115, 107 and 122 games scoring 163, 176 and 188 goals against 61, 56 and 60 different opponents, who scored totals of 173, 141 and 174 goals, respectively. Matches were contested in Asia and Europe, either as friendly games (when one

team is played at home and other away) or at small tournaments (both teams playing away). The age difference between Aspire team and their opponents was up to three years.

Year and season trends

The analysis presented an opportunity to trial the generalized linear mixed modelling procedure, *Proc Glimmix*, recently available in the Statistical Analysis System (Version 9.2, SAS Institute, Cary, NC), (see APPENDIX A). This procedure can model complex repeated-measures structures that cannot be accommodated than the established form of the generalized linear model known as generalized estimating equations; although these could have used with our data. The number of goals scored by each team was modelled as an over-dispersed Poisson distribution to allow for the variance of the counts to be different from the mean count (as justified in (Nevill, Atkinson, Hughes, & Cooper, 2001)). The fixed effects (and their estimates) were as follows: Team (with two levels, estimating a different mean score for Aspire and for the other teams grouped as Opposition), Team interacting with the playing season (allowing for a linear annual trend in performance for Aspire and Opposition), HomeAway interacted with Team (accounting for an advantage when Aspire or Opposition were playing at home), and a linear AgeDifference interacted with Team (reflecting the advantage per year of difference between the mean age of the teams, with a separate estimate for Aspire and Opposition). An annual linear trend in performance rather than quadratic or higher order trend was deemed the most appropriate, based on assessment of the annual mean scores. In the model the estimated mean goals were adjusted to a zero age difference and equal numbers of games played at home and away. The random effect Team interacting with identity of the team in opposition was included to account for opponents' different abilities and Aspire's ability against those opponents.

The analyses were performed individually for each Aspire cohort and for the three cohorts combined. In the combined analysis opposition teams with the same name in different years were treated as independent teams (i.e., not counted as repeated measurements).

Modelling was also investigated for team-performance progression within a season. Dates of each game were not available, but the temporal order was known and used as the time variable. Team performance within-season was predicted with similar Team, HomeAway and AgeDifference effects and an interaction between Team and game order to estimate different within-season rates of progression for Aspire and Opposition.

The effects were derived as ratios from the model but expressed as percentage difference. Magnitudes of effects were categorised in relation to the default thresholds for counts, with small, moderate and large factor effects of $\times/\div 1.10$, $\times/\div 1.40$ and $\times/\div 2.0$ (+10%/-9%, +40%/-29% and 100%/-50%) (W.G. Hopkins, 2010). An inference about the true (large-sample) value of the effect was based on uncertainty in its magnitude: if the 90% confidence interval overlapped small positive and negative values, the magnitude was deemed unclear; otherwise, the magnitude was deemed to be the observed magnitude (Batterham & Hopkins, 2006).

5.3. Results

Results are presented only for the analysis when games data from the three cohorts (Aspire teams born 1991, 1992 and 1993) were combined. The individual analysis for each cohort produced mainly unclear effects.

Age difference had similar effects for Aspire and Opposition, so they were combined into a single effect. A one-year difference between playing teams offered a small advantage of 27% more goals for the older one (confidence interval 13 to 42%). The age effect was modelled as a linear variable with the log of mean number of goals; consequently two- and three-year gaps resulted in moderate and large effects of 61% and 105% more goals scored by the older team.

The Aspire team experienced an advantage of 16% higher scores (-5 to 41%) when playing at home, whereas Opposition scores were higher by 31% (7 to 60%) when playing on their own ground, both home-ground effects were small. The difference between the two effects was unclear (13%, -15 to 49%).

Figure 11 shows the mean number of goals scored per season by the Aspire and Opposition teams over the five years (Season 04/05 through to 09/10). After adjusting for age-difference and home-ground effects, Aspire scored on average 1.5 goals per match, with no change over the five years. On the other hand, the Opposition's mean performance fell from 1.4 goals per match in their first season to 1.0 goals in the last. At the end of the first year (04/05) the difference between the two adjusted means was trivial (5%, -15% to 30%), whereas by the end of 09/10 Aspire scored moderately more goals than the Opposition (40%, 0 to 96%). The comparison of the performance progressions showed a trivial difference between the two teams over one year (7%, -4 to 20%), small over two years (15%, -8 to 44%) and unclear for three years (24%, -11 to 74%) and longer periods.

Within-season team progression was explored using data from the 15-31 games for each season of each of the three cohorts. When the model specifying home-ground and age-difference effects as predictors of mean number of goals was applied, the estimated ratios of progression of Aspire vs Opposition had on average an uncertainty of $\times/\div 4.0$. Thus, for observed differences to be clear, they would have to be at least very large. When a more simplistic model ignoring home-ground and age effects was applied, the uncertainty decreased to $\times/\div 2.7$, which still represents large uncertainty.

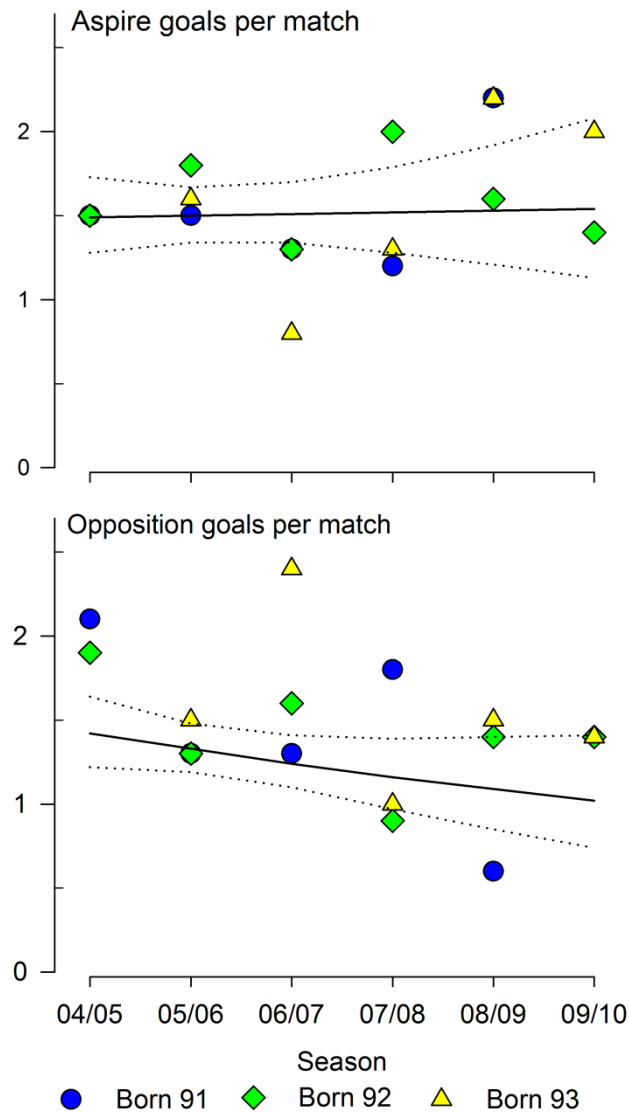


Figure 11. Raw mean number of goals scored in each season for the three Aspire cohort and their Opposition teams. Lines show adjusted means and 90% confident limits for the Aspire and Opposition teams, for the six seasons between 04/05 though to 09/10.

5.4. Discussion

We have investigated the five-year performance progression of three academy soccer-team cohorts using a novel application of generalised linear mixed modelling. The analysis revealed substantial effects on performance for an age difference between teams, for game location, and for differences in progression of the Aspire and Opposition teams. There were no clear outcomes for within-season performance progressions.

An age difference of one year between opposing teams resulted in a small advantage for the older team. This advantage is obviously due to differences in physical maturity, which is highly correlated with performance during puberty (Mujika et al., 2009). Even an age difference of less than a year produces the well-known relative age-effect in performance, which has been demonstrated in soccer (Helsen, Van Winckel, & Williams, 2005) amongst many other sports. The same authors suggest that advantage experienced by older players may also reflect

psychological maturity and longer exposure to practise and matches, resulting on the development of technical and game intelligence skills. Our estimate of the age effect is likely to be biased low, because games between teams differing in age are more likely to have been set up when the perceived abilities of opposing teams were similar.

The estimated small advantage for the team playing at home is consistent with previous studies, in which the home-ground factor represented approximately 40% higher number of goals for the hosting team (Koning, et al., 2003; Lee, 1997). The estimated home-ground effect in our study was a little lower, but differences between the two values may be due to the different nature of players (professional vs youth). The difference between home advantage experienced by the Aspire and Opposition teams was unclear; however, there was an indication of a greater home-ground effect for the opposition. If the true difference between the home advantages is substantial, possible reasons include different climate conditions and different fan support that players experienced in the Qatar venue vs the opposition venues.

Although the analysis for progression for each cohort involved ~100 games, the effects on progression were not clear until all three cohorts were included in the analysis—a sample size of ~300 games. The average performance of Aspire cohorts was fairly constant over the five-year period, while the opposition gradually scored less goals. The most obvious explanation for this outcome is an improvement of Aspire performance through development of their defensive ability. A reduction in the opposition's attacking ability seems a less likely explanation, but this issue could be resolved only by an analysis of scores from games where opposition teams play each other.

The assessment of the magnitude of effects in this study depends on the chosen thresholds. The threshold for small was the default 10% change in the score. However to be consistent with previous research on solo athletes, the threshold should be the smallest change that would increase by 10% the chance of winning against an equally match opponent. Further research is needed to establish this change.

The large uncertainty on the estimates for the within-season progression prevented any investigation of teams' abilities. Indeed, the only useful finding here is that there are insufficient games in a season to quantify anything less than large or very large effects. The removal of predictors from a model normally increases the uncertainty in the estimates of effects, but in the present case collinearity among the predictors and limited sample size resulted in better precision with the simpler model. The resulting uncertainty was still unacceptable for any practical application.

The unclear effects on progression arise from the fundamentally noisy nature of scores with low counts. Evidently, chance is such a major contributor to soccer outcomes that even an entire season of games is insufficient to explore performance progression. Estimates with better precision would be produced using performance indicators with higher numbers of counts as measures of team performance or effectiveness. Scoring opportunities or score box possessions as defined in Tenga et al. (2010) are two examples of such measures for soccer. Modelled progressions could also be extended to other performance indicators describing the different technical aspects of performance, such as defence, passing, crossing and goal attempts

(Oberstone, 2009). Progressions of such performance indicators would then provide evidence and help to explain the progression of game scores. A more detailed match analysis using such performance indicators was beyond of the scope of this study.

5.5. Conclusion

We have presented a novel statistical approach for using objective performance measures to investigate progression of a team. The methodology uses the generalized linear mixed model to account for the different teams' abilities via the repeated-measures structure of the data. This statistical approach will be particularly useful for analyses of other complex performance data. Although limited in its application for soccer scores, the model we have devised should be useful for modelling progression of competitive performance in sports where scores are higher.

CHAPTER 6

USING ATHLETES WORLD RANKINGS TO ASSESS PERFORMANCE OF COUNTRIES

This chapter comprises an article accepted for publication in International Journal of Sport Physiology and Performance:

Malcata, R. M., Vandenbogaerde, T. V., & Hopkins, W. G. (2014, in press). Using athletes' world rankings for assessing country performance. *International Journal of Sport Physiology and Performance*.

Overview

There is a need for fair measures of country sport performance that include athletes not winning medals. **Purpose:** To develop a measure of country performance based on athlete ranks in the sport of swimming. **Methods:** Annual top-150 ranks in Olympic pool-swimming events were downloaded for 1990 through 2011. For each athlete on a given rank, a score representing the athlete's performance potential was estimated as the proportion of athletes on that rank who ever achieved top rank. Country scores were calculated by summing its athletes' scores over all 32 events. Reliability and convergent validity were assessed via year-to-year correlations and correlations with medal counts at major competitions. The method was also applied to ranks at the 2012 Olympics to evaluate country swimming performance. **Results:** The performance score of an athlete on a given rank was closely approximated by $1/\text{rank}$. This simpler score has two practical interpretations: an athlete ranked seventh (for example) has a chance of $1/7$ of ever achieving top rank; and for purposes of evaluating country performance seven such athletes are equivalent to one athlete on the top rank. Country scores obtained by summing $1/\text{rank}$ of its athletes had high reliability and validity. This approach produced scores for 168 countries at the Olympics, whereas only 17 countries won medals. **Conclusions:** We have used the sport of swimming to develop a fair and inclusive measure representing country performance potential. This measure should be suitable for assessing countries in any sports with world rankings or with athletes at major competitions.

6.1. Introduction

Assessing a country's performance in a particular sport is important for monitoring and evaluating effects of sports policies and for providing objective feedback to relevant administrators, coaches, media and the public. In sports where a country is represented by a national team, the measure can be as simple as the team's ranking at a major international competition, such as the Soccer World Cup. In most other sports, medal count at major competitions is the usual measure of country performance (HPSNZ, 2012; Nevill, Balmer, & Winter, 2009, 2012; Steiler, 2010; UKSport, 2012). This approach has several problems. First, a medal count does not reflect a country's talent base, because it excludes performances of athletes not winning medals. Secondly, the count is biased against countries with more talent when (as is usually the case) there is a cap on the number of entries from each country. Finally, medal counts are low and therefore inherently imprecise: in any one year they provide only an approximate assessment of performance. Researchers have also assessed countries by using counts (Colwell, 1982) or proportions (V. De Bosscher, Du Bois, & Heyndels, 2012) of athletes representing a given country in lists of top-finishing placings at major competitions or in world-ranking lists, but this approach gives unacceptable equal importance to all such placings or ranks. Linear weighting scales have been used to accord more importance to better performances (Colwell, 1982; Veerle De Bosscher, De Knop, & Heyndels, 2003), but these scales give unrealistically equal importance to a step change in placings or ranks over the whole range (e.g., a step from 10th to 9th is equal to the step from 2nd to 1st with a linear scale).

The purpose of this project was to develop measures for assessing a country's performance depth in a particular sport by fairly including performances of athletes not winning medals. We chose swimming as a representative sport because of the availability of data for multiple events over many years.

6.2. Methods

Data

Annual top 150 world rankings for all 32 Olympic pool swimming events (excluding open-water events) were downloaded from www.swimnews.com for 1990 through 2011. Each athlete's rank in a calendar year in this database is determined by the athlete's best time in any competition approved by FINA (Federation Internationale de Natation). Performances in the 32 Olympic pool-swimming events at the 2012 London Olympics were also downloaded from www.london2012.com. Informed Consent was not required for approval by our institutional ethics committee, because the top-150 world rankings and Olympic results are in the public domain.

Weighting approaches

To derive annual country performance scores, we first determined appropriate values of importance (weightings) for each rank, assigned these values to the athletes, then summed the values for each country's athletes. Various approaches were investigated to determine suitable weightings, devised to represent the athletes' performance potential.

The first approach involved calculating the proportion (p) of athletes on each rank who ever achieved the top rank over the 22 years and 32 events. For example, 65 of the 566 athletes on the 10th rank achieved top rank; that is, $p = 65/566 = 0.11$. Owing to the fact that the total count of athletes achieving the top rank for athletes on each of the poor ranks (100th to 150th) was only ~5, the observed proportions showed too much scatter for a direct estimation of the weighting for each rank. We therefore defined a continuous and gradually decreasing relationship using logistic regression to model the proportion as a power function of rank: $\text{Odds} = a \cdot \text{Rank}^b$, where $\text{Odds} = p/(1-p)$, and a and b were parameters to be estimated for ranks >1 . We then converted the predicted odds for each rank back to proportions, given by $p = \text{Odds}/(\text{Odds}+1)$. The weighting of each rank was given by the predicted value of p for the rank, except for athletes on Rank 1, where the weighting was assigned the value 1. The modelling was performed via Proc Genmod in the Statistical Analysis System (Version 9.2, SAS Institute, Cary, NC), (see APPENDIX A for the SAS code).

The same linear logistic modelling was used to investigate two other weightings: proportion of athletes ever achieving any top-three rank over the 22-year period, and proportion of athletes ever winning a gold medal at a major competition (Olympic Games and World Championships) in this period. We also investigated a quadratic logistic model with each of these three approaches in an attempt to improve the predicted weightings. Additional analyses were performed by applying a logistic model separately to each stroke and distance. The fourth weighting was given simply by $1/\text{rank}$.

