Monitoring Performance of Elite Rowers

Tiaki Brett Smith

A thesis submitted to

Auckland University of Technology

in fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

July 2011

School of Sport and Recreation

Primary Supervisor: Will G Hopkins

Table of Contents

LIST OF FIGURES

LIST OF TABLES

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 nor material which to a substantial extent has been accepted for the award of any other degree or diploma of a university or other institution of higher learning, except where due acknowledgement is made.

T Brett Smith

Date: 22nd July 2011

PUBLICATIONS, REPORTS AND CONFERENCE PRESENTATIONS ARISING FROM THIS DOCTORAL THESIS

Peer-Reviewed Published Articles

Smith TB, Hopkins WG. Variability and predictability of finals times of elite rowers. Medicine and Science in Sports and Exercise. 2011; 11: 2155-60. (Author contribution percentages: TS, 80%; WH, 20%)

Smith TB, Hopkins WG, Lowe TE. Are there useful physiological or psychological markers for monitoring overload training in elite rowers? International Journal of Sport Physiology and Performance. 2011; 6: 469-84. (TS, 80%; WH, 10%; TL 10%)

Smith TB, Hopkins WG. Heart-rate variability and psychological stress in an elite female rower who developed overtraining syndrome: case report. New Zealand Journal of Sports Medicine. 2011; 38: 18-20. (TS, 90%; WH, 10%)

Smith TB, Hopkins WG. Measures of rowing performance. Sports Medicine 2012; 42: 343-358. (TS, 80%; WH, 20%).

Article Submitted to Peer-Reviewed Journal

Smith TB, Pope CC, Penney, D. Managing athletes on the edge: overtraining, and the complexities of coaches' decision making. Qualitative Research in Sport Exercise and Health 2012; resubmitted April. (TS, 80%; CP, 10%; DP, 10%).

Conference Presentations

Smith TB, Hopkins WG, Lowe TE. Salivary steroid hormones as markers of performance and overreaching in elite rowers. Proceedings of the New Zealand Conference of Sports Medicine, Dunedin, New Zealand. 2008. (TS, 80%; WH, 10%; TL 10%)

Smith TB, Hopkins WG, Lowe TE. Effects of overload training on physiology, psychology and performance of elite rowers. Proceedings of the New Zealand Conference of Sports Medicine, Rotorua, New Zealand. 2009. This presentation

won the award for best emerging researcher at the 2009 SESNZ Conference. (TS, 85%; WH, 10%; TL 5%)

Smith TB, Hopkins WG, Lowe TE. Are there useful physiological or psychological markers for monitoring overload training in elite rowers? Proceedings of the American College of Sports Medicine in Medicine and Science in Sports and Exercise, Indianapolis, USA. 2010. (TS, 90%; WH, 10%)

Smith TB, Hopkins WG. Variability and predictability of finals times of elite rowers? Proceedings of the European College of Sport Sciences, Liverpool, UK. 2011. (TS, 90%; WH, 10%)

Report

Lowe TE, Smith TB, Hopkins WG. Salivary hormone and blood markers of performance in elite rowers. (The Horticulture and Food Research Institute of New Zealand Ltd, Client Report No. 22183 and Contract No. 21618). June 2007. (TL 50%, TS, 40%; WH, 10%)

All co-authors of the articles, presentations and reports indicated above have approved these for inclusion in Tiaki Brett Smith's doctoral thesis.

Brett Smith   Will Hopkins PhD   Tim Lowe PhD   Clive Pope PhD   Dawn Penney PhD

ACKNOWLEDGEMENTS

xiv

In a true show of faith, Dick allowed me to undertake experiments with some of the top rowers in the world as they prepared for the Beijing Olympics. Furthermore, I must thank Dick for including me as full member of his team at all the International events, including the various world championships and the last two Olympics. I am truly indebted to Dick and despite my numerous academic qualifications I have learnt more about elite sport from him than anyone else. I would also like to thank Andrew Matheson (now manager of the Australian rowing program) for refusing to accept my resignation and facilitating this project. Thanks also go to Dr Richard Young, the head of research and technology at SPARC, who provided funding and support for all my research projects. I am truly grateful to all the rowers and coaches of the New Zealand rowing team who volunteered to participate in my various research projects and made this PhD possible. Finally, special thanks to my subjects who were members of the New Zealand 2007 rowing world championship and 2008 Olympic teams, they were truly amazing times.

ABSTRACT

This thesis represents five studies of measurement and monitoring of elite rowers during periods of intensive training and competition. The first study established the smallest worthwhile effect and the variability of competition performance of elite rowers. Study 2 is a review of the literature examining measures of rowing performance and focusing on the errors in these measures using the yardsticks established in the first study. Studies 3 and 4 examined the relationship between changes in performance and changes in various markers during a period of training in a group of elite rowers. The final study focused on how three successful elite rowing coaches monitored their rowers so as to reduce the risk of overtraining.

In Study 1 the finals times for ten men's and seven women's single and crewed boat classes in world-class regattas from 1999 to 2009 were analysed using a linear mixed model. Differences in the effects of venue and of environmental conditions, estimated as variability in mean race time between venues and finals, were extremely large (~3.0%). The race-to-race variability of boat times (~1.0%) and the smallest worthwhile effect (~0.3%) determined from Study 1 were used as yardsticks to determine the accuracy of the various measures of rowing performance examined in the review of literature (Study 2). It was determined that the measurement of boat speed, especially with a good GPS device, has adequate precision for monitoring performance if wind and water current remain consistent. The off-water measure that has error low enough to track an individual's change in physiological performance is the 2000-m time-trial on the Concept II rowing ergometer. Other Concept II measures with acceptably low error for tracking changes include peak power output in an incremental test, some measures of lactate threshold, and measures of 30-s all-out power. Studies 3 and 4 involved a group of elite rowers undertaking a month of overload training followed by a taper. Changes in test performance were compared to changes in various physiological and psychological markers. The changes in many of these markers considered to predict performance maladaptation (e.g., worsening mood, decreasing morning testosterone, increasing inflammation) actually had small to large positive linear relationships with performance. In Study 4 a rower suffered overtraining syndrome but these markers did not show the consistent changes that would have made them useful for predicting overtraining syndrome. In the final study it was found that the strategies used by the coaches to monitor their rowers for overtraining were based

largely on intuition, communication, observation and their subjective analysis of training performance. These strategies had little in common with those promoted in the sport science and medical literature. It could be argued that these differences are the result of the coaches' decision making being primarily based on subjective processes and influenced by various stressors unique to their positions. Future studies should investigate the accuracy of on-water ergometers, the utility of some Concept II test measures, the use of stress markers to prescribe training, and how coaches monitor elite athletes in other sports.

PREFACE

Thesis Rationale

Introduction

In 2000 Richard Tonks was appointed as head coach for the New Zealand Rowing (RNZ) program and brought me in to run the sport science and medicine support program. Richard introduced a very intensive training regime that resulted in many successes but some unexplained disappointments. After the 2006 world rowing championships he requested the exploration of strategies for monitoring the New Zealand rowing team as they prepared for the 2008 Olympic Games. This request became the primary focus of this PhD.

Sport and Recreation New Zealand (SPARC) and the Horticulture and Food Research Institute of New Zealand provided funding for a comprehensive study to examine the relationship between changes in various physiological and psychological markers and changes in performance. This project represented a rare opportunity to examine a range of measures and markers of rowing performance during and following intensive training. The two funding groups required that this project be performed early at the start of the PhD. We undertook this project as a series of two studies. The first study monitored performance and various physiological and psychological markers during the first macrocycle of the 2007 New Zealand Rowing team's training program. Motivation was high as the team was preparing for the world rowing championships, which doubled as the Olympic qualification regatta. It was hoped that the combination of highly motivated elite athletes engaged in intensive training would provide insights into overtraining; indeed it did. The performance measures were undertaken on the Concept II and included a weekly 30-min test during the overload period and an incremental lactate test performed at the beginning of the study and following the taper. The physiological markers chosen included a wide range of salivary and blood markers such as cortisol, testosterone, creatine kinase and various cytokines. A number of psychometric markers such as rating of perceived exertion, positive and negative affect and quality of sleep were also examined. These markers were chosen on the basis of published evidence that they might be useful for predicting overtraining syndrome. The next study was conducted in 2008 and followed a similar methodology to the first study, except the physiological markers were replaced with measures of heart-rate variability. These studies are reported in Chapters 3 and 4.

We next turned our attention to examining the various measures of rowing performance. At this stage there was no information on which measures had the precision to accurately monitor changes. The first stage in answering this question was to determine the variability of performance of elite rowers between international competitions and to derive an estimate of the smallest worthwhile effect there from (Chapter 1). The calculated variability and smallest worthwhile effect were then used as yardsticks to assess the reliability and validity of the various measures of rowing performance reported in the literature (Chapter 2).

During this project I became interested in how the coaches managed the balance between optimal conditioning and overtraining during periods of intensive training. A point of intrigue was how this balance was achieved, given our findings that many of the markers and performance tests in use at that time were not effective monitoring tools. In a first of its kind, a qualitative study of the practices and beliefs for managing elite athletes during periods of intensive training was undertaken with three highly successful professional elite rowing coaches (Chapter 5).

A Review of Overtraining

The following review of literature was written originally to support the detailed proposal for the thesis (the D9) and was included as an appendix. We were requested by the examiners to transfer it to this section of the thesis, since it represents supporting information for the rationale. This section primarily reviews the markers and measurement procedures used in Chapters 3 and 4 and has been updated to address some of the concerns of the examiners.

Intense Training and Overtraining

Elite endurance athlete are required to regularly increase their training workload, and those athletes who can continue to adapt positively to the greater training overload have a greater chance of success in competition.[1, 2] However, the greater the training overload the greater the risk of maladaptation, created primarily by an imbalance between the overload and recovery.[3] The challenge for sport scientists working with elite athletes is to ensure that the hard training required to optimize performance does not cause maladaptation.

Maladaptation during training or competition in elite athletes will manifest itself principally as injuries, illness or overtraining. The focus of this thesis was initially to examine methods for predicting overtraining in elite rowers during periods of hard training. The reason for this interest was the high rates of overtraining reported in Olympic athletes (10-28%),[4] and concerns that the intense training regime employed by the rowing New Zealand coaches could lead to increased incidences of overtraining. Chapters 3 and 4 of this thesis examine various predictors of overtraining.

While overtraining is a much-studied area, there is little consensus over terminology, diagnoses and methods of prediction.[3, 5] In Chapters 3 and 4, I have used the definition presented in the European College of Sport Sciences position statement on the prevention, diagnosis and treatment of the overtraining syndrome.[6] For this definition "overtraining is used as a verb, a process of intensified training with possible outcome of short term overreaching (functional overreaching); extreme overreaching (non-functional overreaching); or overtraining syndrome."[6, p2]. The terms functional overreaching, non-functional overreaching and overtraining syndrome reflect the notion of a fatigue continuum that results in performance changes, which range from temporary decrease, stagnation and decrease. Non-functional overreaching and overtraining syndrome are considered the stage on the fatigue continuum where chronic

performance reduction is combined with maladaptation of various physiological mechanisms. The differences between non-functional overreaching and overtraining syndrome are often subtle and based on a retrospective diagnosis on the period of performance reduction, plus prolonged maladaptation of several biological, neurochemical and hormonal regulation mechanisms.[6]

In Chapter 5 the definitions, interpretations and cues for overtraining described by Meeusen et al.[6] were compared to those used by three elite rowing coaches. An interpretative qualitative methodology was employed to provide an in depth examination of the coaches views on overtraining. Throughout Chapter 5 the words overtrain, overtraining and overtrained were used to denote the process of intensified training with possible outcomes of functional overreaching, non-functional overreaching or overtraining syndrome. This approach was adopted in an attempt to achieve some consistency between the terminology employed by Meeusen et al.[6] and the three coaches.

Predicting Non-Functional Overreaching and Overtraining Syndrome

Various researchers have suggested that a key in the recognition of overtraining syndrome is the assessment of "prolonged maladaptation" not only of the athlete performance, but also of several biological, neurochemical, and hormonal regulation mechanisms.[6-8] Signs and symptoms associated with these maladapted regulatory mechanisms may be grouped into four categories: psychological, physiological, biochemical and immunological.[9] Athletes suffering non-functional overreaching or overtraining syndrome tend to manifest different combinations of these signs and symptoms with varying degrees of severity.[6, 9]

Examination of the causal mechanisms of non-functional overreaching and overtraining syndrome is considered important not only in helping to determine reliable and valid diagnostic tools for predicting these conditions, but also to help identify the specific causalities. The difficulties in effectively monitoring physical conditioning and performance in many endurance sports[10] have focussed researchers on the exploration of physiological and psychological markers to predict non-functional overreaching and overtraining syndrome. Although there are a number of theories to explain non-functional overreaching or overtraining syndrome the search for reliable and valid predictive markers continues.[5, 6, 11]

The various theories of overtraining tend to concentrate on possible pathophysiological changes which include hypothalamic dysfunction;[12] changes in concentration and function of neurotransmitters (amino acid imbalance theory);[13] reduced muscle glycogen;[14] changes in the hypothalamic-pituitary-adrenal axis and pituitary function, and sensitivity to feedback from the periphery;[15] decreased central command to skeletal muscles;[16] changes in autonomic nervous system function, which can have both central or peripheral context[17] and tissue trauma of either the muscle and/or connective tissue and/or bony structures resulting in chronic inflammation.[18] These various theories have spawned a wide variety of potential markers of non-functional overreaching or overtraining syndrome, with O'Toole[3] presenting 84 major markers that are prevalent in the literature. Assessing the efficacy of these various markers in predicting non-functional overreaching and overtraining syndrome is difficult. Urhausen and Kindermann[11] have examined the available research and recommend that the following markers may be suitable for the diagnosis of non-functional overreaching or overtraining syndrome: sport specific performance, ergometric time-trials, mood profile, sleep disorders, rating of perceived exertion, heart rate at rest and during maximal exercise, creatine kinase, testosterone, cortisol, adrenocorticoptrophic hormone and catecholamines.

After examination of the available research, I decided to examine the following markers: cortisol, testosterone, dehydroepiandrosterone (DHEA), creatine kinase (CK), lactate dehydrogenase (LDH), psychological state, sleep quality, rating of perceived exertion (RPE), perception of fatigue, heart rate variability and various cytokines. These markers were chosen because they were likely to be reliable and valid predictors of non-functional overreaching or overtraining syndrome. Furthermore, the tools were either available or being developed so that these markers could be analysed quickly enough to modify the next training session if required.

Performance Measures

In Chapter 3 and 4 the performance measures chosen were a stepwise lactate threshold test and a 30-min maximal rowing ergometer test for which the rating was restricted to 18 strokes per minute. The reliability, validity and smallest worthwhile effect for these (and other) rowing tests are covered extensively in Chapters 1-3 of this thesis and further information on the testing protocols are covered in the report in Appendix C.

Hormone Markers

For athletes involved in endurance training, one would typically expect to see relatively unchanged morning hormone levels, unless an imbalance between training load and recovery causes excessive fatigue resulting in a suppressed endocrine response.[19] This excessive fatigue or functional overreaching is typically reversible within days, however if appropriate modifications to the athletes training are not instituted non-functional overreaching or overtraining syndrome may occur.[20] There is considerable support from many studies that endocrine dysfunction is a characteristic of non-functional overreaching or overtraining syndrome.[9, 15, 19, 21-27] The majority of these studies only measured resting hormone concentrations, and when endocrine responses to exercise has been measured, it is claimed to be superior for indicating overtraining syndrome.[21]

Cortisol and testosterone are controlled by the hypothalamus and changes in these hormones reflect an integrated response to stress and training. The use of testosterone and cortisol as an indicator of the anabolic-catabolic balance in rowers has been examined in a number of studies.[28] Steinacker et al.[29] found that in junior rowers morning testosterone decreased over time with large volumes of training and then increased with lower volume, high intensity training. However, further research with very high volumes of training has showed no changes in either resting testosterone or cortisol.[25] Vervoon et al.[19] monitored the Dutch national team for nine months and found that resting testosterone and cortisol were generally unchanged. Maestu et al.[28] found that three weeks of heavy training in junior rowers induced reductions in resting free testosterone while cortisol showed no change.

Examination of the response of testosterone and cortisol in response to overtraining syndrome in other endurance sports has also shown unclear patterns for both hormones. Reported changes in resting cortisol with endurance training are not consistent between studies with reports of levels going both up and down. For example, male cyclists during the Tour of Spain showed a decrease in cortisol concentrations after only one week of the tour and they continued to decrease over the remainder of the tour,[30] the exercise consisted of 21 consecutive daily stages with only one rest day. It is very difficult to directly compare research in different endurance sports due to the mode of the activity and the various different training regimes eliciting a wide variation in tissue trauma which in turn alters various endocrine responses.

In Chapter 3 the steroid hormones cortisol, testosterone and dehydroepiandrosterone (DHEA) were measured from saliva samples. The fate of cortisol, testosterone and DHEA is complex, and in the blood there is the added complication that these steroids are bound to carrier proteins. Aside from the non-invasive nature of saliva collection compared with blood collection, saliva levels of these hormone are thought to reflect their "free" or "bioactive fraction",[31, 32] thus they afford a more sensitive and relevant measure than blood. Dehydroepiandrosterone was included, as it is the main precursor for testosterone in females and it is also technically easier to measure than testosterone. The responses of these hormones to exercise were also assessed by measuring the change pre and post a weekly 30-min exercise test.

Immune Markers

There is growing support for suppression of the immune system in response to overtraining, and this suppression is reported to increase susceptibility to infection and disease[33]. The most commonly proposed model is that moderate exercise enhances the immune system, but with very high levels of exercise especially intensive endurance training, the immune system becomes compromised, increasing the incidences of recurring infections.[34, 35]

In 2000 Smith[36] proposed that specific cytokines may predict impending non-functional overreaching or overtraining syndrome. Smith[36, p317] argued that intense training with insufficient rest will cause anatomical trauma from which "circulating monocytes are then activated by injury-related cytokines, and in turn produce large quantities of proinflammatory interluekin-1 beta (IL-1β), and/or interluekin-6 (IL-6), and/or tumor necrosis factor-alpha (TNF-α), producing systemic inflammation". Furthermore, the elevated circulating cytokines coordinates a response that elicits mood and behaviour changes that seek to reduce athletes' participation in intensive training as a means of resolving the inflammation and maladapted immune function.

In later research Smith[18, 37] added to this hypothesis by suggesting that anatomical trauma (specifically tissue trauma) created by excessive stress modulates the immune system, and the associated cytokine release affects the hypothalamic control of hormones such as cortisol and testosterone. Steinacker et al.[26] suggested that the activity of specific cytokines is also altered by glycogen depletion associated with severe fatigue. They proposed that the altered cytokine activity associated with

glycogen depletion and excessive tissue trauma impairs the hypothalamic regulatory mechanisms resulting in non-functional overreaching or overtraining syndrome.

In this thesis a range of pro-inflammatory, anti-inflammatory and regulatory cytokines in both blood plasma and saliva were monitored in elite world-class rowers during an intensive phase of training to examine any possible relationships to non-functional overreaching or overtraining syndrome. As there was little research published on the specific cytokines that could predict impending non-functional overreaching or overtraining syndrome, 14 cyokines were chosen that my supervisors and I believed would provide insights into any maladaptations in immune function caused by excessive rowing training. These cytokines were interferon-alpha (IFN-α) interferon-gamma (IFN-γ), interleukin-1b (IL-1b), interleukin-2 (IL-2), interleukin-4 (IL-4), interleukin-5 (IL-5), interleukin-6 (IL-6), interleukin-8 (IL-8), interleukin-10 (IL-10), interleukin-12p70 (IL-12p70), interleukin-18 (IL-18), monocyte chemotactic protein-1 (MCP-1), tumor necrosis factor-alpha (TNF-α), and tumor necrosis factor-beta (TNF-β).

The detailed methods employed for the assay of the cytokines and C reactive protein were not included in the technical report for SPARC, HORT Research and Rowing NZ (see Appendix B), so I have include it here. The cytokines were determined simultaneously using a multiplexing bead assay (FlowCytomix, Bender MedSystems, Vienna, Austria). This assay was performed using a 96 well filter-plate following manufacturer's instructions. Samples and reagents were handled by a liquid handling workstation (Biomek 3000, Beckman Coulter, CA) and fluorescent intensities of beads were measured on a FC500 MPL flow cytometer (Beckman Coulter, FL) (excitation at 488 nm and 633 nm, fluorescent signals recorded at 575 nm and 675 nm). Multiplexing is achieved by using up to 20 different fluorescent beads, each with a different internal dye intensity and a unique antibody. In total, data from 12000 beads were collected (800 beads per cytokine). The median fluorescence intensity (MFI) of standard cytokine/inflammatory markers was used to derive the calibration curve using a five parameter logistic model. For saliva duplicate analysis was performed on 3 samples for each biomarker and the mean intra assay coefficient of variation was 8.7%.

Markers of Muscle Damage

Exercise induced muscle damage is often promoted as a likely cause of reduced performance capacity related to non-functional overreaching or overtraining syndrome,

which has lead to the hypothesis that markers of tissue trauma may be useful indicators of non-functional overreaching and overtraining syndrome.[38] A proposed indicator of muscle damage of athletes involved in intense training is an elevation of muscle proteins namely myoglobin, creatine kinase or lactate dehydrogenase.[38] Examination of the relationship between these muscle proteins and non-functional overreaching or overtraining syndrome has primarily focussed on creatine kinase.

The measurement of creatine kinase to diagnose non-functional overreaching and overtraining syndrome has been used extensively despite a paucity of evidence[5, 25, 39]. In a study on junior rowers it was determined that on any given day, the creatine kinase concentrations were dependent on the training load of the previous day or two[40] as this enzyme typically takes 24 hours to peak after an exercise event or muscle damage.[41] A major problem with creatine kinase is that for non impact sports creatine kinase may not rise in spite of severe fatigue[8] and there have been incidences of non-functional overreaching or overtraining syndrome without elevated creatine kinase.[42]

Creatine kinase and lactate dehydrogenase were measured during this study to examine whether any relationship exists between change in muscle damage and change in performance. If muscle damage was to occur it was more likely in these subjects who were all highly motivated, hard training world-class performers.

Psychological State, Quality of Sleep and Rating of Perceived Exertion

An impaired mood state, subjective complaints, sleep disturbance and perceptions of fatigue are consistently described as sensitive and early signs of non-functional overreaching and overtraining syndrome.[11, 43, 44] The relationship between mood and overtraining has been examined in a number of studies using Profile of Mood States inventory (POMS).[45] POMS involve 65 self report adjectives designed to measure six dimensions of mood: tension-anxiety; depression-dejection; anger-hostility; vigor-activity; fatigue-inertia; and confusion-bewilderment.[46] While athletes suffering non-functional overreaching or overtraining syndrome often have "higher scores for total mood disturbance, depression, tension, and decreased vigor" POMS doesn't predict these conditions "but does provide a validated method for documenting mood changes consistent with the condition".[47, p26]

POMS was developed for the measurement of mood in clinical populations, hence five of its six scales measure negative affect resulting in a focus on stress-related behaviour, potentially reducing its appropriateness in evaluating excessive fatigue in

rowers.[48] For this reason and concerns over the length of POMS,[49, 50] the positive and negative affect score (PANAS) was used to assess changes in mood. PANAS consists of a total of twenty items, one 10-item mood scale for measuring positive affect and another 10-item scale for measuring negative affect.[51, 52] PANAS is "brief, easy to administer, highly internally consistent, largely uncorrelated, and stable at appropriate levels over a 2-month time period".[52, p1067]

In practice, the usefulness of any subjective parameters is questionable especially since deterioration in mood often tracks increased training load, which is a normal requirement of an effective training program.[11] Coaches within the Rowing New Zealand elite program have also raised concerns that that these subjective measures have no reference values and could potentially be used by certain athletes to modify training load for reasons unrelated to maladapted physical conditioning.

The most common tool for estimating the intensity of training is the rating of perceived exertion 10 or 20 point scale developed by Gunnar Borg.[53] Little scientific evidence exists examining the relationship between rating of perceived exertion (RPE) and non-functional overreaching or overtraining syndrome. However it is often suggested that RPE is a useful tool for monitoring conditioning[54] which has lead to its promotion as a monitor of non-functional overreaching or overtraining syndrome.[11]

Sleep disturbance is often cited as an indicator of non-functional overreaching or overtraining syndrome,[33, 55] but similar to the above measures there is little evidence to what specifically defines normal versus disturbed sleep. This study monitored mood, rating of perceived exertion, fatigue and sleep patterns in elite world class rowers during an intensive phase of training examining any possible relationships to non-functional overreaching or overtraining syndrome.

Heart Rate Variability

During normal functioning of the heart there are regular changes in heart rate which are primarily due to changing levels of sympathetic and parasympathetic control of the heart. These fluctuations are caused by the autonomic nervous system, which is constantly monitoring the body's internal environment and modifying its output in an attempt to maintain an optimal physiological state. Therefore, at rest a healthy individual has a continuously varying heart rate that has been termed heart rate variability, and it is believed to be associated with genetic factors, oscillatory

fluctuations in blood pressure and frequency oscillations due to thermal regulation and respiration.[56]

Autonomic nervous system dysfunction, specifically imbalances between the sympathetic and parasympathetic nervous systems, have been associated with overtraining syndrome.[57, 58] This has led to the suggestion that heart rate variability will predict non-functional overreaching and overtraining syndrome.[57, 59-61] While there is a belief among various researchers that autonomic dysfunction associated with non-functional overreaching or overtraining syndrome will manifest itself in substantial changes in heart rate variability, results to date are equivocal.[62, 63] In Chapter 4, heart rate variability was examined in elite world class rowers during an intensive phase of training, to determine if any possible relationship exists between heart rate variability and non-functional overreaching or overtraining syndrome.

Heart rate variability can be quantified using time-domain or frequency-domain methods. The easiest are the time-domain methods, which are computed using simple statistical procedures often from short collection periods.[63] There are a number of different time-domain methods, but since many of these measures correlate highly, the three recommended methods are: standard deviation of the NN interval (SDNN); the square root of the mean squared differences of successive NN intervals (RMSSD); and HRV triangular index which is a geometric method where the integral of the density distribution (i.e. the number of all NN intervals) is divided by the maximum of the density distribution.[56] Another commonly used time-domain method in studies with athletes is the proportion of interval differences of successive NN intervals greater than 50 ms divided by the total number of NN intervals (pNN50).[62, 63]

The frequency-domain methods record heart rate variability as harmonic oscillations calculated by parametric or non-parametric power spectral-density analysis. Three main spectral components consist of a high frequency (HF) region (0.15–0.50 Hz), low frequency (LF) region (0.04-0.15 Hz) and very low frequency region (0.003-0.04 Hz).[56] It has been proposed that the high frequency region is mediated solely by the parasympathetic system, while the low frequency region is mediated by both parasympathetic and sympathetic nervous systems.[56] A common frequency domain measure is the LF/HF ratio, which is considered by some investigators to mirror the balance between the sympathetic and parasympathetic system;[56] this measure is potentially useful for predicing non-functional overreaching or overtraining syndrome.[64]

For the study in Chapter 4, the methods discussed above that could calculate heart rate variability from samples lasting five minutes were used. Five-minute samples were chosen, as this is the smallest timeframe to collect meaningful data,[56] and I considered any measure that would be useful for an elite rower would need to be quick enough to fit into their normal schedule. The methods chosen were the LF/HF, SDNN, RMSSD, pNN50; the mean RR interval was also reported as a measure of (the inverse of) the resting heart rate.

Originality of the Thesis

- In a novel approach, linear modelling using inferential statistics with smallest worthwhile effects was used to estimate the magnitude of the change in various markers of non-functional overreaching and overtraining syndrome, against changes in performance.