Validity

To investigate the validity of the measures of country performance derived by each approach, country performance scores obtained by summing the weighting for the ranks of its athletes were correlated with the number of medals won at Olympic Games or World Championships, after log-transformation. Mean correlations over the 22 years of data were obtained by averaging Fisher-transformed annual correlations and by back-transforming the mean. Validity of individual swimming events could not be investigated, owing to the low medal count (three) in any given event in any given year.

Reliability

The year-to-year reliability of the log-transformed country scores for each approach was analysed for years 2008 through 2011 using a spreadsheet (downloaded from newstats.org/xrely.xls) and presented as mean standard errors of measurement (expressed as a coefficient of variation) and mean intraclass correlation coefficients for consecutive pairs of years (W.G. Hopkins, 2000). These measures of reliability were derived for country scores summed over all 32 swimming events and for each event separately. A novel set of thresholds for assessing the magnitude of the intraclass correlation coefficient was derived by assuming that threshold values for assessing standard deviations are one half those of the modified Cohen scale for standardized differences in means (0.2, 0.6, 1.2, 2.0, and 4.0 SD) (W.G. Hopkins, 2010). Since the intraclass correlation coefficient is given by $\text{SD}^2/(\text{SD}^2 + \text{error}^2)$, substitution of the error term in

this formula with half the standardized thresholds results in the following thresholds: >0.99, extremely high; 0.99-0.92, very high; 0.92-0.74, high; 0.74-0.50, moderate; 0.50-0.20, low; <0.20, very low. If the standard error of measurement is the only source of error in a measure, it follows from the formula for the Pearson correlation $\sqrt{(SD^2/(SD^2+error^2))}$ that the square roots of these correlations are thresholds for validity correlations with an error-free criterion: >0.995, extremely high; 0.995-0.96, very high; 0.96-0.86, high; 0.86-0.71, moderate; 0.71-0.45, low; <0.45, very low. These thresholds were used to interpret the validity correlations, taking into account the fact that the criterion medal count is not free of error.

Application to 2012 London Olympic Games

To demonstrate the applicability of any of these measures for assessing country performance at major competitions, we applied the inverse-rank approach to results from the 2012 London Olympic Games. We assigned the weighting of $1/\text{rank}$ to the athletes' ranks at the Olympics, and summed these weightings into a country score. We then used these scores to rank countries and compared this ranking with country rankings based on medal count.

6.3. Results

Figure 12 shows the observed proportions of athletes on a given world rank who ever achieved top rank in the 22-year period with the best fitting logistic regression curve and the simpler inverse-rank function. Similar regression curves were obtained for the proportion of athletes who ever achieved top-three rank, for the proportion of athletes who ever won a gold medal, and for all three models applied to individual strokes and distances (data not shown). Figure 13 illustrates the weightings assigned to athlete ranks using the four approaches: inverse-rank, top-rank, top-three rank and gold-medal. Although the inverse-rank and top-rank approaches assigned similar weightings to the first ~10 ranks, the inverse-rank approach assigned relatively more weight for ranks >10. The top-three-rank approach produced the highest performance scores.

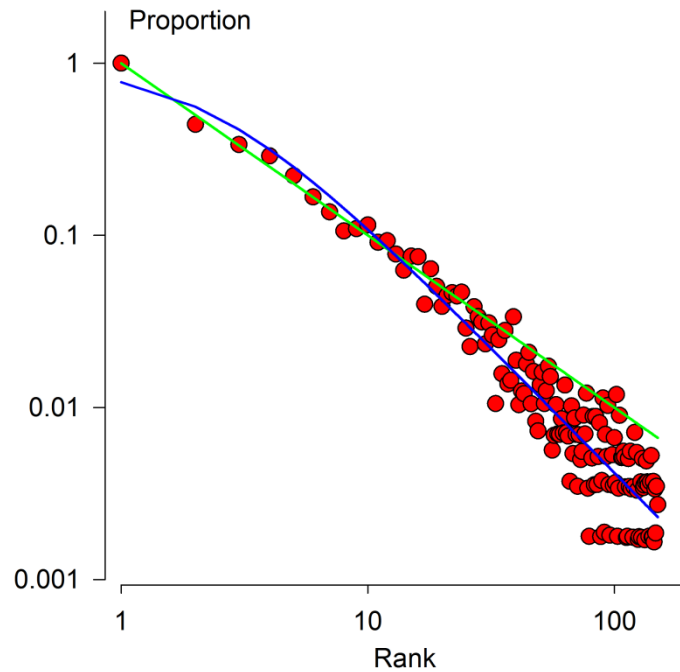


Figure 12. Observed proportion of athletes on a given world rank who ever achieved top rank in the 22-year period. The curve derived by logistic regression is $p = 3.49 \cdot \text{Rank}^{-1.46} / (1 + 3.49 \cdot \text{Rank}^{-1.46})$ (blue line), while the line is the simpler inverse-rank function (green line).

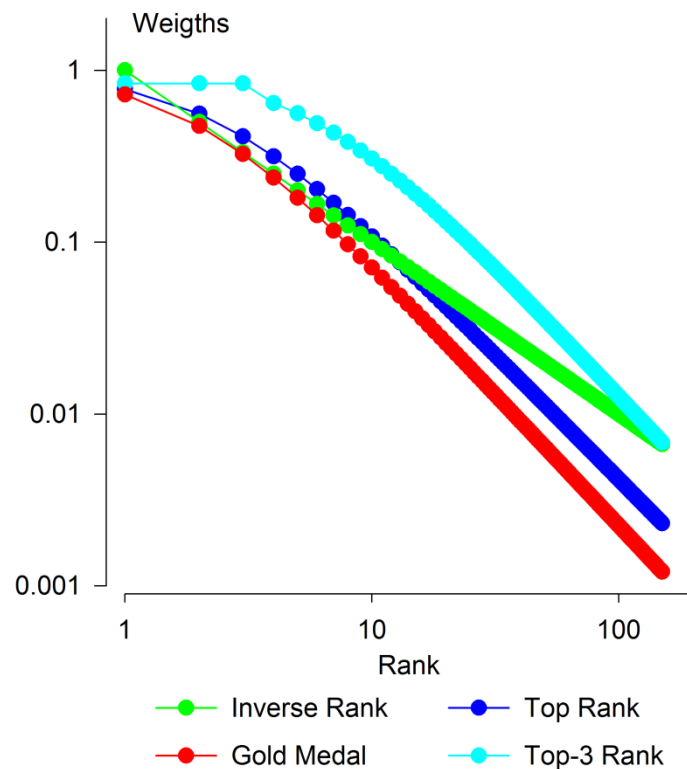


Figure 13. Four different weightings used to derive country scores from athletes' ranks based on: Inverse Rank ($w=1/\text{Rank}$), Top Rank (weightings given by the above proportion p , for $\text{Rank}>1$), Top-3 Rank (based on modelled proportion of athletes on a given world rank who ever achieved any top-three rank, $w = 11.39 \cdot \text{Rank}^{-1.48} / (1 + 11.39 \cdot \text{Rank}^{-1.48})$, for $\text{Rank}>3$) and Gold Medal (based on the modelled proportion of athletes on a given world rank who ever achieved a gold medal at major competitions, $w = 3.07 \cdot \text{Rank}^{-1.14} / (1 + 3.07 \cdot \text{Rank}^{-1.14})$, for $\text{Rank}>1$).

Table 9 presents the performance scores, rankings and medal counts for countries ranked in the top 20 in 2011 using the inverse-rank approach; performance scores and rankings are also shown for the three other approaches. The scores derived from the top-rank, top-three rank and gold-medal approaches represent the predicted number of athletes ever achieving top rank, top-three rank and gold medal respectively. The scores derived from the top-rank and inverse-rank approaches appear similar. Indeed, over the 22 years of data the correlation between the rankings derived with these two approaches was extremely high (0.996); furthermore, the inverse-rank approach predicted the ranks derived with the top-rank approach with a typical uncertainty (standard error of estimate) of $\sim\pm 0.7$ of a rank for the lowest and highest ranked countries and $\sim\pm 3$ ranks for the middle ranked countries.

As shown in Table 9, USA and China were ranked consistently first and second in 2011. Japan was ranked higher than Australia with the inverse-rank and the top-three-rank approaches (Japan 3rd Australia 4th), but lower with the top-rank and gold-medal approaches (Japan 4th, Australia 3rd). Japan had twice as many top-150 performances in that year: 692 vs 343 (data not shown in the table).

The list of countries ranked top 20 in 2011 using the inverse-rank approach includes countries not winning medals at the World Championships (e.g., Spain and New Zealand) and excludes some countries winning medals (e.g., Republic of Korea and Belarus, each winning one medal). Overall mean correlations between country scores and medal counts at Olympics and World Championships were 0.82 (90% confidence limits, 0.75-0.87), 0.84 (0.78-0.89), 0.80 (0.72-0.86) and 0.86 (0.80-0.90) for weightings based on the inverse-rank, top-rank, top-three rank and gold-medal approaches. Applying quadratic logistic models to estimate proportions improved correlations only by ~ 0.01 . The reliability of performance scores for weightings based on the inverse-rank approach was very high: a coefficient of variation of 61% (90% confidence interval 55-68%) and an intraclass correlation coefficient of 0.96 (0.94-0.97). Reliability for single events was high: coefficients of variation of $75\% \pm 18\%$ (mean \pm SD for the 32 events) and intraclass correlation coefficients of 0.86 ± 0.05 . We observed similar reliability of performance scores for the other three approaches (data not shown).

Figure 14 represents the progression in performance scores and rankings since 1990 for the three countries with the highest scores in 2011 based on the inverse-rank approach. The USA was always ranked first over this period, but the gap between first and second was much smaller in 2011. China's performance fluctuated between ranks of 10 and 2. Japan showed an upward trend between 1990 and 1998, but its performance thereafter was relatively stable.

We also used the inverse-rank approach to assess country performance at the 2012 London Olympic Games. Table 10 shows the resulting country performance scores and rankings. This assessment allowed comparing 168 countries whereas ranking countries based on medal count would have allowed comparison of 17 countries only.

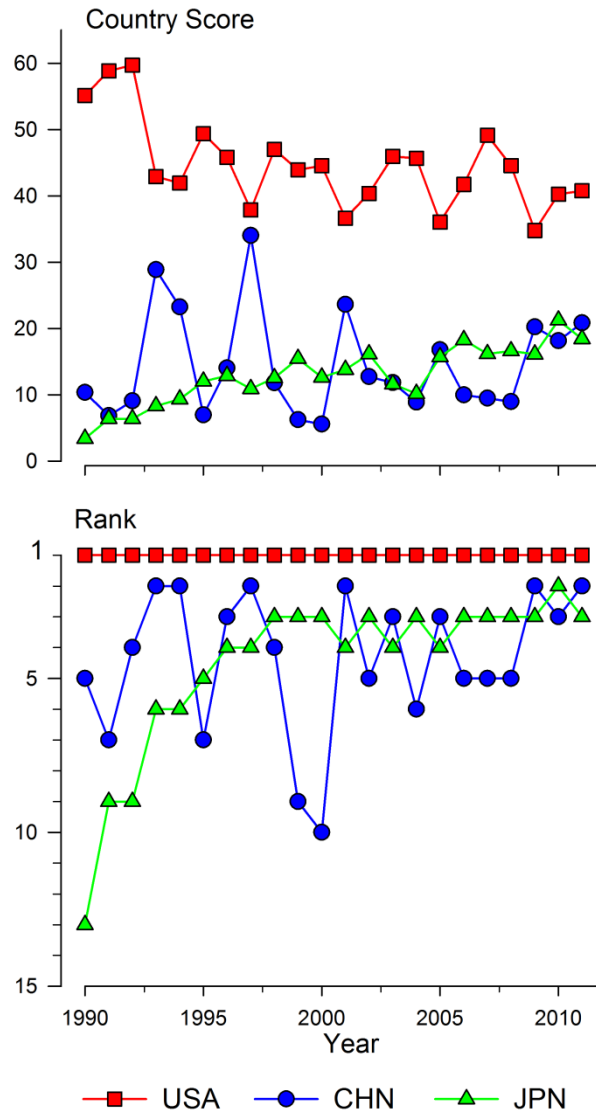


Figure 14. Progression of country performance scores and rankings from 1990 to 2011 for the top-three ranked countries in 2011, USA, China and Japan, by summing the inverse of the world ranks for all the swimmers represented in the top 150 for all of the 16 female and 16 male pool swimming events.

Table 9. Performance scores, rankings and medal counts for countries ranked in the top 20 in 2011 by summing the inverse of the world ranks for all the swimmers represented in the top 150 for all of the 16 female and 16 male pool swimming events. The scores and resulting country rankings based on three other approaches for deriving weightings are also shown: the modelled proportion of all athletes on the rank who ever achieved top rank (p top 1) and similar modelled proportions for achieving top three (p top 3) and gold medal at Olympics or World Championships (p gold).

Country	Performance score based on sum of...				Country ranking based on...				No. of medals	Medal ranking
	1/rank	p top 1 ^a	p top 3 ^b	p gold ^c	1/rank	p top 1 ^a	p top 3 ^b	p gold ^c		
USA	40.8	37.2	71.9	24.0	1	1	1	1	25	1
CHN	20.8	19.4	42.6	12.4	2	2	2	2	12	2
JPN	18.4	15.2	39.6	9.5	3	4	3	4	5	5
AUS	16.0	15.3	35.2	10.0	4	3	4	3	11	3
GBR	9.6	8.9	21.8	5.6	5	5	5	5	4	8
FRA	7.8	7.4	16.8	4.9	6	6	6	6	8	4
ITA	6.5	5.4	12.9	3.4	7	7	10	8	4	8
BRA	6.0	5.4	12.5	3.5	8	8	11	7	1	14
CAN	5.9	5.0	13.3	3.1	9	11	8	12	3	10
RUS	5.9	5.0	12.9	3.2	10	12	9	11	2	13
GER	5.4	5.1	13.3	3.2	11	10	7	10	5	5
NED	5.2	5.2	11.4	3.4	12	9	12	9	5	5
HUN	4.3	4.3	10.8	2.8	13	13	13	13	3	10
SWE	4.1	3.7	6.9	2.4	14	14	14	14	1	14
RSA	2.4	2.0	5.6	1.2	15	17	16	17	1	14
ESP	2.4	2.0	5.7	1.3	16	16	15	16	0	>21
DEN	2.1	2.1	4.8	1.4	17	15	17	15	3	10
POL	1.8	1.6	4.0	1.0	18	18	18	18	1	14
NZL	1.2	1.1	3.1	0.6	19	20	19	22	0	>21
NOR	1.1	1.1	1.2	0.7	20	21	25	20	1	14

^aScores based on proportion of athletes ever achieving top rank

^bScores based on proportion of athletes ever achieving any top-three rank

^cScores based on proportion of athletes ever winning a gold medal at Olympic Games or World Championships

Table 10. Rank-based performance scores, medal count and rankings of the top-20 countries for pool-swimming events at the 2012 London Olympic Games. The rank-based score was obtained by summing the inverse of the ranks of each country's athletes in the 32 events and can be interpreted as gold-medal equivalents.

Country	Rank-based score	Medal count	Score ranking	Medal ranking
USA	27	30	1	1
CHN	11	10	2	3
AUS	9.9	10	3	3
FRA	7.4	7	4	5
JPN	6.7	11	5	2
GBR	6.6	3	6	8
HUN	4.9	2	7	10
RSA	4.5	3	8	8
NED	4.5	4	9	6
RUS	4.3	4	10	6
CAN	4.0	2	11	10
GER	3.4	0	12	>17
ITA	2.7	0	13	>17
BRA	2.7	2	14	10
ESP	2.4	2	15	10
DEN	2.1	0	16	>17
KOR	1.8	2	17	10
SWE	1.6	0	18	>17
POL	1.5	0	19	>17
LTU	1.4	1	20	15

For country abbreviations see www.london2012.com

6.4. Discussion

Using the sport of swimming, we have successfully developed scores to track and compare the sport performance of countries. The scores are based on performances of athletes with a top-150 world rank, and therefore represent depth of performance at the top level. To accommodate different notions of how athletes' performances should be assessed for generating country performance scores, we assigned weightings (importance values) to athlete world ranks using four approaches. The top-rank, top-three-rank and gold-medal approaches were based on proportions of athletes on a given rank ever achieving each of these standards, while the fourth approach was based on the inverse of the rank.