- This was one of the first studies to investigate whether "salivary" measures of cortisol, testosterone, dehydroepiandrosterone and various cytokines predicted changes in performance of elite endurance athletes during a period of intensive training.

- Contrary to accepted theory, the results of this thesis suggest that for elite endurance athletes, excessive change in one or more stress markers during a period of intensive training are useful markers of positive adaptation rather than non-functional overreaching or overtraining syndrome.

- No previous study has investigated the magnitude of the smallest worthwhile enhancement in performance for elite rowers.

- This is the first study to examine predictability of performance of elite rowers, and the effect that environment, venue, boat class and levels of final has on international rowing competitions.

- No previous study has determined the standard error of the estimate and the standard error of measurement of an extensive range of off-water and on-water rowing performance measures.

- This is the first study to explore in-depth the practices and beliefs of elite professional endurance coaches as they manage the risk of athlete overtraining during crucial periods of intensive training.

Thesis Organization

This thesis consists of a preface and six chapters. The references for the preface and each chapter are collated at the end of the thesis. The preface introduces the thesis and contains the literature review for the various psychological and biological markers examined in Chapters 3 and 4. Chapters 1-5 are original investigations. Chapter 1 has been published in Medicine and Science in Sports and Exercise. Chapter 2 is the review of literature for the performance measures and has been accepted for publication in Sports Medicine. Chapter 3 has been published in the International Journal of Sports Physiology and Performance. Chapter 4 has been published in the New Zealand Journal of Sports Medicine. Chapter 5 has been resubmitted to Qualitative Research in Sport, Exercise and Health after replies to the reviewers' comments. Chapter 6 comprises a general summary in which practical applications of this research are discussed.

The appendices are presented in chronological order of their development in the PhD. Appendix A contains the ethics approval, participant information sheet and the participant consent form for the project titled "Are there useful physiological or psychological markers for monitoring?" (Chapter 3). Appendix B contains the weekly recall inventories used to record the positive and negative affect scores (PANAS), quality of sleep, fatigue levels and training times used in Chapter 3. Appendix C is the report presented to Sport and Recreation New Zealand and Rowing New Zealand following the study in Chapter 3. This report was a contractual requirement of the funding received by these two organizations. Appendix D comprises the ethics approval, participant information sheet and the participant consent form for the case report titled "Heart-rate variability and psychological stress in an elite female rower who developed overtraining syndrome" (Chapter 4). Appendix E contains the pages for one day from the diary used to record the PANAS score, quality of sleep, fatigue levels, rating of perceived exertion and training diary used in chapter 4. Appendix F comprises the abstract published in the proceedings of the New Zealand Sport Science and Medicine Conference, Dunedin 2008. Appendix G is the abstract published in the proceedings of the New Zealand Sport Science and Medicine Conference, Rotorua 2009. This presentation won the award for the best emerging researcher from that conference. Appendix H contains the abstract for the American College of Sports Medicine Annual meeting, Indianapolis, 2010. This was delivered as a podium presentation and the abstract was published in the conference proceedings published in Medicine and Science in Sports and Exercise. Appendix I is the abstract for the European College of Sport Science annual conference, Liverpool, 2010. Appendix J

comprises the ethics approval, participation information sheet and the participant consent form for the qualitative project titled "Managing athletes on the edge: overtraining and the complexities of coaches' decision making" (Chapter 5). Appendices K and L comprise the technical reports of the study I undertook to examine the validity of the Nielsen Kellerman XL4 impeller and GPSports Spi Elite GPS during 2000-m rowing regatta competition. Appendix M consists of two tables that report the standard deviation of the 4-week change in the various biomarkers and the technical error of the biomarker assays. Appendix N is an article written about my PhD research published in the International Rowing Federation magazine in 2011.

CHAPTER 1

Variability and Predictability of Finals Times of Elite Rowers

Running title: Variability and predictability in rowing

Overview

Purpose: Little is known about the competitive performance characteristics of elite rowers. We report here analyses of performance times for finalists in world-class regattas from 1999 to 2009. **Methods:** The data were official race times for the 10 men's and seven women's single and crewed boat classes, each with ~200-300 different boats competing in 1-33 of the 46 regattas at 18 venues. A linear mixed model of race times for each boat class provided estimates of variability as coefficients of variation after adjustment for means of calendar year, level of competition (Olympics, world championship, world cup), venue, and level of final (A, B, C…). **Results:** Mean performance was substantially slower between consecutive levels of competition (1.5%, 2.7%) and consecutive levels of finals (~1-2%). Differences in the effects of venue and of environmental conditions, estimated as variability in mean race time between venues and finals, were extremely large (~3.0%). Within-boat race-to-race variability for A finalists was 1.1% for single sculls and 0.9% for crewed boats, with little difference between men and women and only a small increase in lower-level finalists. Predictability of performance, expressed as intraclass correlation coefficients, showed considerable differences between boat classes, but the mean was high (~0.63), with little difference between crewed and single boats, men vs women, and within vs between years. **Conclusion:** The race-to-race variability of boat times of ~1.0% is similar to that in comparable endurance sports performed against water or air resistance. Estimates of the smallest important performance enhancement (~0.3%) and the effects of level of competition, level of final, venue, environment, and boat class will help inform investigations of factors affecting elite competitive rowing performance.

Keywords: athlete, rowing, intraclass correlation, reliability

Introduction

The variability of performance of top athletes between competitions has become a topic of interest to sport scientists since its relationship to medal winning was established.[65] For sports in which athletes compete as individuals against other athletes for a best time, distance or other performance score, 0.3 of the standard deviation of a top athlete's race-to-race performance provides an estimate of the smallest worthwhile enhancement in performance that affects medal prospects substantially.[65] Estimates of the variability of performance and the associated smallest worthwhile effects are available for many sports, but rowing is not yet one of them. There are several published studies of variability of performance in sports similar to rowing, where athletes develop maximal sustainable power and overcome water or air resistance, including flat-water kayaking, (0.7-1.5%),[66] swimming, (0.6-1.0%),[67] kilo cycling (1.2%) and 40-km time-trial cycling (1.7%).[68] It is unclear how variability in rowing would compare with these sports, given the technical demands and environmental effects unique to each sport. We have therefore analyzed competitive performance times of single and crewed boats over 11 years of international competition, with the aim of quantifying smallest worthwhile effects and investigating factors effecting variability of performance.

We have also investigated the predictability of the rowers' performance from competition to competition. This concept addresses the issue of the stability of the ranking of athletes, and it has started to appear in published studies of performance variability.[69, 70] The statistic that is being used to quantify predictability is the intraclass correlation coefficient, which is calculated from variability within and between athletes across numerous competitions and is equivalent to the usual correlation that would be observed between performances in two competitions. The intraclass correlation could be useful for identifying sports in which it is easier to predict medalists. To date the published research on the predictability of performance comprises only two studies reporting correlations that ranged from 0.06 to 0.35 for elite skeleton athletes[69] and from 0.05 to 0.61 for elite slalom canoe-kayak athletes.[70]

Methods

Performance Data

There are up to 22 boat classes in international rowing regattas (competitions), consisting of heavyweight or lightweight rowers, male or female rowers, singles or

crewed boats, coxed or coxless boats, and sweep or sculling oars. In sculling, each rower has two oars with boat types involving one rower (single), two rowers (double) or four rowers (quad). In sweeping, each rower has one-oar and boat classes include two rowers without a coxswain (pair), four rowers without coxswain (four) or eight rowers with a coxswain (eight). All races at international regattas are staged on a 2000-m six-lane course.

We analyzed the 14 Olympic boat classes and three non-Olympic boat classes that regularly included at least B finals and sufficient subject numbers to permit the same analysis as for the Olympic boat classes. One world championship regatta and 3-4 world cup regattas are held every year. In the Olympic years only the non-Olympic boat classes can compete at the world championships.

Official final times for the finalists in world cups, world championships and Olympics from 1999 to 2009 (46 regattas) were obtained from the International Rowing Federation website (worldrowing.com). See Table 1.1 for a list of the boat classes and sample sizes. The University of Waikato, Faculty of Education Research Ethics Committee declared that informed consent was not required as these data were in the public domain and no individuals were named.

Statistical Analysis

The mixed linear-modeling procedure (Proc Mixed) in the Statistical Analysis System (Version 9.2, SAS Institute, Cary, NC) was used for most analyses. Finals times were log transformed before analysis, since after back transformation this approach yields variability and differences as percents of the mean (coefficients of variation), which is the appropriate method for quantifying changes in this kind of athletic performance.[71] We performed separate analyses for each boat class. The fixed effects were: Final (*A,B,C,D...,* to estimate differences between finals); Complevel (*world cup, world championships, non-Olympic world championships, Olympics,* to estimate differences between levels of competition); and Year (*1999, 2000...,* to estimate differences between years). The random effects were: Boat (the name of the athlete in singles or the concatenated alphabetically ordered names of the athletes in crewed boats, to estimate differences in ability between boats); Boat*Year (to estimate within-boat variation between years); Venue (*Athens, Beijing...,* to estimate variation between venues); Competition*Final (to estimate variation from final to final within competitions, assumed due to environmental factors); and the Residual (to estimate within-boat within-year final-to-final variation). The random-effect solution for Venue

provided estimates of the relative mean times at each venue. Analyses were performed allowing only positive variances (the default in Proc Mixed). Different Boat and Residual variances were estimated for the A finals and the other finals to allow separate estimation of variability and predictability of the top and other competitors.

Table 1.1 Simple statistics for the number of races entered by boats in the various boat classes at world cup, world championships and Olympic regattas from 1999-2009.

Boat class	Number of boats		Entries per boat	
	1 entry	>1 entry	Mean	Max
Singles				
Lightweight men's single scull (LM1X)[a]	120	144	2.6	14
Lightweight women's single scull (LW1X)[a]	95	118	2.8	23
Men's single scull (M1X)	125	147	3.2	33
Women's single scull (W1X)	84	100	3.5	32
Pairs and doubles				
Lightweight men's coxless pair (LM2-)[a]	160	87	1.7	9
Lightweight men's double sculls (LM2X)	134	135	2.6	25
Lightweight women's double sculls (LW2X)	149	142	2.3	16
Men's coxless pair (M2-)	131	120	2.5	20
Women's coxless pair (W2-)	120	129	2.2	25
Men's double sculls (M2X)	129	98	2.5	18
Women's double scull (W2X)	151	83	2.1	20
Fours and quads				
Lightweight men's coxless four (LM4-)	188	126	1.9	10
Men's coxless four (M4-)	204	100	1.9	15
Men's quadruple sculls (M4X)	159	114	1.9	13
Women's quadruple sculls (W4X)	118	73	1.7	10
Eights				
Men's coxed eight (M8+)	202	67	1.4	10
Women's coxed eight (W8+)	130	52	1.4	5

[a]Non-Olympic boat classes.

Plots of residual versus predicted values from the analyses showed no evidence of non-uniformity of error. A total of 23 finals times spread across eight boat classes had standardized residuals >4.5, representing unusually slow times; only one belonged to a female boat and nine occurred in A finals. We believe these outliers signal incidences of either equipment failure or rowers "throwing in the towel"; these times were therefore deleted before re-analysis.

Thresholds for interpreting magnitudes of differences in mean performance times as being small, moderate, large, very large and extremely large were respectively 0.3, 0.9, 1.6, 2.5, and 4.0 of the within-boat race-to-race (residual) coefficient of variation of the A finalists in each boat class; these thresholds represent enhancements that would provide a top athlete with an extra 1, 3, 5, 7 and 9 medals in every 10 competitions[71]. To interpret the magnitude of a coefficient of variation (other than the residual) representing typical differences or changes in performance times, we doubled the coefficient of variation before assessing it on the above scale, in the same manner that the effect of a linear covariate should be considered as the effect of twice its SD.[71] A further justification for choosing a value larger than the coefficient of variation itself is that the various practical ways of considering differences between normally distributed values are all larger than one SD: the SD of the difference between two randomly chosen values is 1.4 SD; the difference between the mean of the lower and upper halves is 1.6 SD; the difference between the mean of the lower and upper tertiles is 2.2 SD; and the mean absolute difference between two randomly chosen values is 2.3 SD (WGH, unpublished observations).

Uncertainty in all estimates is shown as 90% confidence limits, in ± form for differences in means and in ×/÷ form for standard deviations and their ratios. Mechanistic magnitude-based inferences were made for the comparisons of double the values of the coefficients of variation, as described elsewhere[71] using a published spreadsheet[72]. The comparison was performed using the ratio of the coefficients of variation (CV), which was presumed to have a log-normal sampling distribution. In previous publications,[e.g., 73] a ratio of ~1.1 was considered a default value for the smallest factor difference, based on the effect CV representing errors of measurement have on sample size. Here we are interpreting CV representing modification of the dependent variable, so we have set magnitude thresholds for ratios such that the difference between double the CV is equal to the magnitude threshold for the dependent variable. For example, for a CV of ~3.0% (environmental and venue variation in this study) and a threshold for the smallest effect on the dependent variable of ~0.3% (this study), the other CV would have to be (2×3.0+0.3)/2 = 3.15% to be substantially different, so the threshold ratio is 3.15/3.0 = 1.05. Similarly, for CV of ~1.0% (within-boat race-to-race variation in this study), the threshold ratio is 1.15.

The within-year intraclass correlation coefficient (ICC) (final-to-final reproducibility in any given year) was calculated as the pure between-boat variance in a given final (sum of the variances represented by Boat and Boat*Year random effects)

divided by the observed between-boat variance in a given final (sum of the pure between-boat variance and the within-boat variance represented by the mean residual). The between-year ICC (reproducibility between finals across calendar years) was calculated as the pure between-boat variance in a given year (represented by Boat alone) divided by the observed between-boat variance in a final (as above). Confidence limits for the ICC were derived by assuming the within-boat/between-boat ratio of the sample/population variance ratio had an F sampling distribution.

A spreadsheet[72] was used to make mechanistic magnitude-based inferences[71] for the comparison of correlations. To assess the magnitude of the ICC and their differences, the usual thresholds of 0.1, 0.3, 0.5, 0.7 and 0.9 for low, moderate, high, very high and nearly perfect[71] may not apply to athletic performance. We have therefore devised a set of thresholds by assuming that a 2SD difference in performance between athletes in one race should predict a small, moderate, large... difference in performance between those athletes in another race, if the correlation between the races (the ICC) is small, moderate, large… Given the relationship $\Delta Y/SD_Y = r.\Delta X/SD_X$ for variables Y (performance in Race 2) and X (performance in Race 1) with correlation r (the ICC), and given $SD_Y = \sqrt{(SD_B^2 + SD_W^2)}$ (where SD_B and SD_W are between- and within-athlete SD) and ICC = $SD_B^2/(SD_B^2 + SD_W^2)$, then with $\Delta X/SD_X = 2$ and $\Delta Y =$ the fraction f of SD_W for small, moderate, large… effects, the relationship between Y and X yields a quadratic for the ICC in terms of f, with the positive solution $f^2(\sqrt{(1+16/f^2)}-1)/8$. The threshold values of 0.3, 0.9, 1.6, 2.5 and 4.0 for f therefore result in thresholds for low, moderate, high, very high and nearly perfect ICC of 0.14, 0.36, 0.54, 0.69 and 0.83.

Results

The mean time for the various A finals estimated from the mixed model ranged from 5:42 to 8:07 min:s. These times were on average 7.8 ± 1.2 percent slower (mean ± SD) than 2010 world-best times for each boat class posted at the Fédération Internationale des Sociétés d'Aviron website (worldrowing.com). For boat classes in which females and males competed (1X, 2X, 2-, 4X, 8+), mean A final times were 9.8 ± 0.4 percent slower, and overall there were 15% fewer boats for females. There was no consistent trend in the overall mean time from year to year.

Olympics and world championships for non-Olympic boat classes were overall 1.5% faster than world championships, which were in turn 2.7% faster than world cups. The mean increases between times in the finals were 0.9% (A-B), 1.7% (B-C), 1.4% (C-

D) and 1.7% (D-E); in terms of the magnitude thresholds defined by the within-boat final-to-final variability (see below), these differences were all moderate to large. There were trivial differences between males and females (~0.2%) in these increases. Uncertainties in all these estimates were negligible.

Table 1.2 Mean of the relative mean final times of all boat classes at each regatta venue, ranked by mean time (fastest to slowest).

Venue	Mean time (%)	No. of boat classes	No. of regattas
Lucerne	-2.8	17	11
Poznan	-2.7	17	4
Amsterdam	-2.0	17	1
Princeton	-1.4	12	1
Linz	-1.4	17	2
Beijing[a]	-1.0	14	1
Seville	-0.8	17	2
Gifu	-0.7	17	1
Banyoles	-0.1	17	2
Milan	0.0	17	2
Hazelwinkel	0.1	17	2
Sydney[a]	0.1	14	1
Eton	0.7	17	2
Athens[a]	0.9	14	1
Munich	1.2	17	10
Zagreb	1.8	3	1
St Catherine's	2.7	17	1
Vienna	6.4	17	3

Mean times were provided by the random-effects solutions for Venue, which have an overall mean of zero. SD of the 3-17 values contributing to the mean at each venue are ~1.0%.

[a]Olympic games.

The random variation in mean time from final to final (representing presumably the effects of environmental variation) expressed as CV ranged from 2.1% to 3.7% (each with 90% confidence limits ~×/÷1.20) across the 17 boat classes–all extremely large when doubled to interpret their magnitudes–with an overall mean CV of 2.8% (×/÷1.04). Singles showed a possibly moderately greater effect of the environment than crewed boats (by a factor of 1.19, ×/÷1.09). Female boat classes showed a likely small

increase in environment-related variability compared with males (by a factor of 1.11, ×/÷1.09).

Venue-to-venue differences in mean time expressed as CV were similar to those of final-to-final variation: a range in CV of 1.7% to 4.1% (~×/÷1.8) across the 17 boat classes, with an overall mean CV of 2.9% (×/÷1.13). The CV were possibly much greater for boat classes with fewer rowers in the boats (e.g., CV for singles vs eights and fours, 3.7% vs 2.5%; a ratio of 1.48, ×/÷1.37), but the observed trivial difference between sexes was unclear (female/male ratio of 0.99, ×/÷1.29). The random-effects solution for the Venue effect, representing percent differences in the speed of each venue from the mean venue, is shown in Table 1.2 There were very large to extremely large differences from the mean venue for the fastest (Lucerne, Poznan) and slowest (Vienna) venues. The Olympic venues are scattered towards the middle of the venues, but the actual speed of Olympic venues needs to include the fixed effect of Olympics noted at the beginning of the Results section. With this added effect, Beijing had the fastest times.

Given the uncertainties in the estimates of the race-to-race variabilities within a year (within-boat CV in Table 1.3), some averaging was considered appropriate and revealed a very likely small increase in variability in singles compared with crewed boats (singles 1.1%, pairs 0.9%, fours and eights 0.9%; ratio singles/crewed 1.28, ×/÷1.09). The threshold for smallest worthwhile enhancements in rowing performance is therefore ~0.3% overall (0.3× ~1.0%)–slightly more for singles (0.33%) and slightly less for crewed boats (0.27%). Thresholds for other magnitudes are moderate 0.9%, large 1.6%, very large 2.5%, and extremely large 4.0%. There was a likely trivial difference in variability for males vs females (ratio of 1.08 (×/÷1.09), There was possibly a small increase in variability of the O finalists overall compared with the A finalists (by a factor of 1.19, ×/÷1.06), but the factor difference decreased with increasing number of rowers in the boats (singles 1.28, ×/÷1.09; pairs 1.21, ×/÷1.10; fours and eights 1.11, ×/÷1.12). Boats generally showed trivial to moderate additional variation between races from one year to another (within-boat between-years CV, range 0.0-0.8%; see Table 1.3).

Differences between boats in A finals expressed as CV ranged from 0.6% to 1.7% (see Table 1.3), which are moderate to very large in magnitude when doubled. There was a possibly trivial-small increase in differences for females vs males (by a factor of

1.12, ×/÷1.15). Differences between boats in O finals (data not shown) were greater than those in A finals by a possibly trivial-small amount (a factor of 1.10, ×/÷1.11).

Table 1.3 Within- and between-boat variability of final times expressed as coefficients of variation (CV), and resulting predictability of performance in A finals expressed as within-year and between-years intraclass correlation coefficients (ICC) for each boat class.

| | Within-boat CV (%) | | | Between-boat CV (%) in A finals | ICC for A finals | |
| | Within-year | | Between-years | | | |
	A finals	Other finals			Within-year	Between-years
Singles						
LM1X[a]	1.4	1.6	0.7	1.4	0.57	0.46
LW1X[a]	1.1	1.2	0.3	1.0	0.50	0.45
M1X	1.2	1.5	0.3	1.4	0.57	0.54
W1X	1.0	1.5	0.5	1.7	0.78	0.71
Pairs and doubles						
LM2-[a]	0.9	0.8	0.4	1.1	0.64	0.57
LM2X	0.7	1.0	0.4	0.9	0.63	0.55
LW2X	1.0	1.4	0.0[b]	0.8	0.38	0.38
M2-	0.9	1.8	0.3	1.5	0.75	0.72
W2-	0.9	0.6	0.6	1.1	0.68	0.54
M2X	1.0	1.3	0.8	0.9	0.61	0.32
W2X	0.9	1.2	0.5	1.2	0.69	0.61
Fours and quads						
LM4-	0.6	0.9	0.6	0.8	0.73	0.50
M4-	0.7	0.8	0.3	0.9	0.64	0.57
M4X	1.1	0.9	0.0[b]	0.6	0.22	0.22
W4X	0.9	0.9	0.0[b]	1.7	0.79	0.79
Eights						
M8+	0.8	1.1	0.5	1.1	0.71	0.59
W8+	0.8	0.8	0.0[b]	1.1	0.65	0.65

90% confidence limits for CV (as ×/÷ factors): within-boat within-year, ~1.15; within-boat between-years and between boat, ~2.0. 90% confidence limits for ICC: ~±0.13.

[a]Non-Olympic boat classes.

[b]Negative variance (arising from sampling variation) was set to zero by Proc Mixed.

The within-boat final-to-final variability within and between years, when combined with the between-boat differences, gave the intraclass correlations representing the predictability of performance shown in Table 1.3. For races within a year the predictability was low to very high, with a mean of 0.63 (very high). There was a likely

trivial difference between the predictability of singles vs crewed boats (ICC of 0.62 vs 0.64; 90% confidence limits for difference, ±0.12) and a similar difference for females vs males (0.66 vs 0.60; ±0.06). The between-year ICC were less than the within-year, but the difference was probably trivial (0.08, ±0.03).

Discussion

We performed this study primarily to investigate the variability of performance in finals of elite rowers and crews. The estimates of final-to-final variability have provided estimates of the smallest worthwhile changes in performance. Other measures of variability and predictability in this study have shed light on the nature of international rowing competition.

The overall differences of 1.5% and 2.7% between mean times in the three levels of competition (Olympics and non-Olympic world championships, world championships, world cups) are remarkable, considering that linear modeling estimates such effects with "other things being equal". These differences are therefore not explained by the inevitably substantial differences in mean caliber of competitors between the three levels of competition. One obvious explanation is the construction or choice of faster venues for the more important events. It also seems likely that there are differences in motivation or preparation for these events: most international rowing programs aim to peak annually for the world championship, with the highest peak in the year of the Olympics (TBS, unpublished observations).

The differences in mean times of ~1-2% between consecutive levels of finals (A, B, C…) represent performance targets for athletes who generally end up in the lower-level finals. However, the variability in mean time from final to final is of such a magnitude that B, C or even D finals could end up faster than A finals at a given venue. Such outcomes are not uncommon in international rowing regattas, where finals separated by hours or days are often performed under different environmental conditions. Our finding of more variability in mean time between finals for boat classes with fewer rowers in the boat supports the belief amongst many in rowing that adverse environmental conditions have a greater effect on finals times of the smaller boats (TBS, unpublished observations). Consistent differences in environmental conditions are presumably responsible at least in part for the huge differences in overall mean times between the different venues (Table 1.2). For example, in the experience of one of the authors (TBS), the slowest course (Vienna) has a prevailing headwind that is generally the strongest of all the courses.

The within-boat final-to-final variability of elite rowing single and crewed boats is similar to that of elite athletes of comparable high-intensity endurance sports that are performed against water or air resistance.[66-68]   Although the 0.8% performance variability of elite middle distance track runners[74] is similar to that of rowers, the variation in physiological power output the athlete must generate to achieve this variability in time differs dramatically between these sports. In running a 1% change in running speed requires a 1% change in physiological power output, whereas a 1% change in the speed of a rowing boat requires approximately a 3% change in physiological power output to overcome the effect of fluid resistance.[75] The physiological race-to-race variability in power output of an elite rower is probably similar to the 0.8% of an elite middle-distance track runner, so the variability in the rower's times should be only ~0.3% (0.8/3) rather than the 0.9% to 1.4% we observed in singles, other things being equal. We believe the additional variability arises predominantly from environmental conditions; for example, a tailwind could benefit some rowers more than others, and a side wind could unfairly benefit rowers in lanes in which there is a wind shadow. The effects of wind could also explain why performance times in flat-water kayaking and time-trial cycling are inherently more variable than in running and similar to those in rowing, but some other factor or factors must explain why swimming times are almost as variable as those in rowing. Water turbulence, transient bulk currents, and the start and turn times may have a role in swimming. Swimming, kayaking and rowing may also differ in the extent to which athletes can maintain the efficiency of their highly skilled movements between competitions.

The lower within-boat variability of the A finalists in the single sculls is consistent with other sports where the better athletes tend to have less variability.[66-70, 74] In these sports it was suggested that the top-ranked individual athletes have less variability because they are better prepared, are more experienced at racing, or have more consistent motivation. Surprisingly, this difference in variability between A and O finalists was less evident in the crewed boats. It seems unlikely to us that crewed boats are relatively better prepared and more experienced across the various finals than competitors in other sports. We suspect it is more likely that motivation is enhanced in a lower-tier final by the threat engendered by the common practice of replacing rowers with perceived poor form in underperforming crewed boats. It is also possible that competition motivation is more consistent within a crewed boat, because crew dynamics create an obligation for each rower to compete maximally or risk censure from crewmates.

If the only source of within-boat race-to-race variability in finals times were the independent variability of the power output of the rowers, the variability of boat performance would be the standard error of the mean performance of the individual rowers. For equal variability of rowers in the crewed boat the error in their mean is this variability divided by the square root of the number in the crew. The larger boats did indeed have less variability than the smaller boats, but the decrease is less than that predicted from just the error in their mean. One obvious explanation is the substantial contribution of environment to the variability, which as argued above is less in a crewed boat but apparently not in proportion to the root of the crew number. In any case, the individual rowers in a crewed boat are unlikely to perform independently, owing to similarity of their training and a group dynamic that entrains their performance during competition.

Our finding of possibly lower variability in race-to-race performance in female rowers relative to the males is in contrast to other studies in which females tended to have greater within-subject race-to-race variability.[68-71] We can only speculate that, in rowing, females are at least as consistent as males in preparation for competition, pacing in the competition, and motivation to perform.