The higher weightings and consequent higher performance scores given by the top-three-rank approach are consequence of more athletes achieving any of the top three ranks compared with only top rank or gold. The gold-medal approach produced the smallest performance scores because one year in four has neither World Championships nor Olympic Games. Another difference between the approaches apparent in Figure 13 is relatively greater weight accorded to ranks better than ~15 with the inverse-rank. These differences in the weightings between approaches produced the small differences in country rankings evident in Table 9. For example, in spite of Australia winning more medals at the 2011 World Championships compared to Japan (11 vs 5), Japan had double the number of performances in the top 150, resulting in an overall higher performance score for Japan with the inverse-rank and top-three-rank approaches.

The correlations of country scores with medal counts at major competitions were moderate for all approaches. We expected a reasonable correlation between our scores and some measure of medal winning, though the highest correlation would not necessarily determine the best approach, for the very reason that medal count is not an ideal way to appraise a country's performance depth.

Our measures are more comprehensive than measures based on medal counts or ranks at major competitions: they include performances of each country's athletes with a top-150 world ranking, whereas the country is represented by a limited number of athletes (if any) at major championships. Our measures also allow for tracking and comparison of performance depth in many more countries than those based on each country's medals. For example, only 21 countries won medals at the 2011 Swimming World Championships, whereas our measures have allowed assessment of 78 countries. Linear weighting scales that include a wide range of ranks would also produce comprehensive measures, but our measures are more appropriate, because they capture the greater importance of transitions between better ranks.

Many sports are already combining ranks at various competitions into athlete scores (tennis ATP World Tour, International Triathlon Union Rankings, Official World Golf Rankings) (rather than country scores), and in golf the approach is very similar to our inverse-rank approach. A golfer's points in a given competition are calculated by weighting the points awarded to the winner by the inverse of the golfer's rank in that competition; the points from each competition are then summed to give the golfer's score ("Official World Golf Ranking," 2012). In this way the importance of the competition is also taken into account.

Our measures were developed to evaluate country performance depth annually. Applying weighting in a similar fashion to athletes' ranks at a major competition and summing these weightings into country scores provides an inclusive measure of country performance, reflecting the conversion of performance depth into ability of athletes when it really matters.

There are several limitations to the new measures. First, to produce a performance score, a country must have at least one athlete in the list of world rankings. Countries that are not represented on the list are all ranked equally last. Secondly, performance scores are less accurate for countries with fewer athletes: the score for a country with 10 athletes on the list would be expected to change typically by ~10% if it gained or lost another athlete in the following year, whereas a country with only one athlete would change by 100%. A third limitation is that the country performance measure is only as good as the validity of the athletes' world rankings: ranking systems that depend on the number of competitions that athletes have entered (e.g., Tennis ATP World Tour list) seem to us inappropriate for applying our method to. For these sports, researchers and managers will need to devise a ranking system based purely on athletes' ability, similar to that of swimming. A fourth limitation is that discontinuity in country performance scores would be shown in any sport that has changed its method for assigning ranks to their athletes. To remove the discontinuity, the country score would need to be calculated by working backward with the new system or working forward with the old. The measures of performance depth in swimming do not include performances in open water swim events. If these need to be included, we suggest adding the inverse of the athlete's rank in major competitions to the country's performance score.

6.5. Conclusion

Using rankings of individual athletes we developed scores to track and compare the sport performance of countries. Our scores provide a more inclusive, fair and transparent measure of performance than current measures based on medal counts. Each country will have to choose the most suitable approach according to what is perceived as positive outcome. For the sport of swimming, the inverse-rank approach is our preferred method, because it does not require complex analyses, and it gives practically the same country ranking as the more sophisticated approaches based on logistic models. More importantly, the weight provided by the inverse-rank approach has two practical interpretations best understood with an example: an athlete ranked 7 has a 1-in-7 chance of ever achieving top rank, and this athlete's performance is worth only one-seventh that of a top-ranked athlete for assessing the performance of their country. A country's performance score derived by summing these weightings also has a practical interpretation: it represents the strength of the country expressed as the equivalent number of top-ranked athletes.

The measure should be applicable for tracking performance in any sport that has athlete world rankings published annually or seasonally. Statistical analyses would be required to define the weightings assigned to the ranks in the sport, as we have done for the sport of swimming. It is our suspicion that the simple and practical inverse-rank approach will closely approximate the logistically modelled top-rank and medal-winning approaches and therefore will be the best choice for any sport. The inverse-rank approach could then be applied to assess performance of

countries in any or all sports at major events, such as the Olympics. Such overall assessments of performance in research and applied settings should be an important component of the evaluation of policies relating to the allocation of resources to athletes and their medical, scientific and coaching support personnel.

CHAPTER 7

EVALUATING SPORTS PERFORMANCE OVER OLYMPIC CYCLES

This chapter comprises an article that is under embargo, unable to be submitted to a journal and to be viewed in this thesis.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

Chapter 7 is under embargo.

CHAPTER 8

DISCUSSION AND CONCLUSION

In this doctoral thesis, I have addressed the question of how to track changes in competitive performance. I started by reviewing the topic of variability in competitive performance, because estimates of variability provide thresholds for assessing magnitude of performance changes. I then developed five analytical linear models to address the research question for five levels of performance: athlete, sport, team, sport-specific squad for a country and all Olympic sports of a country (Table 15). The different kinds of performances measure—triathlon times, soccer scores, world ranking of swimmers, and athletes' placings at annual main competition—required the development of different models to characterise the relationship between performance and the time variable. Mixed linear models were used throughout, following the principle of parsimony. Goodness of fit of the models was assessed by inspection of plots of residuals (performance not explained by the model) and by evaluating the uncertainty of the estimates of effects on performance. In this discussion, I will summarise the key features of each chapter, focusing first on the topic of modelling, followed by practical applications and concluding with limitations and directions for further research.

Firstly, the topic of variability of performance was introduced in a systematic review to alert the reader that any measure will have random variation between measurements. For competitive sport performance this random variation shown by athletes from one competition to the next, also known as within-athlete variability, is the irreducible error associated with measurements of competitive performance. These estimates of variability are dependent on the sport and reflect the contribution of several factors affecting performance additional to the natural physiological variability in an athlete ability to produce power. Furthermore, estimates of typical variability shown by top athletes from competition to competition define the smallest important/worthwhile changes (W.G. Hopkins, et al., 1999) and other magnitude thresholds for assessing performance changes (W.G. Hopkins, et al., 2009). From the review of literature, estimates of variability were evaluated mainly for performances between competitions within seasons, providing thresholds for assessing magnitudes of short-term changes in competitive performance. These thresholds are used for assessing performance-enhancing strategies and other factors affecting competitive performance. For assessing long-term changes in performance, there was only a small number of studies (Bullock & Hopkins, 2009; Fulton, et al., 2009; Smith & Hopkins, 2011; Spencer, et al., 2014) in which estimates of variability of performance between competitions between seasons were reported. For purposes of assessing long-term performance changes (for example, at annual world championships) there is a need for estimates of variability of performance for competitions between seasons in a large range of sports.

Modelling

Variability of competitive performance shown by athletes from competition to competition has been estimated almost invariably with mixed linear models (Bonetti & Hopkins, 2010; Bullock, et al., 2009; Fulton, et al., 2009; W.G. Hopkins, 2005; W.G. Hopkins & Hewson, 2001; Nibali, et al., 2011; Paton & Hopkins, 2005, 2006; Pyne, et al., 2004; Smith & Hopkins, 2011; Spencer, et al., 2014). Fixed effects in the models have been used to adjust for factors affecting mean performance (e.g. competition identity, race distance or environmental conditions) so that any change in performance between competitions that is the same for all athletes does not contribute to the estimates of variability of performance. Random effects have included athlete identity to estimate differences between athletes and athlete identity interacted with year to estimate within-athlete between-year variability. The residual error in such models is the estimate of the typical variability shown by athletes from competition to competition.

In the first of my original research studies, I developed individual career trajectories of elite triathletes, investigating performance changes for the swim, cycle and run race stages and total times. Performance was modelled as a quadratic function of age and a linear function of calendar year in a mixed linear model that accounted for difference in mean race times arising from environmental and other course-related factors. Individual trajectories were achieved by specifying in mixed linear model clusters for repeated measurements for athletes and for athletes within the year. This mixed model included also a random effect for race clusters to allow adjustment to an overall mean race time, thereby excluding changes in performance that are the same for all athletes as result of environmental conditions and other venue-related factors.

In this model for triathlon performance, clusters for races provided race performance times normalized to an overall mean race time (across the full dataset of races). Because mixed modelling allows analysis of datasets with missing values, this normalization of race times is performed as if every athlete had competed in every race. It follows that the normalization of race times takes into account the ability of the athletes competing. The resulting adjusted mean represents an improvement compared with the simple rescaling to the median time used by Alam et al. (2008), in which authors did not take into consideration the ability of the athletes in each race. Additionally, the quadratic function of age used in our model proved to be sufficient to produce well-fitting trajectories over the athletes' careers (Figure 3). Applying the principle of parsimony, I did not investigate the non-linear function to describe performance and age suggested by Berthelot et al. (2012) and Guillaume et al. (2011). Furthermore, the repeated-measure structures available for non-linear mixed models in SAS may not have allowed the use of random effects to adjust for environmental and other venue-related factors, but further investigation of non-linear model for performance analysis is needed.

In Chapter 4, the model for the analysis of calendar-year trends in triathlon performance was similar to that in Chapter 3. With the adjustment for repeated measurements for athletes and estimating the extra variation arising from environmental and other external factors, I was able to investigate performance trends combining performances from various international races within a year, as opposed to previous studies in which only one race was considered (e.g., Ironman Hawaii world championship (Lepers, 2008)). The single race approach had been used to minimize

changes in performance due to variation on race-course and calibre of athletes competing, which the mixed model adjusts for. While having limitations, I assumed a linear relationship of performance with calendar year because of easy interpretation; for example, coefficients of the model provide rates of improvement (or decline) of the mean performance of the top-16 “place getters”. Linear trends have been criticised for their unrealistic predictions in the distant future (Reinboud, 2004). For the 13 years of performance I explored higher degree polynomials, but over-fitting of the data was evident and therefore deemed unsuitable to provide a general understanding of progression of mean top times. I also extrapolated the linear trends four years out to estimate predicted mean performances for the year of the coming Olympic Games. Predictions had a large uncertainty, limiting the practical application of these estimates of mean top times. Uncertainty of predictions ought to be improved using a detailed model to estimate the effects of specific environmental and course-related factors on performance, Hollings et al.(2012) have estimated such effects for track-and-field events.

When applying linear models to measures of performance that are different from the usual continuous measures of time or distance, one of the varieties of generalised linear models has to be used. Chapter 5 and 6 consisted of two examples of such models. In Chapter 5, I modelled the football game scores with an over-dispersed Poisson mixed model to compare the number of goals scored by Aspire and its opponents to answer the question of success (or failure) of a youth-talent development programme. The use of an over-dispersed Poisson to model counts such as game scores, which are likely to have a variance higher than the mean was first reported by Nevill and colleagues (2001). Game scores were modelled as a linear function of calendar year, adjusted for age difference between teams and home advantage. In agreement with previous research (Helsen, Van Winckel, & Williams, 2005; Koning et al., 2003; Lee, 1997), these two factors had clear effects on game performance. Furthermore, the model included clustering of game scores when Aspire played “repeatedly” the same team. This clustering accounted for different abilities among opposition teams, and consequently, relative ability difference between Aspire and each of the opposition teams. Accounting for teams’ abilities was a main challenge in this study. Previous research dealt with game scores either in a championship, where all teams play each other the same number of times, allowing estimation of each team’s attacking and defensive abilities (Karlis & Ntzoufras, 2003, 2009; Lee, 1997; Maher, 1982; Rue & Salvesen, 2000), or in a World Cup tournament, using FIFA World rankings to adjust for teams’ abilities (Dyte & Clarke, 2000; Koning, Koolhaas, Renes, & Ridder, 2003).

In Chapter 6, I applied a generalised linear model to develop scores for tracking performance of a country's swimming squad. The model was applied to swimming because of the availability of a relatively large amount of data in this sport. The generalised linear model was used to perform a logistic regression for estimating the proportion of swimmers with a particular world ranking that ever achieved the top rank. The proportions were the weightings (importance value) assigned to each athlete, which were then summed for each country’s athletes into a country score. It was found that the relationship between proportion and rank could be approximated by the simpler inverse of the rank function (Figure 13). In this study, I derived and justified weighting scales used

to combine individual ranks, as opposed to arbitrary scales presented in previous research to combine rankings (Veerle De Bosscher, et al., 2003) or medal colours (Wood, 2012).

The final research study (Chapter 7) was an application of simple linear regression. I used placings at annual main competitions to derive performance metrics for assessing progression of performance of individual athletes and teams over an Olympic cycle. The study is an example of a simple linear model outperforming a more sophisticated mixed linear model. The mixed model failed to produce outcomes when each event was analysed separately because of the limited amount of performance data. It produced outcomes when events within a sport were combined, but there was a poor fit for some athletes within some events. The poor fit highlights a problem with mixed modelling: it finds values of parameters of the model that give the best overall fit of the model to the data (by maximizing the predicted likelihood of the data), but the best overall fit does not imply the best fit within each cluster of data. This problem appears to become evident when there is a limited amount of data with each cluster, as here with the performances for each athlete.

The use of a linear relationship to describe progression of athlete's rankings with time in Chapter 7, in contrast to the quadratic model presented in Chapter 3 for tracking individual triathletes, is justified by the narrower time window and the limited number of performances available for each athlete (each athlete will have a maximum of five performances over an Olympic cycle). The contribution of a quadratic curvature over this time window (only four years) is likely to be negligible, and the use of a linear model allows athletes to be more easily compared and categorised in terms of improving, declining or being stable in their performance over the Olympic cycle.

An important feature of Chapter 7 was the use of log transformation. The dependent variable in this analysis was the log-transformed placings at annual main competitions. Use of log transformation implicitly assigns equal importance to the same factor differences in placings, whereas use of the untransformed variable assigns equal importance to absolute differences. For example, the difference between second and first is equivalent to the difference between tenth and fifth with log transformation, whereas it would be equal to the difference between tenth and ninth without transformation. Use of log transformation therefore gives greater and presumably more appropriate importance to changes in placings at the top end of the field.

Log transformation was used in every other research study in this thesis. As highlighted by Keene (1995) several features advocate for analysis of log-transformed data. Log transformation of triathlon performance times (Chapter 3 and 4) yielded the effects and errors as percent changes from the mean performance after back transformation. Log transformation was also applied in Chapter 4 for evaluating correlations between triathlon finishing position and rankings in each of the triathlon phases, for the reasons discussed in the previous paragraph (equal importance for a factor difference). In Chapter 5, a log link function was specified to model the over-dispersed Poisson-distributed games scores, presenting effects as percentage differences. The log link function is obligatory in generalized linear modelling to transform a count variable into a continuous variable (W.G. Hopkins, 2010). Finally, in Chapter 6, I used logistic regression to model proportion of athletes with a particular world ranking who ever achieved first rank. In this version of the generalized linear model, the dependent variable predicted by the model is the log

of the odds ($p/(1-p)$, where p is the proportion). Again, the use of log transformation is obligatory to transform a binary outcome variable (1 if achieved top rank, 0 otherwise) into a continuous variable (W.G. Hopkins, 2007b).

Practical Applications

In addition to the theoretical contribution represented by the various models for analysing progression of competitive performance, my research also led to information and practical applications useful for an applied setting. In Chapter 2, thresholds for assessing magnitudes of effects for intervention and other performance enhancing strategies in terms of competitive performance are reported and an explanation for deriving the thresholds for lab-based performance tests is included. In Chapter 3, career trajectories are themselves an analytical tool that allows a visually clear evaluation of athletes' performance changes and a direct comparison between athletes. Deviations of observed values from the modelled trajectory allow assessment of the success or failure of new training or other enhancement strategies in competitive performance outcomes. Additionally, the analyses of career trajectories of successful athletes provide the typical pattern of progression and estimates of age of peak performance (see Appendix E for examples). This evidence should assist with the setting of benchmarks for talent development programmes and guidance of athlete's career (Figure 5).

In Chapter 4, performance trends in triathlon showed improvements in men's swimming and in men's and women's running performances. The performance trends represent the changes observed for mean top 16 times and reflect the evolution of athletes' ability and any other systematic change that occurs in race courses. For example, if cycling courses are gradually being more difficult, with higher number of elevations and more technically demanding, performance times will be longer (as the decline of cycling ability of athletes is unlikely). By assessing mean performances in for the three stages of a triathlon race, it was found that for elite triathletes, running has increased in importance while cycling has decreased over the last 13 years; surprisingly the opposite trend was observed for the male juniors. Performance trends and importance of each of the race stages provide useful information for planning and preparation of triathletes for the Rio Olympics and beyond.