Perhaps the most important result in the present study is the estimate of the smallest worthwhile enhancement in performance time for top-ranked finalists competing in rowing: ~0.3%. Enhancements aimed at improving physiological power output need to be ~3× these values (~1.0% to 0.5%) to effect a change in medal prospects. These smallest meaningful effects on power output also represent desirable targets for the error of measurement in rowing performance tests that would be sufficiently sensitive to quantify trivial-small changes in performance, either when monitoring individual athletes or performing controlled trials with realistic sample sizes.[76] Two estimates of error of measurement of mean power in 2000-m time trials on the Concept II rowing ergometer, 1.3%[77] and 2.0%,[78] fall somewhat wide of this target but with considerable uncertainty (95% confidence intervals 0.9 to 2.9% and 1.3 to 3.1%). Thus, it is possible that the Concept II 2000-m maximal performance test does have an error close to the smallest worthwhile effect for single sculls, and the error might be smaller with more familiarization trials and certainly by combining several tests. If the error of measurement equals the smallest worthwhile effect, the resulting sample size for adequate precision with the worst-case scenario of a trivial outcome would be as low as ten subjects for a crossover trial and 13 subjects in each group for a parallel-groups

controlled trial; samples of at least four times these numbers are needed for estimation of individual responses or effects of covariates.[79]

The additional within-boat variability between races separated by a year or more was trivial overall, which indicates that rowers' performances are generally consistent over more than one season. Over several seasons the year-to-year variability will be additive in some rowers and result in substantial changes in performance. Authors of the only comparable study with a measure of within-subject between-year variability[66] reported similar results and suggested that athletes at the top of the field are in a consistent state of preparation from year to year.[66]

To date the only sports to report intraclass correlation coefficients to assess predictability are slalom canoe-kayak[70] and skeleton,[69] where the correlations were somewhat lower (range 0.01 to 0.61), possibly reflecting a greater influence of changes in the environment or the course on performance of individual athletes between races. More research, perhaps involving simulation, is needed to provide evidence that the magnitude thresholds we have devised for the ICC are appropriate for this measure of predictability. Meanwhile, with reference to these thresholds, predictability of rowing performance was overall high and showed showed little reduction for races between years compared with those within a year; that is, changes in finals placings between races would be similar whether the races were in the same year or in consecutive years. Although there was a possibility of different variabilities within and between boats for males compared with females, when these variabilities were combined, the resulting predictability for females was only a little higher than that for males and the difference was inconsequential.

In conclusion, the race-to-race variabilities and smallest worthwhile effects in elite rowing are similar to those for elite athletes in comparable endurance sports where environment and technical demands can affect performance. Differences in these statistics between male and female rowers are at most small. Crewed boats showed less variability than singles, probably because of less effect of the environment and some averaging of the variability of the individual crew members. The potential for environmental conditions such as wind shadow to have unfair effects on rowing performance warrants further investigation.

Acknowledgments

CHAPTER 2

Measures of Rowing Performance

Running head: Reliability and validity of rowing measures.

Overview

Accurate measures of performance are important for assessing competitive athletes in practical and research settings. We present here a review of rowing performance measures, focusing on the errors in these measures and the implications for testing rowers. The yardstick for assessing error in a performance measure is the random variation (typical or standard error of measurement) in an elite athlete's competitive performance from race to race: ~1.0% for time in 2000-m rowing events. There has been little research interest in on-water time trials for assessing rowing performance, owing to logistic difficulties and environmental perturbations in performance time with such tests. Mobile ergometry via instrumented oars or rowlocks should reduce these problems, but the associated errors have not yet been reported. Measurement of boat speed to monitor on-water training performance is common; one device based on GPS technology contributes negligible extra random error (0.2%) in speed measured over 2000 m, but extra error is substantial (1-10%) with other GPS devices or with an impeller, especially over shorter distances. The problems with on-water testing have led to widespread use of the Concept II rowing ergometer. The standard error of the estimate of on-water 2000-m time predicted by 2000-m ergometer performance was 2.6% and 7.2% in two studies, reflecting different effects of skill, body mass and environment in on-water vs ergometer performance. However, well-trained rowers have a typical error in performance time of only ~0.5% between repeated 2000-m time trials on this ergometer, so such trials are suitable for tracking changes in physiological performance and factors affecting it. Many researchers have used the 2000-m ergometer

performance time as a criterion to identify other predictors of rowing performance. Standard errors of the estimate vary widely between studies even for the same predictor, but the lowest errors (~1-2%) have been observed for peak power output in an incremental test, some measures of lactate threshold, and measures of 30-s all-out power. Some of these measures also have typical error between repeated tests suitably low for tracking changes. Combining measures via multiple linear regression needs further investigation. In summary, measurement of boat speed, especially with a good GPS device, has adequate precision for monitoring training performance, but adjustment for environmental effects needs to be investigated. Time trials on the Concept II ergometer provide accurate estimates of a rower's physiological ability to output power, and some submaximal and brief maximal ergometer performance measures can be used frequently to monitor changes in this ability. On-water performance measured via instrumented skiffs that determine individual power output may eventually surpass measures derived from the Concept II.

Introduction

Rowing competitions usually involve races lasting 6-8 min over a 2000-m regatta course. Rowing demands a high level of endurance,[80] estimates of the aerobic contribution being 70-87%.[81-83] Successful rowers tend to be tall, heavy and lean.[83-85]

There are up to 22 boat classes in international rowing regattas, consisting of heavyweight or lightweight rowers, male or female rowers, singles or crewed boats, coxed or coxless boats, and sweep or sculling oars. In sculling, each rower has two oars, and boat types involve one rower (single), two rowers (double) or four rowers (quad). In sweeping, each rower has one oar, and boat types include two rowers without a coxswain (coxless pair), two rowers with a coxswain (coxed pair), four rowers without coxswain (four), and eight rowers with a coxswain (eight).

Measuring changes in performance is important for monitoring the progress of rowers during training and for research assessing the effect of training and other interventions. In the only review[43] of the tools and tests available for monitoring rowing performance, the authors focussed on monitoring for overtraining, with only a cursory examination of the accuracy of the various measures of rowing performance. In the present review we describe these measures, their errors, and the practical implications.

We used Google Scholar, Sportdiscus and reference lists in reviews and research articles to search for investigations on measures of rowing performance published since 1970. To examine whether we had missed any relevant material we undertook follow up searches in Google Scholar and Sportdiscus where we combined the term *row* with the specific measure/s of interest: *GPS, Concept II, time-trial, lactate tests,* etc.

The yardstick for evaluating a measure of athletic performance is the variability that top athletes show from one race or competition to the next. This variability, which is expressed as a within-subject standard deviation, is analogous to the standard error of measurement (or typical error of measurement) in a reliability study of a performance test, with the repeated trials replaced by races. The race-to-race variability in finish times for elite rowers in world cups, world championships and Olympic competitions is ~1%.[86] These races are maximal efforts for highly motivated and well-conditioned rowers, so the ~1% race-to-race variability is at first glance an irreducible error for any measure of rowing performance. However, variability arising from environmental and other factors probably adds to the rower's inherent physiological variability in performance from race to race,[86] so measures of performance derived from a rowing ergometer can and do have standard error of measurement less than 1% in reliability studies.

A useful measure of rowing performance must have acceptable validity as well as reliability. Validity of a measure is determined by comparing its values with those of a criterion measure, which in rowing is competitive performance time over 2000 m. The single best validity statistic is the standard error of the estimate (or typical error of the estimate), which is a standard deviation representing the error in an individual's criterion value predicted by the test or other measure. The standard error of the estimate and the other regression validity statistics (the correlation coefficient and the regression or calibration equation) are specific to the population represented by the sample from which they are derived, and the standard error of the estimate is misleadingly smaller in samples with a narrower between-subject standard deviation. Indeed, with a homogeneous sample the standard error of the estimate is simply the noise (standard error of measurement) in the criterion. For this reason we have used a method[86] to adjust the standard error of the estimate from each study to a population with the widest possible range of values (infinite standard deviation). See Appendix 1 for the formulae. The standard error of the estimate is then an unbiased estimate of the error in the criterion value arising from error in the test measure and as such can be compared between studies for the purpose of choosing measures with the smallest errors. The

adjusted standard error of the estimate of a test measure is inevitably larger than the standard error of measurement of the criterion, because the standard error of the estimate includes contributions from the standard error of measurement of the criterion, the standard error of measurement of the test measure, and differences between subjects that are not accounted for by the test measure.

The smallest worthwhile change in performance is another important consideration in the assessment of measures of performance. When the standard error of measurement of a performance measure is similar in magnitude to the smallest worthwhile change, the measure is sufficiently sensitive to quantify small but meaningful changes, either when monitoring individual athletes or when performing controlled trials with realistic sample sizes.[86] The smallest worthwhile enhancement of an elite athlete's performance has been defined as the change in performance time or other score that results on average in one extra medal in every 10 competitions.[65] Simulations showed that this enhancement is a factor of 0.3 of the standard deviation of within-athlete race-to-race variability in performance, which for rowing is therefore a 0.3% change in race time (0.3×1.0%). Thresholds for quantifying magnitudes based on winning 3, 5, 7 and 9 extra medals per 10 races are 0.9% (moderate), 1.6% (large), 2.5% (very large), and 4.0% (extremely large).[71] To interpret the magnitude of an error (standard error of measurement or standard error of the estimate) on this scale of magnitudes, the error should be doubled.[86] A good measure of rowing performance would therefore need an standard error of measurement of less than 0.3% if one wanted to be confident about trivial changes in performance. We will see that no rowing tests reach this level of precision.

Measures of On-Water Rowing Performance

Although the 2000-m on-water time is the criterion measure of rowing performance, this measure has a number of limitations. Environmental conditions are the most important limitation, which contributes substantial variability to the competitive performance of elite rowers[86] Even for a group of rowers competing together in a single race and therefore under seemingly identical environmental conditions, rowers could be affected differently; for example, a headwind could hinder some rowers more than others, and a side wind could unfairly benefit rowers in lanes in which there is a wind shadow. Furthermore, without special instrumentation in the boat, it is not possible to quantify an individual rower's performance from the speed of a boat with

two, four or eight rowers. Rowers from a crewed boat could be tested individually in single sculls, but the faster movement speeds in a crew boat could easily create a difference in technique. There are also biomechanical differences between single sculling and sweeping.[87, 88] Instrumenting the boat (see later) may solve the problems of assessing the on-water performance of individual rowers in single or crewed boats. Many coaches will not permit regular maximal 2000-m on-water tests during the earlier stages of the season, owing to concerns about impairing aerobic development, but performance would be available from the competitions over 2000 m that occur regularly during the ~4 month competitive season (TBS, personal observations).

Despite problems with the 2000-m on-water performance tests, maximal boat speed over distances other than 2000 m is a commonly used measure in rowing.[89] The distances range from 250 m to 15 km, depending on the energy system(s) being trained in a given phase of the season (TBS, personal observations). A common practice in this context is the use of "prognostic speeds", which are either world record or some other target speed for each boat class.[89, 90] The performance of each boat is evaluated as a percentage of its prognostic speed to provide a ranking within a boat class and a comparison for boats from different boat classes. Environmental conditions can still affect the accuracy of the ranking, because the environment has a greater effect on boats with fewer rowers and boats with female crews.[86] The boats are often handicapped for these tests in an effort to improve the sense of competition. While handicapping does reduce environmental effects, provided of course that conditions do not change substantially between boats, the turbulence (boat wash) can disadvantage the trailing boats.

The development of various rowing speedometers has increased the popularity of the measurement of boat speed during training and competition.[91] These devices are more convenient than stop-watches, which require timing of at each end of the course and accurate measurement of the distance. The two common speedometers are based either on an impeller or the global positioning system (GPS). The impeller measures speed relative to water (true speed), while GPS measures speed relative to land.

Impeller Measurement of Boat Speed

The two popular devices that measure rowing boat speed via impellers attached to the hull are the Nielsen Kellerman Speed Coach (Nielsen Kellerman, Boothwyn, PA) and the Coxmate (Coxmate, St Peters, South Australia). The impeller has two

advantages over the GPS. One advantage is that it can give accurate readings over any distance, whereas GPS has unacceptable error over short distances (see below). The other is that the impeller measures the speed of the boat relative to any water current, so impeller speed more accurately reflects the performance of the rowers. Windy conditions can create water currents,[92, 93] and while the boat speed calculated by the impeller accounts for these currents, the impeller obviously does not adjust for the direct effect of the wind on the boat, rowers and oars;[94, 95] for example when there is a headwind, an impeller will indicate appropriately a slower boat speed.

The impellers are calibrated upon installation and checked regularly thereafter, as there is anecdotal evidence that weed or other debris in the water can upset the calibration. In an effort to increase the accuracy of the impellers in flowing water, they are calibrated by travelling a known land distance upstream and downstream. The combined land distance for the two runs is compared to the combined impeller distance, which allows the appropriate calibration to be calculated for that stretch of water. Accurate measurement of speed with an impeller requires a constant speed and direction of the water current over the period of testing and through all parts of the waterway that the boats travel.

There are no published studies examining the reliability and validity of the impeller in rowing. To gain some understanding of their accuracy, one of the authors examined the Nielsen Kellerman impeller (NK) during regattas over 2000 m on various international rowing courses. In 61 observations of NK versus true boat speed over 2000 m, the NK units showed a negligible fixed error (0.1%), but there was a moderate random error of 1.2% even when wind direction and speed were taken into account (TBS, personal observations). While this amount of error is only slightly larger than the 1% yardstick considered appropriate to accurately monitor training, the NK is not sufficiently accurate to quantify small but meaningful changes in competitive performance.

GPS Measurement of Boat Speed

GPS requires an unobstructed view of its satellites, but this requirement is seldom a problem on the waterways where rowers train and compete. The devices are easily swapped between boats and do not require calibration.

The early GPS devices sampled and calculated position once per second (1 Hz), but with recent technical developments units that sample as high as 20 Hz are now

available. Higher sampling frequencies are needed for accurate speed measurement over shorter intervals or distances, but the 1-Hz units are still in principle adequate (<1% error) for quantifying boat speed over durations in excess of 100 s.

Proprietary algorithms employed by the various manufacturers of GPS are also considered to influence accuracy, so the findings from a particular GPS model should be considered to apply only to that model.[96] Previous research on inter-unit reliability has established that there is little difference between units of the same model of GPS, at least over long durations and distances,[97] so it is probably safe to assume that findings with one unit apply to all such units of a given model.

GPS technology has undergone a series of rapid advances since May 2000, when the US government made full precision available. We have therefore limited this review to GPS studies published since then. We have included data from studies with movement patterns similar to that of rowing, that is, straight-line movements. We have also included data from movements around 400-m running tracks but have excluded zigzag or T-shaped shuttle runs. In all, three studies provided useful information for rowing,[98-100] and we have included some unpublished observations. See Figure 2.1

Figure. 2.1 Plot of standard error of estimate (SEE) for each measurement device over different distances
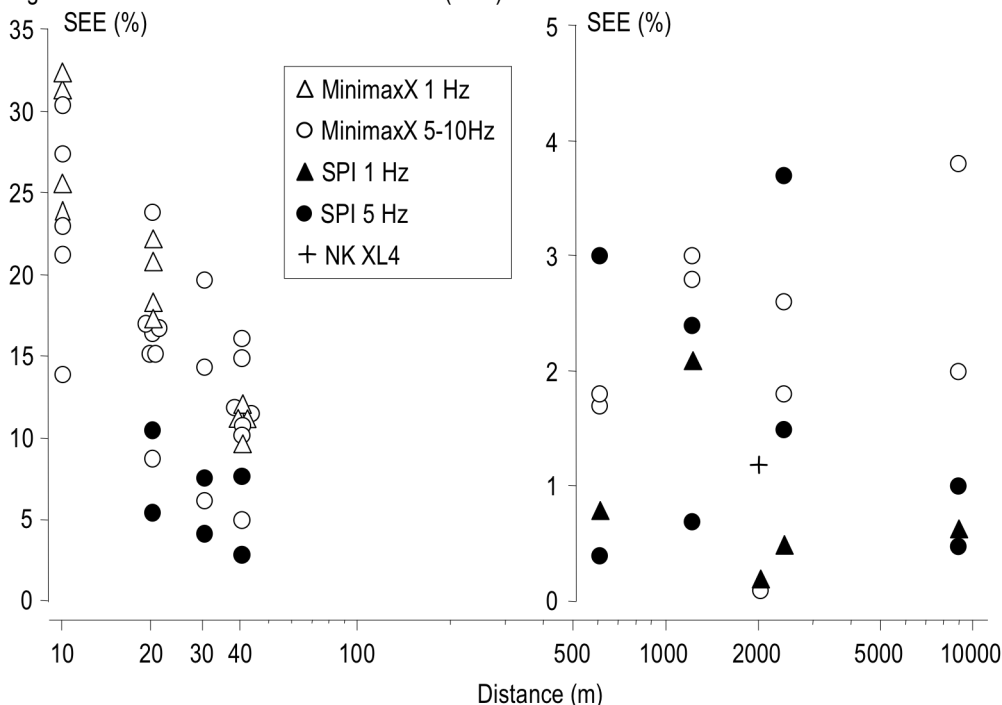
The accuracy of the 5-Hz MinimaxX (Catapult, Melbourne, Victoria), the 1-Hz SPI-10 and the 5-Hz SPI-Pro (GPSports, Fyshwick, ACT) was examined over a range of

speeds and distances by Petersen et al.[99] The distances recorded by the GPS were compared to the actual distances travelled on a 400-m running track for walking 8800 m through to sprinting 20 m. It is apparent in Figure 2.1 that the accuracy of the GPS is more dependent on the manufacturer than the signal frequency. Figure 2.1 also shows the 5-Hz version of the MinimaxX improved the error relative to the 1-Hz version regardless of distance and speed.[98] Pyne et al.[100] compared the 10-Hz MinimaxX with previously published results for the 5-Hz version for straight line sprinting over 10, 20 and 40 m. Not apparent in Figure 2.1 is the fact that the accuracy of the 10-Hz version was better than that of the 5-Hz.

Rowing has a unique problem for speed measured over a short distance. The average speed for the various boat types during competition is 4-6 m.s^{-1}, but speed varies by 2-3 m.s^{-1} during each stroke.[101, 102] An aliasing error may arise from a combination of a 1-Hz sampling frequency, the large oscillations in velocity and a short sample period, and this error may worsen the already poor accuracy of the GPS over shorter durations and distances. Over the shortest distances of interest to rowers (~250 m), GPS sampling at greater frequencies (>1 Hz) would overcome this aliasing error.

The accuracy of GPS in rowing has been examined in one published study, which is available only as a conference abstract.[103] In this study the race time recorded by 5-Hz MinimaxX was compared to the official race time for 244 rowing boats during major competitions over 2000 m. The standard error of the estimate was 0.45 s, but not enough data were presented to convert this error to a percentage. However, if we assume the race time was 6-8 min, the standard error of estimate would be ~0.1%.

One of the authors also examined the accuracy of the GPS in rowing by comparing the distance recorded by 10 SPI-Elite 1-Hz units with the 2000-m distance travelled by 22 rowing boats during various regattas. The standard error of the estimate was negligible (0.2%, TBS personal observations) and similar to the ~0.1% estimated above from the 5-Hz MinimaxX for 2000-m rowing. The standard error of the estimate for these two GPS are shown in Figure 2.1 as the point with the smallest error. These error are well within the 1% yardstick and are sufficiently low to quantify small but meaningful changes in boat speed in 2000-m time trials. Further research is required to determine the GPS accuracy during rowing over distances shorter than 2000 m, when aliasing might begin to make a substantial contribution.

From the data presented in Figure 2.1 we make the following conclusions. The SPI elite and MinimaxX GPS are more accurate than the NK over 2000 m and the various

SPI GPS units are more accurate than the MinimaxX except for 2000-m rowing. The MinimaxX has less error at higher sampling frequencies but the effect of frequency is not clear for the various SPI units. During rowing the SPI-Elite has a smaller error than that determined over similar distances on a 400-m running track for other SPI models, even for those with a higher sampling frequency. It is therefore likely that cornering on athletics tracks increases GPS error.[96]

Environment changes will cause changes in boat speed, so regardless of how speed is measured, it can be an inaccurate method for tracking rowing performance. Even if ideal environmental conditions could be guaranteed between trials (no wind, no water currents, no changes in the composition, depth and temperature of the water), boat speed is still not an accurate measure of an individual rower's performance within a crew boat.

On-Water Ergometry

The measurement of a rower's power output in a boat is now possible with on-water ergometers, which have been constructed for both sculling and sweeping.[104] These ergometers calculate power output from kinetic data measured by sensors in the rowlock and/or oar(s).[105] Although in theory the power output from these devices should correlate strongly with boat velocity, findings have been mixed.[87, 105, 106] On-water ergometers are also expensive, time-consuming to install and calibrate, and often fragile (TBS personal observations). Despite the potential benefits of these ergometers, it remains to be determined how well the power measured by their sensors represents power propelling the boat forward. We therefore advise caution in the use of on-water ergometers until the associated errors have been reported.

Measures of Off-Water Rowing Performance

The difficulties associated with assessment of on-water performance have led to widespread use of stationary ergometers that simulate the action of on-water rowing. Studies of various rowing ergometers have found some differences in arm motion,[107] handle force and acceleration profiles[108] and consistency in stroke timing[109] between off-water and on-water rowing performance. Despite these differences the rowing ergometer is widely used by rowers, and the 2000-m ergometer time-trial is the most common measure of rowing performance.[43, 110] The Concept II air braked rowing ergometer (Morrisville, USA) has led the market since the development of the IIb

model in 1986. The three subsequent models (c,d and e) have maintained the same rowing motion and method for calculating work output, but have made changes to improve comfort, safety, robustness, damper settings and display options. A study to compare two Concept II models[111] was far too underpowered to make meaningful conclusions, but it is reasonable to assume that the only differences between the models are cosmetic (TBS, personal observations).

In a recent development the Concept II has been placed on a slide to allow back and forth motion that simulates more closely the dynamics of on-water rowing.[81, 112, 113] Comparisons of the static version of the Concept II with the new dynamic "slider" version suggested negligible differences in mean power output in time-trials of 2000-m and 6 min, but on the slider peak and mean stroke force were lower and stroke rate was higher.[81, 112, 113] The slider is becoming increasingly popular in training, as there is evidence that dynamic rowing ergometry puts less strain on the lower back, which is beneficial to rowers who commonly suffer back injury.[114] Other advantages include a better "on-water feel" and the capacity to link devices together to simulate crew rowing. The few studies comparing the static and slider versions of the Concept II lack the data to calculate the standard error of the estimate.[81, 112, 113] The only other rowing ergometer in contention is the Rowperfect (Devon, United Kingdom); while it may more closely simulate on-water rowing movement[115] performance on this ergometer is less reliable (see below).

Off-Water versus On-Water Time-Trials

To examine how accurately rowing-ergometer performance predicts on-water performance, we reviewed two studies where comparisons were made between 2000-m time-trials conducted on-water versus on a Concept II (see Table I). We had to exclude two further studies comparing Concept II performance with rankings from world championships,[116, 117] because the authors did not provide competitive performance times that would allow computation of a standard error of the estimate.

In a study of 10 junior males whose on-water tests were single scull competitions,[118] the standard error of the estimate was 2.6%, which in our scale of magnitudes is a very large error. Although the "competition results" were obtained on a "windless day", it is not clear whether the results were obtained from a single race. We therefore suspect environmental conditions contributed to the error. The limited competition experience of the young rowers (18.9 ± 1.7 y) may also have contributed to the error, along with

the substantial uncertainty in the estimate (90% confidence limits x/÷1.54) arising from the small sample size.

In another study[119] 49 junior elite males completed two 1000-m time-trials in single sculls, which were combined to give a 2000-m time and an standard error of the estimate of 7.2%. All tests were undertaken at a national training camp, which presumably ensured high motivation. Both single sculls time-trials were conducted on the same day "with winds of approximately 2–3 m.s^{-1}, the direction being predominately a headwind". Although there are various potential sources of error, we believe that change in wind speed and direction between trials for the different rowers was the main source of the extremely large error.

These large standard error of the estimate do not necessarily mean that performance on the Concept II is invalid; more likely, there is large random error in the criterion measure of 2000-m on-water time, most of which is due to environmental factors[86] When body mass was taken into account in a multiple linear allometric regression equation, the observed standard error of the estimate decreased from 4.1% to 3.1% (see Table 2.1), which is consistent with the observation that body mass provided a substantial contribution to Concept II time (r = 0.68) but a negligible contribution to single-scull time (r = 0.04).[119] The exponent of body mass in the allometric equation was approximately -0.8 so the widespread practice of expressing Concept II performance as mean power per kilogram must produce close to the optimal measure for combining body mass with ergometer performance.

Even when environment and body mass are taken into account, it is inevitable that some rowers perform better on the Concept II than on water and vice versa, so the Concept II cannot predict on-water ability perfectly. For a better estimate of the validity of performance on a Concept II, the standard error of the estimate needs to be obtained with a good sample size of top rowers under ideal environmental conditions.

Table 2.1 Standard error of estimate of 2000-m single-scull performance time derived from correlations of this performance measure with measures from tests on a Concept II rowing ergometer. Measures shown in order of adjusted SEE (lowest to highest).

Test measure	Rowers	Test measure	Correl-ation	SEE (%)	90% CI	Adjusted SEE (%)	Reference
2000-m time (min:s)	10 M	7:28 ± 0:13	0.72	2.1	1.4-3.3	2.6	Jurimae et al.[118]
Peak incremental power (W)	10 M	369 ± 37	-0.70	2.2	1.4-3.4	2.8	Jurimae et al.[118]
VO_2 @ 4-mM lactate (L.min^{-1})	10 M	4.13 ± 0.63	-0.69	2.2	1.4-3.4	2.9	Jurimae et al.[118]
VO_2max (L.min^{-1})	10 M	4.85 ± 0.63	-0.64	2.4	1.5-3.6	3.3	Jurimae et al.[118]
Lactate @ 350 W (mM)	10 M	11.8 ± 4.8	0.64	2.4	1.5-3.6	3.3	Jurimae et al.[118]
Power @ 4-mM lactate (W)	10 M	275 ± 41	-0.61	2.4	1.6-3.8	3.5	Jurimae et al.[118]
40-s all-out mean power (W)	10 M	614 ± 82	0.60	2.5	1.6-3.8	3.6	Jurimae et al.[118]
VO_2max (ml.kg^{-1}.min^{-1})	10 M	61.5 ± 5.6	-0.33	2.9	1.9-4.5	7.8	Jurimae et al.[118]
2000-m time and body mass[a]	48 M	?	0.77	3.1	2.7-3.7	?	Nevill et al.[119]
2000-m time (min:s)	48 M	9:08 ± 0:26	0.54	4.1	3.5-4.9	7.2	Nevill et al.[119]

2000-m single-scull times in Jurimae et al.[118] and Nevill et al.[119] were 7:28 ± 0:13 and 9:08 ± 26 min:s (mean ± SD).

For other subject characteristics, see Table 2.2. Data for test measures are mean ± standard deviation. **90% CI** = 90% confidence interval for the SEE; **M** = Male; **mM** = mmol.L^{-1}; **SEE** = Standard error of the estimate; **VO_2** = oxygen uptake; **?** = not provided or estimable.

a Measures combined via multiple linear regression.

Reliability of the Off-Water Time-Trial

If we accept that the performance on a Concept II has adequate validity, at least for physiological power output, an important issue is whether this ergometer has adequate reliability for tracking changes in performance. Reliability for tests on a Concept II has been reported in two studies. Schabort et al.[78] examined 2000-m time-trial speed on a Concept II for eight well trained rowers who rowed on three occasions at 3-day intervals and reported a 0.6% standard error of measurement. In the second study, 15 elite rowers performed five 500-m time-trials each on a Concept II and Rowperfect ergometer and later performed a 2000-m time-trial on 3 consecutive days on one ergometer.[77] The standard error of measurement for 2000-m time were 0.4% and 1.1% on the Concept II and Rowperfect respectively, while the standard error of measurement for the 500-m trials were 0.7% and 1.1%. Combining these studies, the standard error of measurement of ~0.5% is half the standard error of measurement for competitive on-water performance (our 1% yardstick) and is only just outside the 0.3% threshold to quantify smallest meaningful changes in competitive performance. Although this reliability is not ideal, it is unusual for tests of athletic performance to be this good.[75] This higher reliability of performance on the Concept II is likely a result of less technical demands of the ergometer and environmental effects causing less variability to

performance on a rowing ergometer compared to on-water rowing. In comparison, the Rowperfect was clearly inferior, and it is unclear whether its reliability would reach that of the Concept II with enough familiarisation.