In these two triathlon studies (Chapter 3 and Chapter 4), predictions for both individual performances and mean performance times had limited practical application because of the relatively high uncertainty. For example, the estimated improvement of the mean running performance of top 16 triathletes over four years was 1.8% with $\pm 1.1\%$ for 90% confidence limits. In an applied setting, this information means that the average athlete currently among top 16 might have to improve as little as $\sim 1\%$ or as much as $\sim 3\%$ to continue to remain contention within the top 16 triathletes. This uncertainty arises from two sources of error: random variation in environmental conditions or other course-related factors, and the random variation athletes show from competition to competition. If information on environmental conditions and other factors (e.g., race course distance) is specified for each competition, the model will be able to account for

effects of such factors and the predictions will improve. Unfortunately, there is nothing that a performance analyst can do about the random variation shown by athletes between competitions.

In Chapter 5, I was able to answer the question regarding improvement (or decline) of competitive football team performance to objectively assess the Aspire youth-talent development programme. Although effects of linear trend were unclear over the span of time larger than two years; there was a tendency for the Aspire team to increase the number of goals scored and a tendency for the opposition to reduce the number of goals. This reduction of opposition goals is likely to be related to an improvement of Aspire defensive ability, because they had a reduced number of conceded goals. Furthermore, even with a complex model to analyse football scores, this measure of performance is too noisy to provide precise outcomes and a practical clear answer to coaches and managers. I suggest performance indicators with higher numbers of counts per game, as the ones developed by Tenga et al. (2010) to be used as measures of performance.

The country scores derived in Chapter 6 are a more inclusive, fair and transparent measure of performance than current measures based on medal counts. As stated, the inverse-rank approach is the preferred weighting method. These weightings have two practical interpretations best understood with an example: an athlete ranked 7 has a 1-in-7 chance of ever achieving top rank, and this athlete's performance is worth only one-seventh that of a top-ranked athlete for assessing the performance of their country. A country's performance score derived by summing these weightings also has a practical interpretation: it represents the strength of the country expressed as the equivalent number of top-ranked athletes. This measure of country performance should be applicable for tracking performance in any sport that has athlete world rankings published annually or seasonally. Further statistical analyses will be required to define the weightings assigned to the ranks for the particular sport, but the simple and practical inverse-rank approach will presumably approximate the weightings derived with other complex approaches and therefore will be the best choice for any sport.

In Chapter 7, performance progressions of individual athletes and teams were, again combined to assess progression of sports and countries. The two performance metrics allow identification of athletes with greater improvements compared with those of their peers and identification of athletes who "choke" (underperform) or perform above expectation at the Olympics (see appendix for practical application). These two metrics also supplement information provided by medal counts and opinions aimed at identifying successful and unsuccessful sport campaigns, thereby enhancing the strategic view and planning for the following Olympic quadrennium.

Limitations and future directions

Through the work developed in this thesis and review of relevant literature, there are areas of research needing improvement and further investigation. I have focussed here on the practical implications and applications of research in the applied setting.

Our review of literature was limited to the topic of within-athletes variability of competitive sport performance. Research on estimates of variability of competitive performance should be extended to judge-based sports and Estimates of variability of performance between competitions

between seasons should also be investigated to provide thresholds for assessing long-term changes. Particular attention should be addressed to team sports, because there has been no study on variability of team performance, and therefore there is no scale of thresholds for assessing magnitude of performance changes in such sports. This research would investigate if thresholds of 0.3, 0.9, 1.6, 2.5 and 4.0 of the typical game-to-game variation represent thresholds for a small, moderate, large, very large and extremely large changes in team performance.

Secondly, the approach for modelling performance times of triathletes accounted for effects of external factors on performance but it was unable to quantify specifically such effects. The effects of the factors affecting performance can be estimated in such models if values of the environmental variables in each race are available (e.g., temperature was 20°C for race X), this approach has been used by Hollings et al. (2012) for some of the environmental factors affecting track and field athletics performances. Nevertheless, investigation of progression of competitive performance of individual athletes should be extended to a wide range of sports. As showed here for the sport of triathlon, this objective assessment of performance enables characterization of the typical patterns of progression of successful athletes and evaluation of each athlete against those benchmarks. Modelling competitive performance should always account for effects of external factors affecting performance, because such factors will help explain some of the observed changes in performance that are not directly due to changes in ability of athletes. Either using a similar method to the one presented here (Chapter 3) or by specifically identifying and quantifying effects of such factors (Hollings, et al., 2012), investigation of progression of competitive performance should account for effect of external factors, particular in the sports where performance is directly affected by environmental and other course-related factors, such as in cycling or kayak. Kayak and other sports where athletes compete as a team involve an additional challenge, because effect of boat-crews or team composition needs to be considered.

A third topic of future investigation is a systematic review of the effects of environmental and other external factors in performance. With the coming Olympic Games in Rio de Janeiro, environmental factors, particularly temperature, have become a focus of attention. It is therefore important to establish the effects of such factors on sport performance.

Fourthly, research on performance of team sport in Chapter 4 was limited because of the low number of goals scored per match. The common problem of low counts is present in other sports, such as rugby 7's, with a common low number of counts actions to characterise player's performance. Measures of performance in which chance does not play an important role should be identified and used as alternative measure of team or individual player performance.

The research in Chapter 7 raised my awareness of the importance of progression of athletes' competitive performance for assessing sport campaigns. In retrospect, country scores as derived in Chapter 6 are limited by the fact that athletes with the same rank are equally weighted, regardless of whether the athletes' ranking is improving or declining. There is a sense that an athlete showing improvement should be weighted more than an athlete with the opposite trend. A combination of methods developed with Chapter 6 and Chapter 7 to develop a method that favours improvement could be a fifth topic for future research.

There was a level of competitive performance that was not investigated in this thesis: performance of individual athletes within a team. Such performance could be assessed by combining performance indicators into a player score that summarises the contribution of a player to team performance. The appropriate model for tracking this score would depend on how the score was derived: it may or may not require log transformation or adjustment for team in opposition, but a mixed model would likely be the best way to account for repeated measurement.

Finally, application of non-linear models to competitive performance should be explored for analysis of longitudinal trends. Linear models have provided simple and practical assessment of competitive performance trends, but all phenomena in nature are fundamentally non-linear.

Table 11 Summary of the performance metrics, models, useful findings, limitations and future directions investigated, developed and discussed in this thesis.

	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Chapter 6	Chapter 7
Title	Variability of competitive performance	Tracking career performance of successful athletes	Performance progression in elite OD triathlon ^a	Modelling progression of an academy's soccer team	Using athletes' world rankings to assess performance of countries	Evaluating sports performance over Olympic cycles
Level of competitive performance	Athlete	Athlete	Sport	Team	Country squad	All country's Olympic sport
Performance variable	Time, distance, score	Athlete's performance time	Top 16 mean performance time	Performance score	World rankings	Competition placings
Time window	Within-year and between year ^b	13 years; 2000-2012	13 years; 2000-2012	5 year ^b ; 2000-2012	22 years; 1990-2011	4-year Olympic cycle; 2004-2012
Type of model	Mixed linear model with repeated measurements for athletes	Mixed linear model with repeated measurement for athletes, cluster for race and competition level	Mixed linear model with repeated measurement for athletes and cluster for race	Generalised mixed linear model with repeated measurement for clustering Aspire and team identities	Generalised linear model	Linear model
Transformation	Log	Log	Log	Log	Logit	Log

Useful Feature

- | | | | | | |
|---|--|---|--|---|---|
| <ul style="list-style-type: none"> • Thresholds for assessing magnitude of effects in terms of competitive performance. • Explanation for deriving thresholds for field/laboratory performance tests. | <ul style="list-style-type: none"> • Quadratic trajectories for triathletes adjusting for environmental and other external factors. • Typical patterns of progression for each triathlon stage. • Comparison of an athlete with benchmarks. | <ul style="list-style-type: none"> • Linear calendar year trends for mean of top 16 athletes in Triathlon, for key races and all international races. • Calendar year trend for gender difference in OD triathlon^a • Comparison of the importance of each stage to overall performance for Elite and Junior athletes. | <ul style="list-style-type: none"> • Linear calendar year trend of Aspire and Opposition, adjusting for home-advantage, team's mean age and ability of teams in opposition. | <ul style="list-style-type: none"> • Weighting (Importance) scale for individual ranking in first-place getter equivalent • Country score by summing the weighted rankings of a nation athletes. • New method for ranking countries not limited to medal counts. | <ul style="list-style-type: none"> • Use of log transformation to equally weight factor differences between ranks (10th to 5th = 2nd-1st) • Two new performance metrics: rate of progression and Olympic effect; and percentile ranks for direct comparison to their peers. • Ranking of sports and countries by combining athletes performance metrics. |
|---|--|---|--|---|---|

Findings

- | | | | | | |
|---|--|---|--|---|---|
| <ul style="list-style-type: none"> • Different factors explain difference in variability of performance. | <ul style="list-style-type: none"> • Trajectories for running are the most comparable for overall performance. • Running is the stage where athletes show the largest improvement. | <ul style="list-style-type: none"> • Little change for overall and cycle and improvement for swim (only in the men) and run performances. • Gender differences decrease with the years. • Gender difference is the smallest in | <ul style="list-style-type: none"> • One year difference provided an advantage of 27% more goals. • Home-advantage provided 16-31% more goals. • Aspire appears to have improved their performance. | <ul style="list-style-type: none"> • 1/Rank approach is a simpler approach to quantify the importance value of athletes' ranks. • Chance of an athlete getting to | <ul style="list-style-type: none"> • Comparing NZ with average, NZ athletes won relatively more medals and improvement was as good as the average. • NZL men's rowing is ranked 11th in progression of athletes, being among the 22% |
|---|--|---|--|---|---|

- | | | | | |
|---|--|--|--|--|
| <ul style="list-style-type: none"> • Lower depth of competition for women would explain the wide range of trajectories and the lower predictability compared with that for men. • Age of peak performance is 26-28 years. | <p>swimming and the largest in running. In cycling gender difference decrease the least.</p> <ul style="list-style-type: none"> • Running stage has increased importance for elite men and women. • For junior and elite men, running and cycling appear to have opposite trends for the two age groups. | <ul style="list-style-type: none"> • Performance indicators that are related with performance outcomes, but with higher counts per game, should be used to address performance changes in soccer and other team sports, with low counts per game. | <p>the Top-1 is given by $1/\text{Rank}$</p> <ul style="list-style-type: none"> • For a country performance, an athlete is worth $1/\text{Rank}$ in terms of first-place getters. | <p>countries that improved the most.</p> <ul style="list-style-type: none"> • Women's hockey was the NZL sport that improved the most, with men's triathlon being the sports with the worse progression in comparison to their peers. |
|---|--|--|--|--|

Limitation

- | | | | | | |
|---|--|---|---|--|---|
| <ul style="list-style-type: none"> • Estimates of CVs mainly for individual sports. • Definition of upper-ranked athlete ambiguous. • Publications are mainly done by one author (W Hopkins) and his colleagues. | <ul style="list-style-type: none"> • Model does not estimate the effect of specific environmental and other external factors. • Limited application of race predictions. | <ul style="list-style-type: none"> • 4-years out predictions had a too large uncertainty to be useful. • Correlations for Junior races were done for one race in each year. | <ul style="list-style-type: none"> • Low scores per match lead to unclear effects over the 5-years period. | <ul style="list-style-type: none"> • Country scores are as valid as the ranking list used. • Country scores are more accurate for countries with several athletes in the ranking list. • Athletes in a similar rank are equally weighted, | <ul style="list-style-type: none"> • Limited data points to derived performance progression for all Olympic athletes. • Competition placing is just a measure of comparison of an athlete's performance with that of their competitors, on the day. • Uncertainty in competition placings limits prediction of performances. |
|---|--|---|---|--|---|

regardless of their progression.

- These performance metrics are not intended to evaluate consistent podium performance

Future direction

- | | | | | | |
|---|---|---|---|--|---|
| <ul style="list-style-type: none"> • Estimates of variability for more sports, particular for team sports and sports with subjective scoring. • Review of effects of environmental and other factors affecting performance. | <ul style="list-style-type: none"> • Develop career trajectories for a wide range of sports. • Explore non-linear models to describe the relationship between performance and age. • Improve precision for race predictions in order to help with race strategies. | <ul style="list-style-type: none"> • Improve precision of estimates by providing a more detailed model, with specific information of environmental and other course-related factors. | <ul style="list-style-type: none"> • Define performance indicators that measure important aspects of performance and have higher counts per match, to improve the tracking of performance changes. | <ul style="list-style-type: none"> • Take into consideration athletes' progressions, as defined in Chapter 7, when combining the weighted ranks of athletes of a country. | <ul style="list-style-type: none"> • Applied progression metrics for the log-transformed world rankings (instead of competition placings). • Extend data to lower level competitions, which implies adapting the model to accommodate the difference in competition levels. |
|---|---|---|---|--|---|

^bOlympic distance Triathlon.

^ayear is equivalent to competitive season.

Conclusion

This PhD represented an opportunity to explore the applicability of linear models to assess changes in various kinds of performance outcome using SAS software. I have modelled performance times, game scores, world rankings and placings at annual main competitions from an individual athlete level to all Olympic sports of a country. The modalities of linear models, including simple regression, generalised linear model, mixed model and generalised mixed linear model, appeared to be sufficient and efficient to represent progression of performance.

The objective assessment of performance achieved with the developed models has provided relevant information for the applied setting: characterization of typical patterns of progression of successful athletes; evaluation of performance trends in a sport; assessment of a youth-talent development programme; quantification of a country's performance within a sport; and, formulation of two measures of performance for assessing sports Olympic campaigns. Future research in the area of modelling progression of competitive performance should focus on extending assessments to a wider range of sports, equipping coaches, sport scientists and managers with objective information for evidence-based decisions.

REFERENCES

- Alam, M., Carling, K., Chen, R., & Liang, Y. (2008). How to determine the progression of young skiers? *Chance*, 21(4), 13-19.
- Albert, J. (2002). Smoothing career trajectories of baseball hitters. *Manuscript, Bowling Green State University*.
- Allen, S. V., Hopkins, W. G., & Vandenberghe, T. J. (2012, Jul 4-7). *Career performance trajectories of Olympic swimmers*. Paper presented at the 17th Annual congress of the European College Sport Science, Bruges (Belgium).
- Allison, L., & Monnington, T. (2002). Sport, prestige and international relations. *Government and Opposition*, 37(1), 106-134.
- Baker, A. B., & Tang, Y. Q. (2010). Aging performance for masters records in athletics, swimming, rowing, cycling, triathlon, and weightlifting. *Experimental Aging Research*, 36(4), 453-477.
- Balyi, I. (2001). Sport system building and long-term athlete development in British Columbia. *Canada: SportsMed BC*.
- Bartlett, R., Wheat, J., & Robins, M. (2007). Is movement variability important for sports biomechanists? *Sports Biomechanics*, 6(2), 224-243.
- Batterham, A. M., & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance*, 1, 50-57.
- Bernard, T., Sultana, F., Lepers, R., Hausswirth, C., & Brisswalter, J. (2010). Age-related decline in Olympic triathlon performance: effect of locomotion mode. *Experimental Aging Research*, 36(1), 64-78.
- Berthelot, G., Len, S., Hellard, P., Tafflet, M., Guillaume, M., Vollmer, J. C., . . . Toussaint, J. F. (2012). Exponential growth combined with exponential decline explains lifetime performance evolution in individual and human species. *Age*, 34, 1001-1009.
- Berthelot, G., Thibault, V., Tafflet, M., Escolano, S., El Helou, N., Jouven, X., . . . Toussaint, J. (2008). The citius end: world records progression announces the completion of a brief ultra-physiological quest. *PLoS One*, 3(2), e1552.
- Bird, S., Balmer, J., Olds, T., & Davison, R. (2001). Differences between the sexes and age-related changes in orienteering speed. *Journal of Sports Sciences*, 19(4), 243-252.
- Blight, S., & Rogers, S. (2012). Olympics 2012: the alternative medal table interactive. *The Guardian*, from www.theguardian.com/sport/datablog/interactive/2012/aug/07/olympics-2012-alternative-medal-table-visualised
- Bonetti, D. L., & Hopkins, W. G. (2010). Variation in performance times of elite flat-water canoeists from race to race. *International Journal of Sports Physiology and Performance*, 5(2), 210.
- Bullock, N., & Hopkins, W. G. (2009). Methods for tracking athletes' competitive performance in skeleton. *Journal of Sports Sciences*, 27(9), 937-940.
- Bullock, N., Hopkins, W. G., Martin, D. T., & Marino, F. E. (2009). Characteristics of performance in skeleton World Cup races. *Journal of Sports Sciences*, 27(4), 367-372.
- Bürge, A. (2013). *Career Development of World Class Triathletes*. Paper presented at the ITU Science and Triathlon World Conference, Lausanne.
- Chatard, J. C., Senegas, X., Selles, M., Dreanot, P., & Geyssant, A. (1995). Wetsuit effect: a comparison between competitive swimmers and triathletes. *Medicine and Science in Sports and Exercise*, 27(4), 580-586.
- Churilov, L., & Flitman, A. (2006). Towards fair ranking of olympics achievements: the case of Sydney 2000. *Comput. Oper. Res.*, 33(7), 2057-2082.
- Clarkson, S. (2012). Why are the Olympics important? *Engineering Sport*, from engineeringsport.co.uk/2012/06/26/why-are-the-olympics-important/
- Colwell, J. (1982). Quantity or quality: non-linear relationships between extent of involvement and international sporting success. In A. O. Dunleavy, A. W. Miracle & C. R. Rees (Eds.), *Studies in the sociology of sport* (pp. 101-118): Texas Christian University Press.
- De Bosscher, V., De Knop, P., & Heyndels, B. (2003). Comparing tennis success among countries. *International Sports Studies*, 25(1), 49-68.
- De Bosscher, V., Du Bois, C., & Heyndels, B. (2012). Internationalization, competitiveness and performance in athletics (1984-2006). *Sport in Society: Cultures, Commerce, Media, Politics*, 15(1), 88-102.
- Desgorges, F. D., Berthelot, G., El Helou, N., Thibault, V., Guillaume, M., Tafflet, M., . . . Toussaint, J. (2008). From Oxford to Hawaii ecophysiological barriers limit human progression in ten sport monuments. *PLoS One*, 3(11), e3653.