Other Off-Water Test Measures

The rowing community was aware of the value of performance testing on the Concept II long before any reliability and validity studies were performed.[43, 80] Indeed, the 2000-m time-trial on the Concept II has become the most commonly used selection tool for national rowing organisations (TBS, personal observations). Furthermore, other measures derived from the performance test on the Concept II have been investigated for their ability to predict rowing performance. In one study the relationship of these measures to on-water 2000-m performance has been quantified, but by far the majority of the studies has used the 2000-m Concept II time-trial as the criterion measure.

In the one study that used 2000-m single scull performance,[118] the standard error of the estimate were moderate to large (2.1-2.8%), probably because of the different effect of environment, technique and body mass on performance on-water versus on the ergometer (Table I). Approximately half of the remaining studies had enough data to calculate the standard error of the estimate. For these studies the subject characteristics are included in Table 2.2, while the standard error of the estimate for the various measures are in Table 2.3. The measures that come close to the ~1.0% yardstick are peak incremental power, VO_2max , some measures of lactate power, and power in the 30-s modified Wingate. These measures have adequate validity for assessing moderate differences in rowing performance between rowers.

In four studies multiple linear stepwise regression analyses provided best combinations of measures to predict 2000-m Concept II time-trial performance. The sample sizes were far too low in three of these studies to perform such analyses, so the relatively low standard error of the estimate of 0.5 to 1% we obtained from their data must be substantial underestimates of the true error. In the other study, Nevill et al.[120] combined data from 48 males and 28 females to obtain a reasonable sample size, but the result is effectively a prediction equation for distinguishing between genders. Nevertheless, the resulting standard error of the estimate was relative low (1.6%), so there may still be some value in combining several measures for predicting 2000-m performance on the ergometer and especially on water. Definitive studies need to be performed.

Table 2.2 Subject characteristics for studies used to calculate the standard error of estimate in 2000-m ergometer performance time for measures from rowing tests.

Reference	Rowers	Body mass (kg)	VO_2max (L.min^{-1})	2000-m time (min:s)
Bourdin et al.[121]	31 National and International heavyweight males	88.6 ± 5.1	5.68 ± 0.32	6:04 ± 0:10
Bourdin et al.[121]	23 National and International lightweight males	74.0 ± 1.8	5.05 ± 0.20	6:21 ± 0:07
Bourdin et al.[121]	54 combined	82.4 ± 8.3	5.41 ± 0.42	6:12 ± 0:12
Cosgrove et al.[122]	13 club-level males	73.1 ± 6.6	4.5 ± 0.4	7:05 ± 0:11
Faff et al.[123]	8 male (?) teenagers	85 ± 14	4.97 ± 0.48	6:45 ± 0:14
Gillies & Bell[124]	10 competitive males	82.3 ± 7.5	4.38 ± 0.42	7:07 ± 0:14
Gillies & Bell[124]	22 competitive females	71 ± 10	3.19 ± 0.57	8:17 ± 0:30
Gillies & Bell[124]	32 combined	75 ± 11	3.62 ± 0.84	7:55 ± 0:42
Jurimae et al.[118]	10 experienced males	79.3 ± 7.3	4.85 ± 0.63	6:38 ± 0:18
Nevill et al.[120]	48 current/former Senior A/B males	88 ± 11	5.60 ± 0.56	6:07 ± 0:13
Nevill et al.[120]	28 current/former Senior A/B females	71.7 ± 8.5	4.03 ± 0.33	7:01 ± 0:17
Nevill et al.[120]	76 combined	82 ± 13	5.02 ± 0.91	6:27 ± 0:30
Nevill et al.[119]	48 elite junior males	83 ± 7	?	6:44 ± 0:11
Riechman et al.[125]	12 competitive females	67 ± 12	3.18 ± 0.35	7:47 ± 0:12
Russel et al.[126]	19 elite schoolboys	85 ± 8	4.6 ± 1.5	6:43 ± 0:16
Womack et al.[127]	10 college males (pre-Fall)	86.1 ± 7.3	5.25 ± 0.69	6:48 ± 0:18
Womack et al.[127]	10 college males (post-Fall)	86.1 ± 7.3	5.28 ± 0.62	6:42 ± 0:18

Data are mean ± standard deviation. **VO_2** = oxygen uptake; **?** = not provided.

A low standard error of the estimate for a performance measure is desirable, but measures with higher standard error of the estimate may still be useful if they have standard error of measurement low enough to reliably track an athlete's change in performance. Unfortunately, only one rowing study provided data to calculate the standard error of measurement for performance measures other than the 2000-m time-trial.[128] Owing to a complex research design in this study, we were only able to calculate the standard error of measurement for various measures of lactate threshold and for peak lactate in an incremental test. The standard error of measurement for Lactate threshold were relatively low (0.5 to 1.8%) but there is considerable error arising from the small sample size of 10 elite male rowers.

The lack of reliability studies in rowing led us to examine the reliability of similar performance measures for other modes of exercise in a comprehensive meta-analytic review.[75] The measure of reliability in the review was the standard error of measurement of power output; we have therefore divided the standard error of measurement by three to obtain an equivalent standard error of measurement for performance time, as explained in that review. The most reliable tests that might be applicable to rowing were peak incremental power (standard error of measurement for

time ~0.3%), VO_2max and lactate threshold (~0.5%). While all tests are well within the 1% yardstick, only peak incremental power could track smallest worthwhile changes. These measures also have the lowest standard error of the estimate for predicting 2000-m time-trial performance on a Concept II (see Table 2.3). The only other measure with a low standard error of the estimate in our review is the modified 30-s Wingate test on the Concept II, but the standard error of measurement of Wingate measures was somewhat larger (~1.2%) than that of the other two measures in the meta-analytic review.[75] Thus it is possible that Wingate performance is more reliable on the Concept II than on other ergometers.

In summary, peak incremental power, VO_2max, some measures of lactate threshold power and possibly 30-s power have measurement properties that make them potentially valuable for assessing rowing performance. In our view VO_2max provides no information additional to that provided by peak incremental power, which along with 30-s power has the advantage of requiring no equipment other than the Concept II. These tests can be performed weekly at any time of the year with little impact on the training program. Whether the measures can track performance adequately on the rowing ergometer and more importantly on water is a question that needs to be addressed with further research.

Conclusion

Measures of on-water rowing performance are very noisy, owing to the effects of environment, and they do not measure performance of an individual in a crew. Performance testing on the Concept II eliminates these problems. Peak incremental power and 30-s power on this ergometer are likely to be useful for frequent monitoring of a rower's physiological power output. However, the Concept II does not adequately address the skill component of performance on water. Instrumentation to measure each rower's on-water power output should provide the best measure of rowing performance, but it remains to be seen whether the errors are acceptably low.

Table 2.3 Standard error of estimate of 2000-m ergometer performance time derived from correlations of this performance measure with aerobic physiological test measures. All tests were performed on a Concept II rowing ergometer. Measures shown in order of adjusted SEE (lowest to highest)

Measure	Rowers	Test measure	Correl-ation	SEE (%)	90% CI	Adjusted SEE (%)	Reference
Peak incremental performance							
Power (W)	10 M	369 ± 37	-0.97	1.2	0.9-1.9	1.2	Jurimae et al.[118]
Power (W)	54 M	422 ± 37	-0.92	1.3	1.1-1.5	1.4	Bourdin et al.[121]
Power (W)	31 M, HW	441 ± 34	-0.89	1.3	1.0-1.6	1.4	Bourdin et al.[121]
Power (W)	23 M, LW	396 ± 23	-0.76	1.2	0.9-1.5	1.4	Bourdin et al.[121]
Power (W)	28 F	256 ± 23	-0.92	1.6	1.3-2.0	1.7	Nevill et al.[120]
Speed ?	13 M	?	-0.77	1.7	1.3-2.5	?	Cosgrove et al.[122]
Power (W)	48 M	369 ± 37	-0.84	1.9	1.6-2.3	2.3	Nevill et al.[120]
Power (W)	28 F, 48 M	328 ± 64	-0.96	2.2	1.9-2.5	2.3	Nevill et al.[120]
Speed (m.min^{-1})	10 M	307 ± 17	-0.82	2.7	2.0-4.3	3.3	Womack et al.[127] post-Fall
Speed (m.min^{-1})	10 M	304 ± 17	-0.77	3.0	2.2-4.8	3.8	Womack et al.[127] pre-Fall
Power (W)	22 F	243 ± 45	-0.77	3.9	3.2-5.3	5.1	Gillies & Bell[124]
Power (W)	10 M, 22 F	285 ± 46	-0.81	5.3	4.4-6.7	6.5	Gillies & Bell[124]
Power (W)	10 M	377 ± 72	-0.04	3.5	2.6-5.7	>10	Gillies & Bell[124]
Peak incremental VO$_2$							
VO$_2$max (L.min^{-1})	13 M	4.5 ± 0.4	-0.85	1.4	1.0-2.0	1.6	Cosgrove et al.[122]
VO$_2$max (L.min^{-1})	23 M, LW	5.05 ± 0.20	-0.70	1.3	1.0-1.7	1.7	Bourdin et al.[121]
VO$_2$max (L.min^{-1})	54 M	5.41 ± 0.42	-0.84	1.8	1.5-2.1	2.1	Bourdin et al.[121]
VO$_2$max (L.min^{-1})	48 M	5.60 ± 0.56	-0.82	2.0	1.7-2.4	2.5	Nevill et al.[120]
VO$_2$max (L.min^{-1})	10 M, 22 F	3.62 ± 0.84	-0.96	2.5	2.1-3.1	2.6	Gillies & Bell[124]
VO$_2$max (L.min^{-1})	22 F	3.19 ± 0.57	-0.92	2.4	1.9-3.2	2.6	Gillies & Bell[124]
VO$_2$max (L.min^{-1})	10 M	5.25 ± 0.69	-0.87	2.3	1.7-3.7	2.7	Womack et al.[127] post-Fall
VO$_2$max (L.min^{-1})	28 F, 48 M	5.02 ± 0.91	-0.94	2.7	2.4-3.1	2.8	Nevill et al.[120]
VO$_2$max (L.min^{-1})	31 M, HW	5.68 ± 0.32	-0.68	2.0	1.6-2.5	2.9	Bourdin et al.[121]
VO$_2$max (L.min^{-1})	10 M	5.28 ± 0.62	-0.84	2.5	1.9-4.1	3.0	Womack et al.[127] pre-Fall
VO$_2$max (L.min^{-1})	10 M	4.38 ± 0.42	-0.75	2.3	1.7-3.8	3.0	Gillies & Bell[124]
VO$_2$max (L.min^{-1})	8 M?	4.97 ± 0.48	-0.71	2.6	1.8-4.6	3.6	Faff et al.[123]
VO$_2$max (L.min^{-1})	28 F	4.03 ± 0.33	-0.74	2.8	2.3-3.5	3.7	Nevill et al.[120]
VO$_2$max (L.min^{-1})	10 M	4.85 ± 0.63	-0.76	3.1	2.2-5.0	4.1	Jurimae et al.[118]
VO$_2$max (L.min^{-1})	12 F	3.18 ± 0.35	-0.50	2.3	1.7-3.5	4.5	Riechman et al.[125]
VO$_2$max (ml.kg^{-1}.min^{-1})	22 F, 10 M	48.4 ± 7.4	-0.81	5.3	4.4-6.7	6.5	Gillies & Bell[124]
VO$_2$max (ml.kg^{-1}.min^{-1})	22 F	45.1 ± 5.9	-0.66	4.6	3.7-6.2	7.0	Gillies & Bell[124]
VO$_2$max (ml.kg^{-1}.min^{-1})	10 M	55.8 ± 4.3	-0.39	3.2	2.4-5.2	8.0	Gillies & Bell[124]
VO$_2$max (L.min^{-1})	19 M	4.6 ± 1.5	-0.43	3.7	3.0-5.1	8.4	Russel et al.[126]
VO$_2$max (ml.kg^{-1}.min^{-1})	12 F	47.4 ± 5.3	-0.11	2.7	2.1-4.2	>10	Riechman et al.[125]
VO$_2$max (ml.kg^{-1}.min^{-1})	10 M	61.6 ± 5.6	-0.13	4.8	3.5-7.7	>10	Jurimae et al.[118]
Submaximal lactate-related measures							