- Dumke, C. L., Pfaffenroth, C. M., McBride, J. M., & McCauley, G. O. (2010). Relationship between muscle strength, power and stiffness and running economy in trained male runners. *International Journal of Sports Physiology and Performance*, 5(2), 249-261.
- Dyte, D., & Clarke, S. R. (2000). A rating based Poisson model for World Cup soccer simulation. *Journal of the Operational Research Society*, 51, 993-998.
- El Helou, N., Berthelot, G., Thibault, V., Tafflet, M., Nassif, H., Campion, F., . . . Toussaint, J. F. (2010). Tour de France, Giro, Vuelta, and classic European races show a unique progression of road cycling speed in the last 20 years. *Journal of Sports Sciences*, 28(7), 789-796.
- Etter, F., Knechtle, B., Bukowski, A., Rüst, C. A., Rosemann, T., & Lepers, R. (2013). Age and gender interactions in short distance triathlon performance. *Journal of Sports Sciences*, 31(9), 996-1006.
- Flyger, N. (2009). *Variability in competitive performance of elite track cyclists*. Paper presented at the ISBS-Conference Proceedings, Limerick, Ireland.
- Fröhlich, M., Klein, M., Pieter, A., Emrich, E., & Gießing, J. (2008). Consequences of the three disciplines on the overall result in olympic-distance triathlon. *International Journal of Sports Science and Engineering*, 2(4), 204-210.
- Fulton, S. K., Pyne, D. B., Hopkins, W. G., & Burkett, B. (2009). Variability and progression in competitive performance of Paralympic swimmers. *Journal of Sports Sciences*, 27(5), 535-539.
- Gallmann, D., Knechtle, B., Rüst, C. A., Rosemann, T., & Lepers, R. (2013). Elite triathletes in 'Ironman Hawaii' get older but faster. *Age*, 1-10.
- Green, M., & Oakley, B. (2001). Elite sport development systems and playing to win: uniformity and diversity in international approaches. *Leisure studies*, 20(4), 247-267.
- Guillaume, M., Len, S., Tafflet, M., Quinquis, L., Montalvan, B., Schaal, K., . . . Toussaint, J. F. (2011). Success and Decline: Top 10 Tennis Players follow a Biphasic Course. *Medicine & Science in Sports & Exercise, e-Pub, downloaded*, 25.
- Haake, S. J., Foster, L. I., & James, D. M. (2013). An improvement index to quantify the evolution of performance in running. *Journal of Sports Sciences*(ahead-of-print), 1-13.
- Hauswirth, C., & Brisswalter, J. (2008). Strategies for improving performance in long duration events: Olympic distance triathlon. *Sports Medicine*, 38(11), 881-891.
- Haynes, J. (2000). Socio-economic impact of the Sydney 2000 Olympic Games.
- Helsen, W. F., Van Winckel, J., & Williams, A. M. (2005). The relative age effect in youth soccer across Europe. *Journal of Sports Sciences*, 23(6), 629-636.
- Hinckson, E. A., & Hopkins, W. G. (2005). Reliability of time to exhaustion analyzed with critical-power and log-log modeling. *Medicine and Science in Sports and Exercise*, 37(4), 696-701.
- Hollings, S. C., Hopkins, W. G., & Hume, P. A. (2012). Environmental and venue-related factors affecting the performance of elite male track athletes. *European Journal of Sport Science*, 12(3), 201-206.
- Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, 30(1), 1-15.
- Hopkins, W. G. (2003). *Analysis of repeated measures*. Paper presented at the Annual meeting of the American College of Sport Medicine, San Francisco, USA.
- Hopkins, W. G. (2005). Competitive performance of elite track-and-field athletes: variability and smallest worthwhile enhancements. *Sportscience*, 9, 17-20.
- Hopkins, W. G. (2006). A spreadsheet for combining outcomes from several subject groups. *Sportscience*, 10, 51-53.
- Hopkins, W. G. (2007a). A spreadsheet for deriving a confidence interval, mechanistic inference and clinical inference from ap value. *Sportscience*, 11, 16-20.
- Hopkins, W. G. (2007b). Understanding statistics by using spreadsheets to generate and analyze samples. *Sportscience*, 11, 23-36.
- Hopkins, W. G. (2010). Linear models and effect magnitudes for research, clinical and practical approaches. *Sportscience*, 14, 49-57.
- Hopkins, W. G., Hawley, J. A., & Burke, L. M. (1999). Design and analysis of research on sport performance enhancement. *Medicine and Science in Sports and Exercise*, 31(3), 472.
- Hopkins, W. G., & Hewson, D. J. (2001). Variability of competitive performance of distance runners. / Variabilite des performances de competition des coureurs de fond. *Medicine and Science in Sports and Exercise*, 33(9), 1588-1592.
- Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine and Science in Sports and Exercise*, 41(1), 3-12.

- Hopkins, W. G., Schabert, E. J., & Hawley, J. A. (2001). Reliability of power in physical performance tests. *Sports Medicine*, 31(3), 211-234.
- HPSNZ. (2012). Strategic Plan 2013-2020: High Performance Sport New Zealand.
- ITU. (2013). *ITU competition rules*. International Triathlon Union. Retrieved from www.triathlon.org/about/downloads/category/competition_rules.
- Karlis, D., & Ntzoufras, I. (2003). Analysis of sports data by using bivariate poisson models. *Journal of Royal Statistical Society. Serie D (The Statistician)*, 52(3), 381-393.
- Karlis, D., & Ntzoufras, I. (2009). Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference. *IMA Journal of Management Mathematics*, 20(2), 133-145.
- Keene, O. N. (1995). The log transformation is special. *Statistics in Medicine*, 14(8), 811-819.
- Knechtle, B., Knechtle, P., & Lepers, R. (2010). Participation and performance trends in ultra-triathlons from 1985 to 2009. *Scandinavian Journal of Medicine & Science in Sports*, 1-9.
- Knechtle, B., & Kohler, G. (2009). Running performance, not anthropometric factors, is associated with race success in a Triple Iron Triathlon. *British Journal of Sports Medicine*, 43(6), 437-441.
- Koning, R. H., Koolhaas, M., Renes, G., & Ridder, G. (2003). A simulation model for football championships. *European Journal of Operational Research*, 148, 268-276.
- Kuper, G. H., & Sterken, E. (2003). Endurance in speed skating: The development of world records. *European Journal of Operational Research*, 148(2), 293-301.
- Landers, G., Blanksby, B., Ackland, T., & Monson, R. (2008). Swim Positioning and its Influence on Triathlon Outcome. *International Journal of Exercise Science*, 1(3), 96-105.
- Landers, G., Blanksby, B., Ackland, T., & Smith, D. (2000). Morphology and performance of world championship triathletes. *Annals of Human Biology*, 27(4), 387-400.
- Lee, A. J. (1997). Modeling scores in the premier league: Manchester United really the best? *Chance*, 10, 15-19.
- Lepers, R. (2008). Analysis of Hawaii Ironman performances in elite triathletes from 1981 to 2007. *Medicine and Science in Sports and Exercise*, 40(10), 1828-1834.
- Lepers, R., Knechtle, B., Knechtle, P., & Rosemann, T. (2011). Analysis of ultra-triathlon performances. *Open Access Journal of Sports Medicine*, 2, 131-136.
- Lepers, R., Knechtle, B., & Stapley, P. J. (2013). Trends in Triathlon Performance: Effects of Sex and Age. *Sports Medicine*, 1-13.
- Lepers, R., & Maffiuletti, N. (2011). Age and gender interactions in ultraendurance performance: insight from the triathlon. *Medicine and Science in Sports and Exercise*, 43(1), 134-139.
- Lepers, R., & Stapley, P. J. (2010). Differences in gender and performance in off-road triathlon. *Journal of Sports Sciences*, 28(14), 1555-1562.
- Lepers, R., & Stapley, P. J. (2011). Age-related changes in conventional road versus off-road triathlon performance. *European Journal of Applied Physiology*, 111(8), 1687-1694.
- Lepers, R., Sultana, F., Bernard, T., Hausswirth, C., & Brisswalter, J. (2009). Age-related changes in triathlon performances. *International Journal of Sports Medicine*, 31(4), 251-256.
- Lippi, G., Banfi, G., Favaloro, E. J., Rittweger, J., & Maffulli, N. (2008). Updates on improvement of human athletic performance: focus on world records in athletics. *British Medical Bulletin*, 87(1), 7-15.
- Magnay, J. (2012). Team GB medal target for London 2012 Olympics is fourth place with 48 medals across 12 sports. *The Telegraph*, from www.telegraph.co.uk/sport/olympics/9374912/Team-GB-medal-target-for-London-2012-Olympics-is-fourth-place-with-48-medals-across-12-sports.html
- Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36(3), 109-118.
- McGuigan, M. R., & Kane, M. K. (2004). Reliability of performance of elite Olympic weightlifters. *Journal of Strength and Conditioning Research*, 18(3), 650-653.
- Mendez-Villanueva, A., Mujika, I., & Bishop, D. (2010). Variability of competitive performance assessment of elite surfboard riders. *Journal of Strength & Conditioning Research*, 24(1), 135.
- Mikulic, P. (2011). Maturation to elite status: a six-year physiological case study of a world champion rowing crew. *European Journal of Applied Physiology*, 111(9), 2363-2368.
- Millet, G., & Vleck, V. (2011). Triathlon Specificity. *World Book of Swimming, from science to performance*, 25, 481-495.
- Millet, G. P., & Bentley, D. J. (2004). The physiological responses to running after cycling in elite junior and senior triathletes. *International Journal of Sports Medicine*, 25, 191-197.

- Millet, G. P., Dréano, P., & Bentley, D. J. (2003). Physiological characteristics of elite short-and long-distance triathletes. *European Journal of Applied Physiology*, 88(4-5), 427-430.
- Muehlbauer, T., Schindler, C., & Panzer, S. (2010). Pacing and sprint performance in speed skating during a competitive season. *International Journal of Sports Physiology and Performance*, 5(2), 165-176.
- Mujika, I., Vaeyens, R., Matthys, S. P. J., Santisteban, J., Goiriena, J., & Philippaerts, R. (2009). The relative age effect in a professional football club setting. *Journal of Sports Sciences*, 27(11), 1153-1158.
- Nevill, A. M., Atkinson, G., Hughes, M. D., & Cooper, S. M. (2001). Statistical methods for analysing discrete and categorical data recorded in performance analysis. *Journal of Sports Sciences*, 20(10), 829-844.
- Nevill, A. M., Balmer, N. J., & Winter, E. M. (2009). Why Great Britain's success in Beijing could have been anticipated and why it should continue beyond 2012. *British Journal of Sports Medicine*, 43, 1108-1110.
- Nevill, A. M., Balmer, N. J., & Winter, E. M. (2012). Congratulations to team GB, but why should we be so surprised? Olympic medal count can be predicted using logit regression models that include "home advantage". *British Journal of Sports Medicine*, Online First.
- Nevill, A. M., & Whyte, G. (2005). Are there limits to running world records? *Medicine and Science in Sports and Exercise*, 37(10), 1785.
- Nibali, M., Hopkins, W. G., & Drinkwater, E. (2011). Variability and predictability of elite competitive slalom canoe-kayak performance. *European Journal of Sport Science*, 11(2), 125-130.
- Oberstone, J. (2009). Differentiating the top english Premier League football clubs from the rest of the pack: identifying the keys to success. *Journal of Quantitative Analysis in Sport*, 5(3).
- Official World Golf Ranking. (2012) Retrieved 18/10/2012, 2012, from http://www.officialworldgolfranking.com/about_us/
- Paton, C. D., & Hopkins, W. G. (2005). Competitive performance of elite olympic-distance triathletes: reliability and smallest worthwhile enhancement. *Sportscience*, 9, 1-5.
- Paton, C. D., & Hopkins, W. G. (2006). Variation in performance of elite cyclists from race to race. *European Journal of Sport Science*, 6(1), 25-31.
- Pike, J., Hopkins, W. G., & Nottle, C. (2010, Jun 16-19). Overall trends and individual trajectories of swimming performances in a decade of New Zealand national championships. Paper presented at the XIth International Symposium on Biomechanics and Medicine in Swimming, Oslo (Norway).
- Putt, G. D. (2013). National performance versus population at the Olympic games: a methodology for determining elite performers. *International Journal of Sports Science*, 3(3), 74-80.
- Pyne, D. B., Trewin, C. B., & Hopkins, W. G. (2004). Progression and variability of competitive performance of Olympic swimmers. *Journal of Sports Sciences*, 22(7), 613-620.
- Radicchi, F. (2012). Universality, limits and predictability of gold-medal performances at the Olympic Games. *PloS one*, 7(7), e40335.
- Reinboud, W. (2004). Linear models can't keep up with sport gender gap. *Nature*, 432(7014), 147-147.
- Rue, H., & Salvesen, O. (2000). Prediction and retrospective analysis of soccer matches in a league *Journal of Royal Statistical Society- serie D (The Statistician)*, 49(3), 399-418.
- Rüst, C. A., Knechtle, B., Knechtle, P., Rosemann, T., & Lepers, R. (2012a). Age of peak performance in elite male and female Ironman triathletes competing in Ironman Switzerland, a qualifier for the Ironman world championship, Ironman Hawaii, from 1995 to 2011. *Open Access Journal of Sports Medicine*, 3, 175-182.
- Rüst, C. A., Knechtle, B., Knechtle, P., Rosemann, T., & Lepers, R. (2012b). European athletes dominate performances in Double Iron ultra-triathlons- a retrospective data analysis from 1985 to 2010. *European Journal of Sport Science*, 2012, 1-12.
- Rüst, C. A., Knechtle, B., & Rosemann, T. (2012). Women achieve peak freestyle swim speed at earlier ages than men. *Open Access Journal of Sports Medicine*, 3, 189-199.
- Saaty, T. L. (2010). Who won the Winter 2010 Olympics? A quest into priorities and rankings. *Journal of Multi-Criteria Decision Analysis*, 17(1-2), 25-36.
- Saunders, P. U., Pyne, D. B., Telford, R. D., & Hawley, J. A. (2004). Factors affecting running economy in trained distance runners. *Sports Medicine*, 34(7), 465-485.
- Seiler, S. (2013). Evaluating the (your country here) olympic medal count. *International Journal of Sports Physiology and Performance*, 8(2), 203-210.
- Smith, T. B., & Hopkins, W. G. (2011). Variability and predictability of finals times of elite rowers. *Medicine and Science in Sports and Exercise*, 43(11), 2155-2160.

- Sowell, C. B., & Mounts Jr, W. S. (2005). Ability, Age, and Performance Conclusions From the Ironman Triathlon World Championship. *Journal of Sports Economics*, 6(1), 78-97.
- Spencer, M., Losnegard, T., Hallén, J., & Hopkins, W. G. (2014). Variability and predictability of performance times of elite cross-country skiers. *International Journal of Sports Physiology and Performance*, 9(1), 5-11.
- Steiler, S. (2010). Who won the Olympics? *Sportscience*, 14, 1-6.
- Stevenson, J., Song, H., & Cooper, J. A. (2012). Age and Sex Differences Pertaining to Modes of Locomotion in Triathlon. *Medicine and Science in Sports and Exercise*, 976-984.
- Stiefel, M., Knechtle, B., & Lepers, R. (2012). Master triathletes have not reached limits in their Ironman triathlon performance. *Scandinavian Journal of Medicine & Science in Sports*(ahead-of-print).
- Tatem, A. J., Guerra, C. A., Atkinson, P. M., & Hay, S. (2004). Momentous sprint at the 2156 Olympics?: Women sprinters are closing the gap on men and may one day overtake them. *Nature*, 431(7008), 525.
- Tenga, A., Ronglan, L., & Bahr, R. (2010). Measuring the effectiveness of offensive match-play in professional soccer. *European Journal of Sport Science*, 10(4), 269-277.
- Trewin, C. B., Hopkins, W. G., & Pyne, D. B. (2004). Relationship between world-ranking and Olympic performance of swimmers. *Journal of Sports Sciences*, 22(4), 339-345.
- Truxillo, C., Amrhein, J., Dickey, D., Littell, R., Lucas, B., & Modlin, D. (2012). *Mixed Models Analyses Using SAS: Course notes*. Cary: SAS Institute.
- UKSport. (2012). Annual report and accounts 2011/2012: UK Sport.
- Van Hilvoorde, I., Elling, A., & Stokvis, R. (2010). How to influence national pride? The Olympic medal index as a unifying narrative. *International Review for the Sociology of Sport*, 45(1), 87-102.
- Vleck, V. E., Bentley, D. J., Millet, G. P., & Bürgi, A. (2008). Pacing during an elite Olympic distance triathlon: comparison between male and female competitors. *Journal of Science and Medicine in Sport*, 11(4), 424-432.
- Vleck, V. E., Bürgi, A., & Bently, D. J. (2006). The consequences of swim, cycle, and run performance on overall result in elite olympic distance triathlon. *International Journal of Sports Medicine*, 27(1), 43-48.
- Wainer, H., Njue, C., & Palmer, S. (2000). Assessing Time Trends in Sex Differences in Swimming & Running. *Chance*, 13(1), 10-15.
- Wood, R. (2012). Medal tally ranking systems. *Topendsports*, from www.topendsports.com/events/summer/medal-tally/rankings.htm
- Yahoo. (2012). Paula Findlay breaks down, apologizes after Olympic women's triathlon. *Yahoo! Sports Canada* Retrieved Jan 29, 2013, from ca.sports.yahoo.com/blogs/eh-game/paula-findlay-breaks-down-apologizes-olympic-women-triathlon-144038613.html
- Young, B., & Starkes, J. (2005). Career-span analyses of track performance: longitudinal data present a more optimistic view of age-related performance decline. *Experimental Aging Research*, 31, 69-90.
- Zanoletti, C., La Torre, A., Merati, G., Rampinini, E., & Impellizzeri, F. M. (2006). Relationship between push phase and final race time in skeleton performance. *Journal of Strength and Conditioning Research*, 20(3), 579-583.

APPENDIX A

SAS LINEAR MODELS

Below are the examples of the specification of the linear models in the Statistical Analysis systems for each of the individual studies. I have also provided a small sample of the dataset and variables used in the models.

Chapter 2- Estimation of variability of surf performance scores

Ten observations of dataset topathletes.

Obs	Event	Athlete	Points	logpoints
137	5	Nunes	288	566.296
138	6	Nunes	288	566.296
139	7	Nunes	288	566.296
140	8	Nunes	288	566.296
239	8	Winkler, Lee	480	617.379
240	9	Winkler, Lee	288	566.296
241	10	Winkler, Lee	1032	693.925
243	1	Winter, Russell	480	617.379
244	2	Winter, Russell	480	617.379
245	3	Winter, Russell	288	566.296

```
proc mixed data=topathletes covtest cl alpha=0.1 ;
class Event Athlete;
model logpoints=Event/s outp=predtop residual noint ddfm=kr;
random Athlete/s;
ods output solutionr=solrtop;
ods output solutionf=solftop;
ods output classlevels=clevtop;
ods output covparms=covtop;
run;
```

Chapter 3- Mixed model for individual trajectories for running performance

Ten observation of dataset allmixed.

Obs	Gender	Short kind			RaceId	Year		Athlete
		Long	Top	Type		Year	X	
5635	Men	0	1	Junior	2007HamburgJunior	2007	7	AlistairBrownlee
6503	Men	1	1	WCup	2007Rhodes	2007	7	AlistairBrownlee
6852	Men	1	1	WChamps	2008BeijingOlympics	2008	8	AlistairBrownlee
7207	Men	1	1	WCup	2008Lisbon	2008	8	AlistairBrownlee
7856	Men	1	1	WChamps	2009Kitzbuehel	2009	9	AlistairBrownlee
7905	Men	1	1	WChamps	2009London	2009	9	AlistairBrownlee
8755	Men	1	1	WChamps	2010London	2010	10	AlistairBrownlee
8968	Men	1	1	WCup	2010Athlone	2010	10	AlistairBrownlee
9670	Men	1	1	WChamps	2011Kitzbuehel	2011	11	AlistairBrownlee
9720	Men	0	1	WChamps	2011Lausanne	2011	11	AlistairBrownlee

Obs	LogRun	Run	Age0	Age0Sqr
-----	--------	-----	------	---------

5635	680.128	0:14:59	-6.1485	37.804
6503	757.199	0:32:23	-6.0444	36.534
6852	756.372	0:32:07	-5.1803	26.836
7207	754.433	0:31:30	-5.4536	29.741
7856	751.534	0:30:36	-4.28546	18.3651
7905	745.182	0:28:43	-4.18957	17.5525
8755	750.769	0:30:22	-3.24984	10.5615
8811	753.476	0:31:12	-3.38409	11.4520
8968	752.510	0:30:54	-3.30737	10.9387
9670	752.887	0:31:01	-2.34847	5.5153
9720	677.651	0:14:37	-2.17587	4.7344

```

proc hpmixed data=allmixed ;
class Gender ShortLong Venue KindTop Type RaceId Athlete Year;
Model LogRun=ShortLong*KindTop YearX*KindTop Age0*KindTop
      Age0Sqr*KindTop /noint s;
random int Age0 Age0Sqr/subject=Athlete Group=KindTop s CL alpha=0.1
      type=&variance;
random Year/subject=Athlete Group=KindTop Cl alpha=0.1 s;
Random RaceId/ CL alpha=0.1 s;
Random RaceId*varWCup*Athlete/ CL alpha=0.1 s;
Random RaceId*varJun*Athlete / CL alpha=0.1 s;
output out=predmixed pred=Pred stderr=StdErr resid=Resid
      student=StudentResid LCL(blup)=Lower UCL(blup)=Upper
      pred(noblup)=predMean stderr(noblup)=StdErrMean
      LCL(noblup)=LowerMean UCL(noblup)=UpperMean;
ods output estimates=estmixed;
ods output solutionr=solrmixed;
ods output parameterestimates=solfmixed;
ods output classlevels=clevmixed;
ods output diffs=lsmdifmixed;
ods output covparms=covmixed;
ods output lsmeans=lsmmixed;
by Gender;
run;
ods listing;

```

Chapter 4- Mixed model for running performance trend

Ten observation of dataset allmixed2.

Obs	Gender	Year	YearX	RaceId	Athlete	LogRun	Run	WCup
2608	Men	2012	-4	2012Auckland	JavierGomez	751.425	0:30:34	0
2609	Men	2012	-4	2012Auckland	JonathanBrownlee	751.425	0:30:34	0
2624	Men	2012	-4	2012Banyoles	LukasVerzbicas	752.564	0:30:55	1
2625	Men	2012	-4	2012Banyoles	LaurentVidal	754.009	0:31:22	1
2656	Men	2012	-4	2012Huatulco	SimonDeCuyper	758.223	0:32:43	1
2657	Men	2012	-4	2012Huatulco	RyanSissons	758.883	0:32:56	1
2688	Men	2012	-4	2012Kitzbuehel	AlistairBrownlee	749.053	0:29:51	0
2689	Men	2012	-4	2012Kitzbuehel	JonathanBrownlee	751.589	0:30:37	0
2752	Men	2012	-4	2012SanDiego	JonathanBrownlee	749.554	0:30:00	0
2753	Men	2012	-4	2012SanDiego	SvenRiederer	749.721	0:30:03	0

```

proc hpmixed data=allmixed2;
class Gender RaceId Year Athlete;
Model LogRun= YearX / s;
random int Year RaceID*varWCup /subject=Athlete s;
random RaceId / s;
Output out=predmean pred=Pred stderr=StdErr resid=Resid
      student=StudentResid LCL=Lower UCL=Upper

```

```

Pred(noblup)=PredMean StdErr(noblup)=StdErrMean
LCL(noblup)=LowerMean UCL(noblup)=UpperMean;
estimate "2000Mean" int 1 YearX -16 /cl alpha=0.1;
estimate "2001Mean" int 1 YearX -15 /cl alpha=0.1;
estimate "2002Mean" int 1 YearX -14 /cl alpha=0.1;
estimate "2003Mean" int 1 YearX -13 /cl alpha=0.1;
estimate "2004Mean" int 1 YearX -12 /cl alpha=0.1;
estimate "2005Mean" int 1 YearX -11 /cl alpha=0.1;
estimate "2006Mean" int 1 YearX -10 /cl alpha=0.1;
estimate "2007Mean" int 1 YearX -9 /cl alpha=0.1;
estimate "2008Mean" int 1 YearX -8 /cl alpha=0.1;
estimate "2009Mean" int 1 YearX -7 /cl alpha=0.1;
estimate "2010Mean" int 1 YearX -6 /cl alpha=0.1;
estimate "2011Mean" int 1 YearX -5 /cl alpha=0.1;
estimate "2012Mean" int 1 YearX -4 /cl alpha=0.1;
estimate "2013Mean" int 1 YearX -3 /cl alpha=0.1;
estimate "2014Mean" int 1 YearX -2 /cl alpha=0.1;
estimate "2015Mean" int 1 YearX -1 /cl alpha=0.1;
estimate "2016Mean" int 1 YearX 0 /cl alpha=0.1;
ods output estimates=estmean;
ods output solutionr=solrmean;
ods output parameterestimates=solfmean;
ods output classlevels=clevmean;
ods output covparms=covmean;
by gender;
run;

```

Chapter 5- Generalised linear mixed model for soccer scores

Ten observation of dataset dat1.

Obs	Intake	Team	TeamOpp	Points	Aspire01	Home Away	Age Aspire MinusOpp	AgeOpp Minus Aspire
1	Team91	Aspire	Ajax	2	1	FH	1	0
2	Team91	Other	Ajax	2	0	FA	0	-1
3	Team91	Other	Al Ain	3	0	FH	0	0
4	Team91	Aspire	Al Ain	0	1	FA	0	0
5	Team91	Other	Al Ain	0	0	FH	0	0
6	Team91	Aspire	Al Ain	0	1	FA	0	0
7	Team91	Aspire	Al Arabi	0	1	FH	-1	0
8	Team91	Other	Al Arabi	0	0	FA	0	1
9	Team91	Other	Al Jazeera	1	0	FH	0	0
10	Team91	Aspire	Al Jazeera	4	1	FA	0	0

```

proc glimmix data=dat1;
class Intake Team TeamOpp Aspire01 HomeAway;
model Points=Intake*Team Intake*Team*Year AgeAspireMinusOpp
AgeOppMinusAspire Team*HomeAway/s link=log dist=Poisson
noit;
random int/subject=Intake*Team*TeamOpp s;
random _residual_/s;
estimate "AgeAspireMinusOpp" AgeAspireMinusOpp 1/alpha=0.1;
estimate "AgeOppMinusAspire " AgeOppMinusAspire 1/alpha=0.1;
estimate "Mean age effect" AgeAspireMinusOpp .5 AgeOppMinusAspire
.5/alpha=0.1;
estimate "";
estimate "Aspire Home advantage" Team*HomeAway -1 1 0 0/alpha=0.1;
estimate "Other Home advantage" Team*HomeAway 0 0 -1 1/alpha=0.1;

```

```

estimate "Other-Aspire Home advantage" Team*HomeAway 1 -1 -1
1/alpha=0.1;
estimate "Mean home advantage" Team*HomeAway -1 1 -1 1/alpha=0.1
divisor=2;
estimate "";
estimate "Aspire 2005" Intake*Team 1 0 1 0 1 0 Intake*Team*Year 5 0 5
0 5 0 Team*HomeAway 1.5 1.5 0 0/alpha=0.1 divisor=3;
estimate "Aspire 2006" Intake*Team 1 0 1 0 1 0 Intake*Team*Year 6 0 6
0 6 0 Team*HomeAway 1.5 1.5 0 0/alpha=0.1 divisor=3;
estimate "Aspire 2007" Intake*Team 1 0 1 0 1 0 Intake*Team*Year 7 0 7
0 7 0 Team*HomeAway 1.5 1.5 0 0/alpha=0.1 divisor=3;
estimate "Aspire 2008" Intake*Team 1 0 1 0 1 0 Intake*Team*Year 8 0 8
0 8 0 Team*HomeAway 1.5 1.5 0 0/alpha=0.1 divisor=3;
estimate "Aspire 2009" Intake*Team 1 0 1 0 1 0 Intake*Team*Year 9 0 9
0 9 0 Team*HomeAway 1.5 1.5 0 0/alpha=0.1 divisor=3;
estimate "Aspire 2010" Intake*Team 1 0 1 0 1 0 Intake*Team*Year 10 0
10 0 10 0 Team*HomeAway 1.5 1.5 0 0/alpha=0.1 divisor=3;
estimate "";
estimate "Aspire 09/05" Intake*Team*Year 4 0 4 0 4 0/alpha=0.1
divisor=3;
estimate "";
estimate "Other 2005" Intake*Team 0 1 0 1 0 1 Intake*Team*Year 0 5 0
5 0 5 Team*HomeAway 0 0 1.5 1.5/alpha=0.1 divisor=3;
estimate "Other 2006" Intake*Team 0 1 0 1 0 1 Intake*Team*Year 0 6 0
6 0 6 Team*HomeAway 0 0 1.5 1.5/alpha=0.1 divisor=3;
estimate "Other 2007" Intake*Team 0 1 0 1 0 1 Intake*Team*Year 0 7 0
7 0 7 Team*HomeAway 0 0 1.5 1.5/alpha=0.1 divisor=3;
estimate "Other 2008" Intake*Team 0 1 0 1 0 1 Intake*Team*Year 0 8 0
8 0 8 Team*HomeAway 0 0 1.5 1.5/alpha=0.1 divisor=3;
estimate "Other 2009" Intake*Team 0 1 0 1 0 1 Intake*Team*Year 0 9 0
9 0 9 Team*HomeAway 0 0 1.5 1.5/alpha=0.1 divisor=3;
estimate "Other 2010" Intake*Team 0 1 0 1 0 1 Intake*Team*Year 0 10
0 10 0 10 Team*HomeAway 0 0 1.5 1.5/alpha=0.1 divisor=3;
estimate "";
estimate "Other 09/05" Intake*Team*Year 0 4 0 4 0 4/alpha=0.1
divisor=3;
estimate "";
estimate "Aspire/Other 2005" Intake*Team 1 -1 1 -1 1 -1
Intake*Team*Year 5 -5 5 -5 5 -5
Team*HomeAway 1.5 1.5 -1.5 -1.5/alpha=0.1 divisor=3;
estimate "Aspire/Other 2009" Intake*Team 1 -1 1 -1 1 -1
Intake*Team*Year 9 -9 9 -9 9 -9
Team*HomeAway 1.5 1.5 -1.5 -1.5/alpha=0.1 divisor=3;
estimate "";
estimate "Aspire/Other /4y" Intake*Team*Year 4 -4 4 -4 4 -4/alpha=0.1
divisor=3;
estimate "Aspire/Other /3y" Intake*Team*Year 3 -3 3 -3 3 -3/alpha=0.1
divisor=3;
estimate "Aspire/Other /2y" Intake*Team*Year 2 -2 2 -2 2 -2/alpha=0.1
divisor=3;
estimate "Aspire/Other /1y" Intake*Team*Year 1 -1 1 -1 1 -1/alpha=0.1
divisor=3;
ods output classlevels=clev;
ods output solutionf=solf;
ods output solutionr=solr;
ods output estimates=est;
ods output covparms=cov;
run;

```

Chapter 6- Generalised linear model (logistic regression) for modelling proportion

Ten observation of dataset all_rank2.

Obs	Rank1	Total	LogRank	Gender	Stroke	Distance
1	4	22	0.69315	men	Backstroke	100
2	3	22	1.09861	men	Backstroke	100
3	5	23	1.38629	men	Backstroke	100
4	2	22	1.60944	men	Backstroke	100
5	0	21	1.79176	men	Backstroke	100
6	1	22	1.94591	men	Backstroke	100
7	1	25	2.07944	men	Backstroke	100
8	1	20	2.19722	men	Backstroke	100
9	1	22	2.30259	men	Backstroke	10
10	0	24	2.39790	men	Backstroke	100

```
proc glimmix data=rank_all2 pconv=1E-5;
model Rank1X/TotalX=LogRankX /s link=logit dist=bin;
random _residual_;
output out=glimmixout_all pred(ilink)=PredictedAll
      resid(ilink)=ResidualAll;
ods output covparms=cov_all;
ods output parameterestimates=parmest_all;
ods output lsmeans=lsmean_all;
ods output diffs=lsmdiff_all;
ods output solutionr=solr_all;
ods output classlevels=clev_all;
ods output estimates=est_all;
run;
ods listing;
```

Chapter 7- Linear regression of individual athletes placings at annual main competitions

Ten observation of dataset data5.

Obs	OlyCycle	Gender	Sport	Event	Person_Team	LogRank	Rank	
Yeardif								
56014	2012	Men	Sailing	49er	Burling/Tuke	69.315	2	0.00000
56015	2012	Men	Sailing	49er	Burling/Tuke	69.315	2	-0.23770
56016	2012	Men	Sailing	49er	Burling/Tuke	69.315	2	-0.63945
56017	2012	Men	Sailing	49er	Burling/Tuke	283.321	17	-2.57918
56018	2012	Men	Sailing	49er	Burling/Tuke	325.810	26	-3.05589
56027	2012	Men	Sailing	49er	Chekrygin/Russu	407.754	59	-0.23770
56028	2012	Men	Sailing	49er	Chekrygin/Russu	438.203	80	-3.05589
56037	2012	Men	Sailing	49er	D'Ortoli/Delpech	207.944	8	-0.23770
56038	2012	Men	Sailing	49er	D'Ortoli/Delpech	336.730	29	-2.57918
56039	2012	Men	Sailing	49er	D'Ortoli/Delpech	248.491	12	-3.05589
56043	2012	Men	Sailing	49er	D.Evans/&			

```
proc reg data=data5;
model LogRank=Yeardif;
by sport Gender Event Person_Team OlyCycle;
Ods output parameterestimates=logregression;
run;
```



APPENDIX B

Oral presentation presented at European College of Sport Science, in Barcelona, Spain, 2013.

Career Trajectories of successful triathletes

Rita M Malcata^{1,2}, Simon Pearson², Will G Hopkins¹
¹Sport Performance Research Institute, AUT University, Auckland, New Zealand
²High Performance Sport New Zealand, Auckland, New Zealand

June 2013

Olympic Distance Triathlon




Objective: To develop career trajectories for triathletes for swim, bike and run and overall performance, with adjustment for the different races




Progression of Individual Athletes


- Monitoring athletes' competition performances is important
 - to set performance targets
 - to guide training programs
 - to select talented athletes
 - to support funding decisions




Consistent Race Environments
 Examples: swimming, track-and-field athletics




Straightforward



Variable Race Environments
 Examples: triathlon, rowing




Problem of extra variation


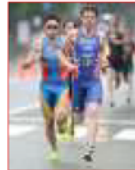


Data

- Splits and finishing times from races between 2000-2012 (www.triathlon.org):
 - Olympic Games
 - World Championships Series (including Junior and U23)
 - World Cup
 - European Championships
- Race date and athlete's date of birth
- Final sample
 - 427 men, 224 races
 - 337 women, 222 races
- Athletes were grouped according to their finishing position in any World Championship or Olympic Games 2008-2012:
 - Top 3**- Finished within top 3
 - Top 16**- Finished between 4th-16th
 - Other**- Finished outside top 16




Olympic Distance Triathlon


→

→



1.5km Swim 40 km Bike 10 km Run

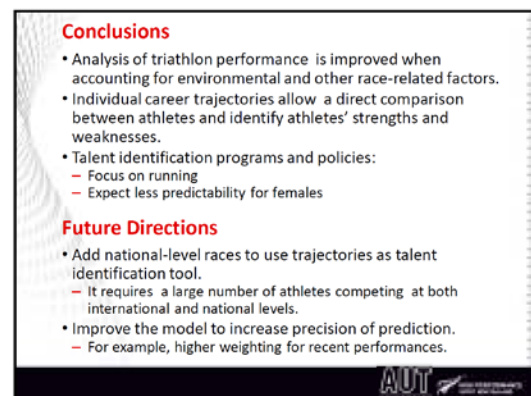
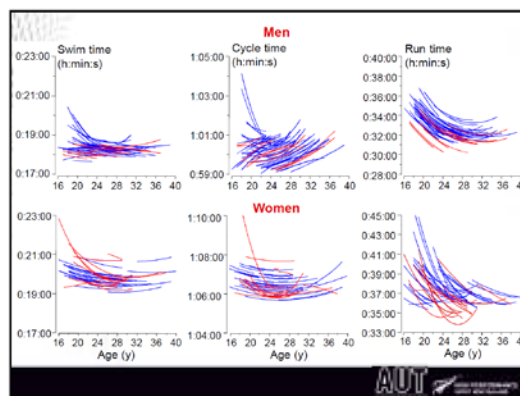
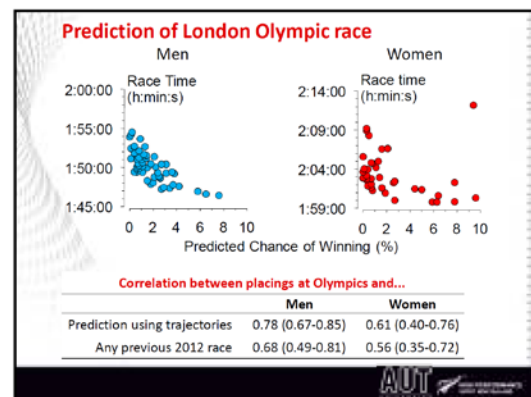
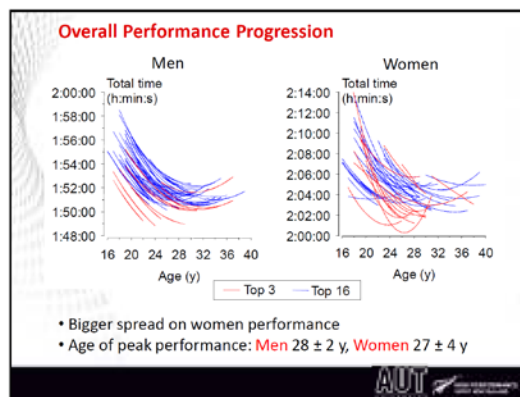
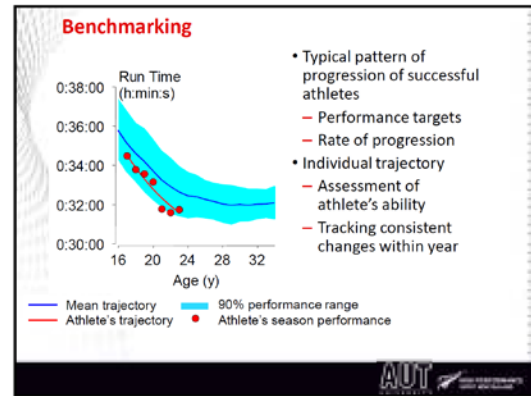
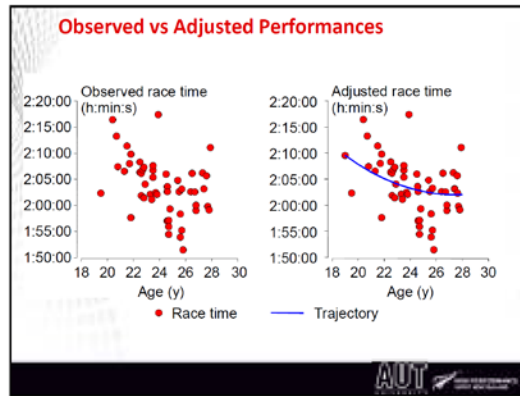
- It's not simply a swim, bike and run!



Model

- Mixed linear model in SAS (Proc Hpmixed)
- Performance times were predicted with the following model:
 - Fixed effects**
 - Mean time for each ability group for sprint- and Olympic-distance
 - Mean quadratic age trend for each ability group
 - Mean linear calendar year trend for each ability group
 - Random effects**
 - Quadratic age trend for each athlete
 - Consistent seasonal performance for each athlete
 - Mean time for each race, accounting for environmental and other course-related factors
 - Different residual errors for each level of competition
- Swim, Bike, Run and Total for men and women times were modelled separately






APPENDIX C

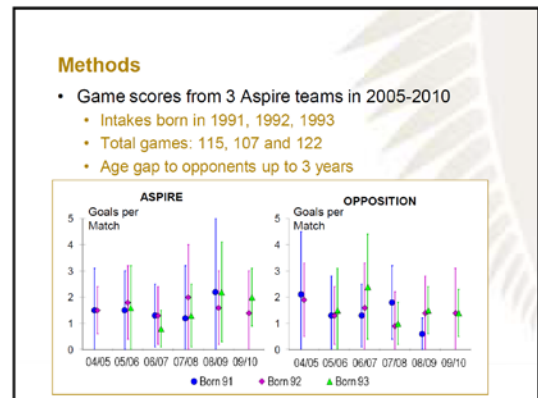
Oral presentation presented at European College of Sport Science, in Liverpool, United Kingdom, 2011.

**MODELLING PROGRESSION OF
COMPETITIVE PERFORMANCE OF AN
ACADEMY'S FOOTBALL TEAMS**

Rita Malcata^{1,2}, Will G Hopkins¹, Scott Richardson³

¹Sport Performance Research Institute, AUT University, Auckland, New Zealand;
²New Zealand Academy of Sport, Auckland, New Zealand;
³Aspire Academy for Sports Excellence, Doha, Qatar





Is the team making
progress?

Methods

- Statistical Analysis System (Version 9.2, SAS Institute, Cary, NC)
- Proc Glimmix
 - Over-dispersed Poisson distribution
 - Fixed effects
 - Mean score; home-ground; age difference
 - Random effects
 - Ability of Aspire vs opposition
- Mechanistic magnitude based inferences (Hopkins, 2010)

Rationale

- Low scores
- Season structure
- Number of games against same opponent
- Quality of opponents

Aim

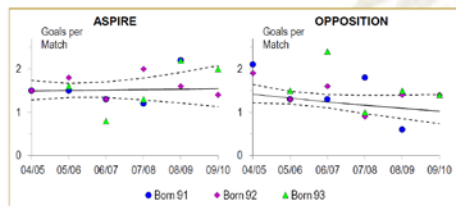
- Tracking team performance using game scores: Aspire vs opponents

Results/ Discussion

- Effects were only clear when the three teams were grouped
- Age Difference: Small effect
 - Team one year older: 27% (90%CI 13 to 42%)
 - Difference in performance ↔ difference physical development stage (Helsen, Van Winckel, & Williams, 2005)
- Home-ground factor: Small effect
 - Aspire: 16% (-5 to 41%)
 - Opposition: 31% (7 to 60%)
 - Maher (1982) attributed the effect to decline in opponent's attacking ability

Results/ Discussion

- Team's performance progression



Conclusion

- A method to study the progression of team's performance scores
 - Generalised mixed linear model
 - Is a team making progress?
- Soccer scores are noisy
 - Measures of soccer performance
 - Scoring opportunities
 - Score box possession
 - Team sport
 - Rugby
 - Netball
- Evaluation of effectiveness of coach policies

Thanks to:



SPORTS PERFORMANCE
RESEARCH INSTITUTE, NEW ZEALAND
AN INSTITUTE OF AUT UNIVERSITY

And, thank you!

Using Athletes' World Rankings to Assess Performance of Countries

Rita M Malcata, Tom J Vandenbogaerde, Will G Hopkins

High Performance Sport New Zealand, Auckland, New Zealand;
Sport Performance Research Institute of New Zealand, AUT University, Auckland, New Zealand



Introduction

- Sporting performance of countries should be assessed with methods that include athletes not winning medals.
- We present a novel and practical approach for ranking countries based of the athletes' world rankings for the sport of swimming.

Methods

- Annual top-150 rankings in the 16 female and 16 male Olympic swimming events were downloaded (www.swimnews.com) for 1990 through 2011.
- Annual country performance scores were derived:
 1. determine importance value (weighting) of each rank;
 2. assign these values to the athletes;
 3. sum values for each country's athletes.
- Four approaches were investigated to determine the weightings:
 - proportion of athletes on each rank who ever achieved the top rank (Figure 1);
 - proportion of athletes on each rank who ever achieved any top-three rank;
 - proportion of athletes on each rank who ever won a gold medal;
 - inverse-rank function (Figure 1).
- The proportions were modeled using logistic regression via Proc Genmod in the Statistical Analysis System (Version 9.2, SAS Institute, Cary, NC).
- Validity of country performance scores was investigated as the correlation between log-transformed scores and number of medals won at Olympic Games and World Champions over the 22-year period.
- Reliability of country performance scores was evaluated for the years 2008-2011 as the intraclass correlation coefficient (ICC) and as the typical within-country year-to-year variation (expressed as a coefficient of variation).

Results

- The modeled proportion of athletes who ever achieved top rank was closely approximated by the much simpler inverse-rank function (Figure 1).
- The inverse-rank approach was used to derive country performance scores and rankings (Table 1 and Figure 2).
- The mean correlation between country scores and medal counts over the 22 years was very large: 0.82 ± 0.06 (mean \pm SD) for scores derived with the inverse-rank approach.
- The reliability of performance scores derived with inverse-rank approach was very high: ICC of 0.96 (90% confidence limits, 0.94-0.97) and coefficient of variation of 61% (55%-68%).
- Scores derived for each swimming event individually had ICC of 0.86 ± 0.05 (mean \pm SD) and coefficients of variation of $75\% \pm 18\%$.
- The other three approaches had similar validity and reliability.

Figure 1. Observed proportions of athletes on a given world rank who ever achieved top rank in the 22-year period. The two curves (almost coincident) are the best fitting logistic regression and the simpler inverse-rank function. Proportions predicted by the curves were used to derive country scores.

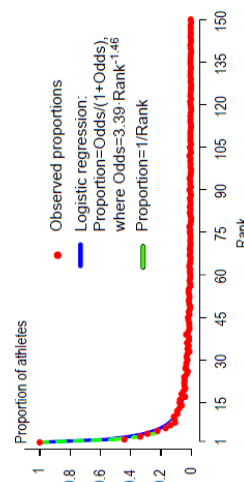
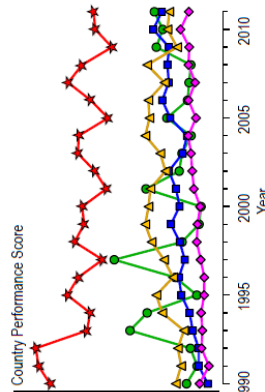


Table 1. Performance scores and the resulting rankings for the 20 countries with the highest scores in 2011.

Country	USA	CHN	JPN	AUS	GBR	FRA	ITA	BRA	CAN	RUS	GER	NED	HUN	SWE	ESP	DEN	POL	NZL	NOR	
Score	41	21	18	16	9.6	7.8	6.5	6.0	5.9	5.9	5.4	5.2	4.3	4.1	2.4	2.4	2.1	1.8	1.2	1.1
Ranking	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

Figure 2. Progression of countries' performance scores from 1990 to 2011, for top five countries in 2011: USA (★), China (●), Japan (■), Australia (▲), Great Britain (◆).



Conclusions

- A score obtained by adding the inverse of the ranks of athletes provides a valid, reliable, practical measure of country performance.
- The value of the inverse rank has two practical interpretations, e.g.:
 - an athlete ranked 17 has a 1-in-17 chance of ever achieving top rank;
 - this athlete's performance is worth 1/17th of a top-ranked athlete for assessing the performance of their country.
- The new measure of country performance is more comprehensive than scores based on medals.
- The measure will be useful in research and applied settings for any sports with seasonal or annual world rankings.

APPENDIX E

Here, I present a section of the report provided to Triathlon NZ to guide the use of an excel-based application for analysis of triathletes running performance. The findings were derived using similar methods as the one in Chapter 3. Sections on the performance trends and predictions for Rio were deleted at the request of Triathlon NZ.

TRIATHLON TRACKING TO 2016: RUN PERFORMANCE

Simon Pearson, Rita Malcata

Overview

This document presents the progression of run performance in triathlon from two perspectives: an overview of elite level performance trends as a whole and also examining individual athlete development. We initially chose to analyse the run performance in isolation as it has the highest correlation with the overall performance (table below), however similar performance trajectories can be developed for the swim and bike (although the effect of pack riding may well confound the usefulness of this).

Table 1. Mean correlations (with 90% confidence limits) between performance in each individual leg and overall race performance

Swim	Bike	Run
0.52 (0.28-0.70)	0.77 (0.53-0.90)	0.87 (0.74-0.94)

Findings presented in this document are based on analysis of run split data from 449 international triathlon races (226 men, 223 women) from 2000 to 2012. These were comprised of World Cup, World (Championship) Series, World Championships (including Junior and U23 races) and Olympic level races.

Because of the confounding nature of triathlon courses on run times (some courses and/or conditions leading to faster or slower run times), these have been corrected to a “standard” time scale by using a combination of average time and strength of field to adjust out (or at least minimise) the effect of course characteristics on run times. Results are currently presented as corrected 10km run times, however could be adjusted to other units (e.g. speed, 5km time) as needed.

Run performance progression (Individual)

A second area of analysis is how individual NZ athletes are tracking in terms of their run performance, in which there are two key elements for consideration:

1. Current gap to the podium and projected rate of progression required to podium in 2016 (or another pinnacle event), as summarised in the year progression above.

2. Age-related development: what is an athlete's current rate of run performance progression, and how much more can they be expected to improve, based on retrospective data. This is particularly relevant for development athletes, as it looks to address whether they are “on track” based on their current age, and current rate of progression.

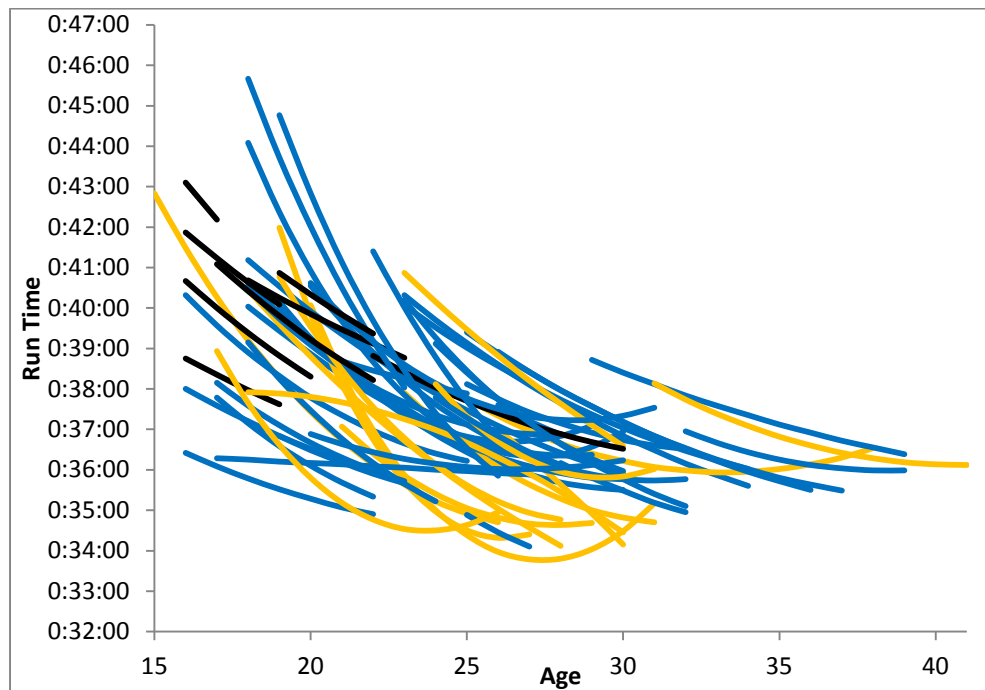


Figure 1. Individual trajectories for the top 3 (yellow), top 16 (blue) and NZ athletes who do not classified in the previous groups (black).

This second point has been addressed by analysing the age-related run performance progressions of any athlete who has had a top 16 finish at a pinnacle event in the last Olympic cycle, starting from and including the Beijing 2008 Olympics. These criteria were selected as it enabled us to benchmark against a reasonably large group of recent elite performers and create a “typical” rate of age-related progression in run performance for those that make it to a high level of international success. While it is acknowledged that 16th can be a long way from a medal performance, this was considered to be the best approach based on:

- Relatively small difference between the average run progression values for top 3 vs 10 vs 16.
- Much better certainty around the model with greater numbers (top 3 created a very noisy model)
- A certain alignment with current funding criteria (medals being the primary aim but top 16 still acknowledged as a notable level of international performance).
- The perspective that because of individual characteristics there can always be exceptions and as such this type of benchmarking should be used as a filter rather than an absolute criterion in athlete evaluation. Therefore a slightly more general model (with better certainty) still fulfils this need.

Below (Figure 2) are a couple of examples using the models for men as a visual assessment tool of how an athlete is progressing based on age. In both cases these examples compare an established athlete who has medalled at a pinnacle event with a developing athlete.

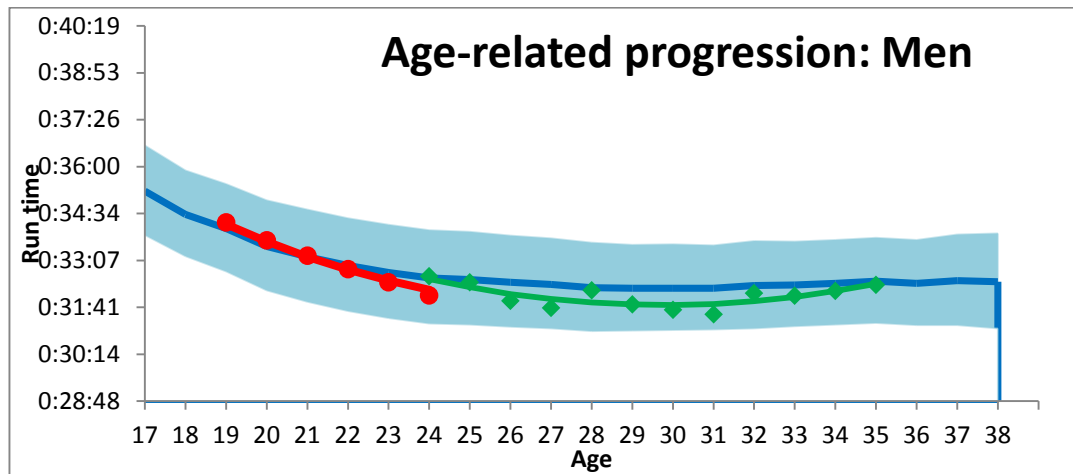


Figure 2. Plotting of individual athletes' progression in run performance against each other (red, green lines) and the typical progression (mean, 90% range) of a pinnacle top 16 athlete (blue line and shading).

An additional way in which this data can be used is in getting an idea of “check points” in term of results. Table 2 outlines the average age at which a (future) pinnacle medallist attained notable finishing positions in lower level races. While the World (Championship) Series data is not particularly useful at this stage, what this shows is that typically a successful athlete will have achieved a World Cup podium by around the age of 23 (men) or 25 (women) and their first Pinnacle event podium will occur around 3 years later. (Note that there are expected levels of variation around these values, this is just providing a basic snapshot).

Table 2. Average ages at which a pinnacle event medallist obtained their first top 16 and top 3 results at different levels of racing.

	Pinnacle	World Series*	World Cup	
	Top 3 / 16	Top 3 / 16	Top 3 / 16	Peak Age
Men	25.6 / 25.0	26.9 / 26.9	23.0 / 22.6	28
Women	27.9 / 27.9	27.8 / 27.8	25.1 / 24.9	27

*Note: World Series results will be skewed due to only being in existence for the last four year.

APPENDIX F

Initial section of report presented to Swimming NZL and board of HPSNZ for assessing performance of New Zealand swimming team at the 2012 London Olympic Games. The remaining section of the report was not relevant for this appendix because they included work not performed by me.

LONDON OLYMPICS 2012: NZ SWIM PERFORMANCE REPORT

Tom Vandenberghe, Sian Allen, Rita Malcata, Will Hopkins

High Performance Sport New Zealand, Auckland, New Zealand and Sport Performance
Research Institute of New Zealand, AUT University, Auckland, New Zealand

We first present an analysis of the performance of New Zealand in Swimming, using a novel method developed by Will Hopkins, Rita Malcata and Tom Vandenberghe. In many sports such as Swimming, medal count at major competitions is the usual measure of country performance. This approach has several problems. First, a medal count does not reflect a country's talent base, because it excludes performances of athletes not winning medals. Secondly, the count is biased against countries with more talent when (as is usually the case) there is a cap on the number of entries from each country. Finally, medal counts are low and therefore inherently imprecise: in any one year they provide only an approximate assessment of performance. The novel method, which solves these problems, combines world rankings of individual athletes into a country score by summing the inverse of the athletes' ranks. The resulting score is equivalent to the top-rank or gold-medal capability of the country. A paper on this method has been submitted for publication in *Medicine and Science in Sports and Exercise*. Please contact Rita or Tom for more info.

We then show performance times of the New Zealand Olympic Team swimmers at the London Olympics and other major competitions 2010 through 2012. We've compared performance progression rates between trials and Olympics, personal best and Olympics, and heats vs semi-finals vs finals, for New Zealand, Great Britain, Australia, USA, China and Japan. We've also included comparisons in number and percentage of total number of performances in individual events that had improved between competitions and from heats to semi-finals and/or finals at the Olympics.

We then present performance trajectories of our London Olympic Team swimmers that qualified for an individual event, and trajectories of medal winners and of the top 16 in the respective event. These modelled trajectories include only best performances each year and do sometimes not reflect performance at the major competition. Nevertheless, we believe these trajectories are a useful tool to track performance progression. We also provide some statistics on the mean age of medalists, finalist and semi-finalists, and on times required to win medals, make finals and make semi-finals.

Finally, we report some general comments and observations, and we've included additional figures and tables in addenda.

Performance of New Zealand in Swimming

Table 1 presents our top-rank performance scores, rankings and medal counts for countries ranked in the top 20 in 2012. The list includes countries not winning medals at the London Olympics (e.g., Italy, Germany and New Zealand) and excludes some countries winning medals (e.g., Tunisia and Lithuania, each winning one medal, and Belarus, winning two medals).

Table 1. Top-rank performance scores, rankings and medal counts for countries ranked in the top 20 in 2012. Top-rank scores are the sum of the inverse of the world ranks for each country's swimmers represented in the top 150 for all 16 female and 16 male pool swimming events.

Country	Top-rank Score	Score Ranking	Medal count	Medal Ranking
USA	39.6	1	30	1
CHN	16.6	2	10	3
AUS	16.3	3	10	3
JPN	13.4	4	11	2
GBR	11.2	5	3	8
FRA	10.2	6	7	5
RUS	6.3	7	4	6
CAN	6.0	8	2	10
ITA	5.2	9	.	.
GER	5.2	10	.	.
HUN	5.2	11	2	10
NED	4.9	12	4	6
RSA	4.7	13	3	8
BRA	4.2	14	2	10
ESP	3.1	15	2	10
SWE	3.1	16	.	.
DEN	2.3	17	.	.
POL	1.7	18	.	.
KOR	1.7	19	2	10
NZL	1.6	20	.	.

Figure1 represents the progression in performance scores and rankings since 1990 for New Zealand and the four countries with the highest scores in 2012 based on our inverse-rank method. The USA was always ranked first over this period. China's performance fluctuated between ranks of 10 and 2. Japan showed an upward trend between 1990 and 1998; its performance stabilized afterwards (rank 2, 3 or 4). Figure 2 has been included to show the ranking of New Zealand on a more appropriate scale. Our country has been ranked between 14th and

28th over the last 23 years. It was ranked 20 in 2008, 21 in 2009, 19 in 2010, 19 in 2011 and 20 in 2012.

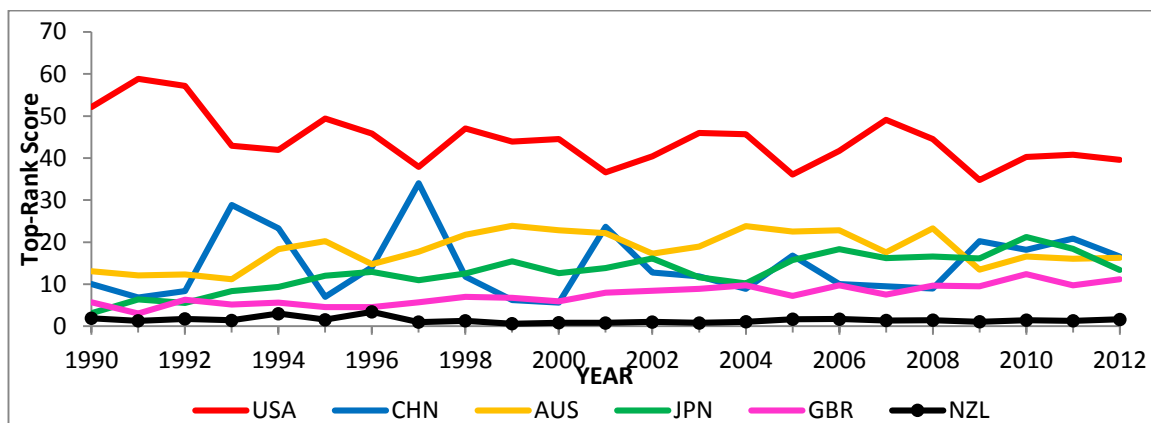


Figure 1. Progression of country scores from 1990 to 2012 for New Zealand and the top-five ranked countries in 2012, USA, China, Australia, Japan and Great Britain, by summing the inverse of the world ranks for all the swimmers represented in the top 150 for all of the 16 female and 16 male pool swimming events.

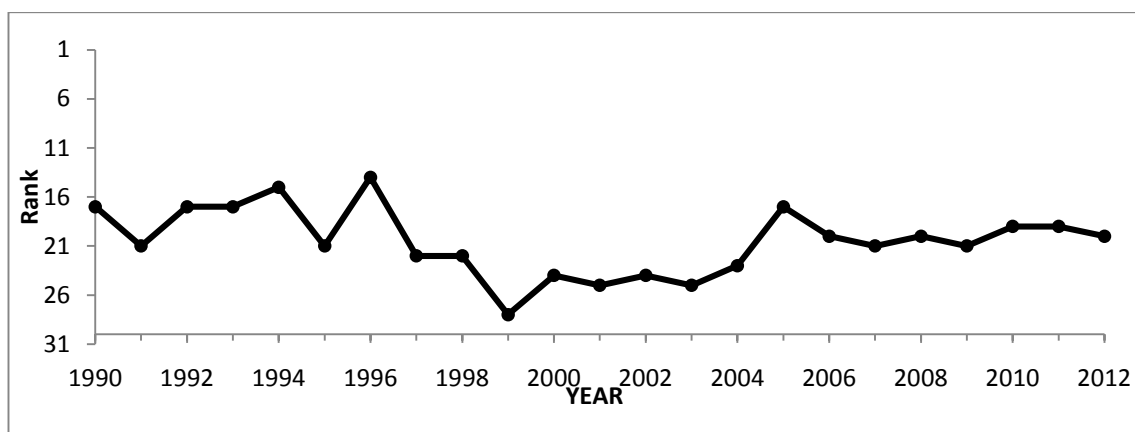


Figure 2. Progression of New Zealand ranking from 1990 to 2012, according to the performance score calculated by summing the inverse of the world ranks for all the 16 female and 16 male pool swimming events.

Appendix G

EXCEL APPLICATION TO ASSESS PROGRESSION OF PERFORMANCE OVER AN OLYMPIC CYCLE

Chapter 7 and its excel application are under embargo.