Power @ 4-mM (W)	10 M	275 ± 41	-0.96	1.3	0.9-2.1	1.4	Jurimae et al.[118]
Lactate @ 350 W (mM)	10 M	11.8 ± 4.8	0.96	1.3	0.9-2.1	1.4	Jurimae et al.[118]
VO2 @ 4-mM (L.min-1)	10 M	4.66 ± 0.75	-0.94	1.6	1.2-2.6	1.7	Womack et al.[127] pre-Fall
Power @ 1 mM above baseline (W)	12 F	138 ± 27	-0.82	1.5	1.1-2.3	1.8	Riechman et al.[125]
Speed @ 4-mM (m.min-1)	10 M	282 ± 17	-0.93	1.7	1.2-2.7	1.9	Womack et al.[127] post-Fall
Power @ 4-mM (W)	8 M?	222 ± 23	-0.89	1.7	1.2-2.9	1.9	Faff et al.[123]
%VO2max @ 4-mM (%)	31 M, HW	89.9 ± 5.2	-0.79	1.7	1.4-2.1	2.1	Bourdin et al.[121]
VO2 @ 1 mM above baseline (L.min-1)	12 F	2.24 ± 0.36	-0.77	1.7	1.3-2.6	2.2	Riechman et al.[125]
Speed @ 4-mM (m.s-1)	13 M	?	-0.73	1.8	1.4-2.7	?	Cosgrove et al.[122]
Speed @ 4-mM (m.min-1)	10 M	274 ± 18	-0.90	2.0	1.5-3.3	2.2	Womack et al.[127] pre-Fall
VO2 @ lactate inflection (L.min-1)	48 M	4.3 ± 0.5	-0.83	2.0	1.8-2.5	2.4	Nevill et al.[120]
VO2 @ 4-mM (L.min-1)	13 M	?	-0.68	2.0	1.5-2.9	?	Cosgrove et al.[122]
Power @ 4-mM (W)	28 F	256 ± 25	-0.84	2.2	1.8-2.8	2.6	Nevill et al.[120]
VO2 @ 4-mM (L.min-1)	10 M	4.13 ± 0.63	-0.87	2.4	1.7-3.9	2.7	Jurimae et al.[118]
Power @ 3-mM (W)	28 F	240 ± 24	-0.82	2.4	2.0-3.0	2.8	Nevill et al.[120]
Power @ 2-mM (W)	48 M	309 ± 36	-0.77	2.3	2.0-2.8	2.9	Nevill et al.[120]
Power @ 3-mM (W)	48 M	335 ± 40	-0.75	2.4	2.1-2.9	3.1	Nevill et al.[120]
Power @ 2-mM (W)	28 F	221 ± 24	-0.78	2.6	2.1-3.3	3.3	Nevill et al.[120]
VO2 @ 4-mM (L.min-1)	10 M	4.74 ± 0.71	-0.82	2.7	2.0-4.3	3.3	Womack et al.[127] post-Fall
Power @ 4-mM (W)	48 M	355 ± 42	-0.73	2.4	2.0-2.9	3.3	Nevill et al.[120]
Power @ 3-mM (W)	28 F, 48 M	300 ± 58	-0.92	3.1	2.8-3.6	3.3	Nevill et al.[120]
VO2 @ lactate inflection (L.min-1)	28 F, 48 M	3.8 ± 0.8	-0.92	3.1	2.8-3.6	3.3	Nevill et al.[120]
Power @ 2-mM (W)	28 F, 48 M	276 ± 54	-0.92	3.1	2.8-3.6	3.3	Nevill et al.[120]
Speed @ lactate inflection (m.s-1)	13 M	?	-0.39	2.5	1.9-3.7	?	Cosgrove et al.[122]
VO2 @ lactate inflection (L.min-1)	13 M	?	-0.39	2.5	1.9-3.7	?	Cosgrove et al.[122]
Power @ 4-mM lactate (W)	28 F, 48 M	319 ± 61	-0.91	3.3	2.9-3.8	3.6	Nevill et al.[120]
%VO2max @ 4-mM lactate (%)	54 M	90.0 ± 4.8	-0.49	2.8	2.4-3.3	5.7	Bourdin et al.[121]
VO2 @ lactate inflection (L.min-1)	28 F	3.0 ± 0.3	-0.45	3.7	2.1-4.8	8.0	Nevill et al.[120]
Submaximal VO2-related measures							
Gross efficiency (%)	23 M, LW	18.6 ± 0.8	-0.51	1.5	1.2-2.0	2.7	Bourdin et al.[121]
Gross efficiency (%)	31 M, HW	18.5 ± 1.0	-0.64	2.1	1.7-2.6	3.2	Bourdin et al.[121]
VO2 @ 4.00 m.s-1 (L.min-1)	13 M	3.29 ± 0.13	0.62	2.1	1.6-3.1	3.3	Cosgrove et al.[122]
VO2 @ 3.85 m.s-1 (L.min-1)	13 M	2.99 ± 0.16	0.51	2.3	1.7-3.4	4.3	Cosgrove et al.[122]
VO2 @ ?? W (ml.min-1.W-1)	28 F	15.73 ± 0.84	0.46	3.7	3.1-4.8	7.8	Nevill et al.[120]
VO2 Gross efficiency (%)	54 M	18.5 ± 0.9	-0.35	3.0	2.6-3.5	8.5	Bourdin et al.[121]
VO2 @ 3.70 m.s-1 (L.min-1)	13 M	2.66 ± 0.09	0.20	2.6	2.0-3.9	>10	Cosgrove et al.[122]
VO2 @ ? W (ml.min-1.W-1)	48 M	15.21 ± 0.86	0.02	3.6	3.1-4.3	>10	Nevill et al.[120]
VO2 @ ? W (ml.min-1.W-1)	28 F, 48 M	15.40 ± 0.88	0.33	7.4	6.6-8.6	>10	Nevill et al.[120]
Ventilatory-threshold measures							
VO2 (L.min-1)	22 F	2.41 ± 0.49	-0.72	4.3	3.4-5.7	5.9	Gillies & Bell[124]
VO2 (L.min-1)	22 F, 10 M	3.30 ± 0.64	-0.83	5.0	4.2-6.4	6.0	Gillies & Bell[124]
VO2 (L.min-1)	10 M	3.31 ± 0.47	-0.49	3.0	2.2-4.9	6.0	Gillies & Bell[124]

Power (W)	22 F	149 ± 31	-0.65	4.7	3.8-6.3	7.2	Gillies & Bell[124]
Power (W)	22 F, 10 M	170 ± 46	-0.74	6.0	4.9-7.6	8.1	Gillies & Bell[124]
Power (W)	10 M	215 ± 39	-0.13	3.4	2.6-5.6	>10	Gillies & Bell[124]
Anaerobic measures							
30-s Wingate minimum power (W)	12 F	358 ± 60	-0.89	1.2	0.9-1.9	1.3	Riechman et al.[125]
30-s Wingate mean power (W)	12 F	368 ± 60	-0.87	1.3	1.0-2.0	1.5	Riechman et al.[125]
30-s Wingate peak power (W)	12 F	380 ± 63	-0.85	1.4	1.0-2.2	1.6	Riechman et al.[125]
5-stroke all-out mean power (W)	48 M	596 ± 72	-0.82	2.0	1.7-2.4	2.5	Nevill et al.[120]
5-stroke all-out mean force (N)	48 M	738 ± 75	-0.81	2.1	1.9-2.6	2.5	Nevill et al.[120]
5-stroke all-out mean power (W)	28 F, 48 M	523 ± 117	-0.94	2.7	2.4-3.1	2.9	Nevill et al.[120]
5-stroke all-out mean force (N)	28 F, 48 M	662 ± 121	-0.93	2.9	2.6-3.4	3.1	Nevill et al.[120]
40-s all-out mean power (W)	10 M	614 ± 82	-0.76	3.1	1.6-3.3	4.1	Jurimae et al.[118]
5-stroke all-out mean force (N)	28 F	532 ± 54	-0.69	3.0	2.5-3.8	4.3	Nevill et al.[120]
5-stroke all-out mean power (W)	28 F	398 ± 55	-0.69	3.0	2.5-3.8	4.3	Nevill et al.[120]
30-s Wingate fatigue (%)	12 F	6.2 ± 4.8	0.24	2.6	2.0-4.1	>10	Riechman et al.[125]
Accumulated oxygen deficit (L)	19 M	2.1 ± 1.4	0.10	4.1	3.3-5.6	>10	Russel et al.[126]
Other measures							
Lactate 5 min post 2-km test (mM)	13 M	?	-0.58	2.2	1.7-3.2	?	Cosgrove et al.[122]
Lactate 5 min post VO2max test (mM)	13 M	?	-0.58	2.2	1.7-3.2	?	Cosgrove et al.[122]
Critical power (W)	8 M?	275 ± 28	-0.74	2.5	1.8-4.4	3.3	Faff et al.[123]
Heart rate @ VT (beats.min-1)	10 M	164.2 ± 9.4	0.54	2.9	2.1-4.7	5.3	Gillies & Bell[124]
Lactate @ VO2max (mM)	12 F	14.1 ± 2.7	-0.37	2.5	1.8-3.8	6.6	Riechman et al.[125]
Power @ 170 heart rate (W)	8 M?	242 ± 20	0.45	3.3	2.4-5.7	7.3	Faff et al.[123]
Heart rate @ VT (beats.min-1)	22 F	166 ± 13	0.08	6.2	5.0-8.3	>10	Gillies & Bell[124]
Heart rate @ VT (beats.min-1)	22 F, 10 M	166 ± 12	0.15	8.9	7.4-11.3	>10	Gillies & Bell[124]

Data for test measures are mean ± standard deviation. **90% CI** = 90% confidence interval for the SEE; **F** = female; **M** = Male; **HW** = heavyweight rowers; **LW** = lightweight rowers; **mM** = mmol.L^{-1}; **pre-Fall, post-Fall** = pre and post the autumn competitive season; **SEE** = standard error of the estimate; **VO$_2$** = oxygen uptake; **VT** = ventilatory threshold; **?** = not provided.

Appendix: Calculation of standard error of the estimate[129]

If X and Y are the practical and criterion in the validity study, r is their correlation, e_X and e_Y are their random errors, n is the sample size, SD is standard deviation, and SEE is standard error of the estimate, then:

$e_X = \sqrt{[SD_X^2(1-r^2SD_Y^2/(SD_Y^2-e_Y^2))]}$;
observed slope of regression line = $r(SD_Y/SD_X)$;
observed SEE = $SD_Y\sqrt{[(1-r^2)(n-1)/(n-2)]}$;

true slope = (observed slope)/$(1-e_X^2/SD_X^2)$;
true SEE without criterion error = (true slope)e_X;
true SEE with criterion error = $\sqrt{[(\text{true SEE})^2 + e_Y^2]}$.

The adjusted SEE shown in the tables is the true SEE with criterion error. The random error in the criterion, e_Y, was assumed to be 1% for 2000-m single-scull performance time (see Table 2.1) and 0.5% for 2000-m Concept II performance time (see Table 2.3).

This approach cannot be used for measures derived by multiple linear regression unless the authors provide the mean and SD of the predicted values.

CHAPTER 3

Are There Useful Physiological or Psychological Markers for Monitoring Overload Training in Elite Rowers?

Running head: Monitoring of overload training

Overview

There is a need for markers that would help determine when an athlete's training load is either insufficient or excessive. In this study we examined the relationship between changes in performance and changes in physiological and psychological markers during and following a period of overload training in 10 female and 10 male elite rowers. Change in performance during a 4-wk overload was determined with a weekly 30-min time-trial on a rowing ergometer, while an incremental test provided change in lactate-threshold power between the beginning of the study and following a 1-wk taper after the overload. Various psychometric, steroid-hormone, muscle-damage, and inflammatory markers were assayed throughout the overload. Plots of change in performance versus the 4-wk change in each marker were examined for evidence of an inverted-U relationship that would characterize under-training and excessive training. Linear modeling was also used to estimate the effect of changes in the marker on changes in performance. There was a suggestion of an inverted U only for performance in the incremental test versus some inflammatory markers, due to the relative under-performance of one rower. There were some clear linear relationships between changes in markers and changes in performance, but relationships were inconsistent within classes of markers. For some markers, changes considered to predict excessive training (e.g., creatine kinase, several pro-inflammatory cytokines) had small to large *positive* linear relationships with performance. In conclusion, some of the markers investigated in this study may be useful for adjusting the training load in individual elite rowers.

Introduction

A period of intense training prior to a competition is an important phase in the preparation of elite athletes in high-intensity sports.[1] The optimum training load for this phase could be defined by the apex of an inverted-U relationship between training and subsequent performance: below the optimum load the lower training stimulus results in less gain in performance, whereas the stress of training above the optimum load results in the maladaptation that researchers are now referring to as non-functional overreaching that can lead to overtraining syndrome.[6] Direct determination of an athlete's optimum load is therefore desirable in theory but logistically impossible in practice: it would require systematic manipulation of training load before each competition in a series of competitions. In any case, the optimum load might change in an unpredictable way between competitions or even between training sessions, owing to the modifying effects of diet, lifestyle, illness, injury, and psychological state.

Sport scientists nevertheless nurture the hope that there is an indirect way to train athletes close to their optimum load. The hope is based on the notion that there are physiological or psychological markers that are closely associated with whatever mechanism is responsible for maladaptation. For example, systemic inflammation arising from illness or from tissue damaged by training or injury is a possible cause of overtraining syndrome,[36] so the level of some marker of systemic inflammation (such as IL1beta, IL6 and TNFalpha) might have an inverted-U relationship with performance. If the relationship is essentially the same in each athlete, the optimum level of the marker could be identified in a single study of a sample of athletes monitored through an intense phase of training. Here we have performed such a study with a squad of elite rowers performing four weeks of intense overload training in preparation for a world championship. We chose to explore markers that could also be analyzed quickly enough to help the coach make decisions on whether to modify subsequent training sessions. With this consideration, and on the basis of previous research, the potential markers of maladaptation we monitored were cortisol, testosterone, dehydroepiandrosterone (DHEA),[20, 130, 131] creatine kinase (CK), lactate dehydrogenase (LDH),[132] mood state, sleep quality, perception of fatigue,[133-135] C-reactive protein (CRP) and a suite of cytokines.[36, 136]

Methods

Subjects and Design

Ten females and 10 males from the New Zealand elite rowing squad, of whom 11 were current or former world champions, gave informed consent to participate in accordance with requirements of the AUT ethics committee. The mean (and range) for height, weight, and age of the females were 178.0 (176-182) cm, 74 (69-93) kg, and 23 (19-31) y; those of the males were 191 (181-200) cm, 90 (72-101) kg, 24 (21-30) y.

Markers were monitored during a 4-wk overload period (Figure 3.1), during March and April 2007. The rowers performed a stepwise lactate test at the beginning of the study and after a 1-wk mini taper following the 4-wk overload. They also performed a 30-min rowing ergometer test near the end of each week of the overload (midday Friday). Data from performance tests in a following overload period (April-May) and from similar overload phases in the previous 6 years were available and also contributed to the analysis.

The rowers undertook ~12 aerobic rowing sessions and two weight training sessions per week, and volume of rowing training increased by ~10% each week during the overload, with a ~20% reduction in volume during the subsequent mini taper. The prescribed training plan was adhered to reasonably rigidly throughout the study. The coaches did not give permission to release other details of training and individual test results.

Figure 3.1 Timeline of 4-mM and 30-min performance tests, markers of inflammation and muscle damage (M: cytokines, CRP, CK & LDH), steroid hormones (H: cortisol, testosterone & DHEA) and weekly recalls (Ψ: mood, fatigue & sleep).
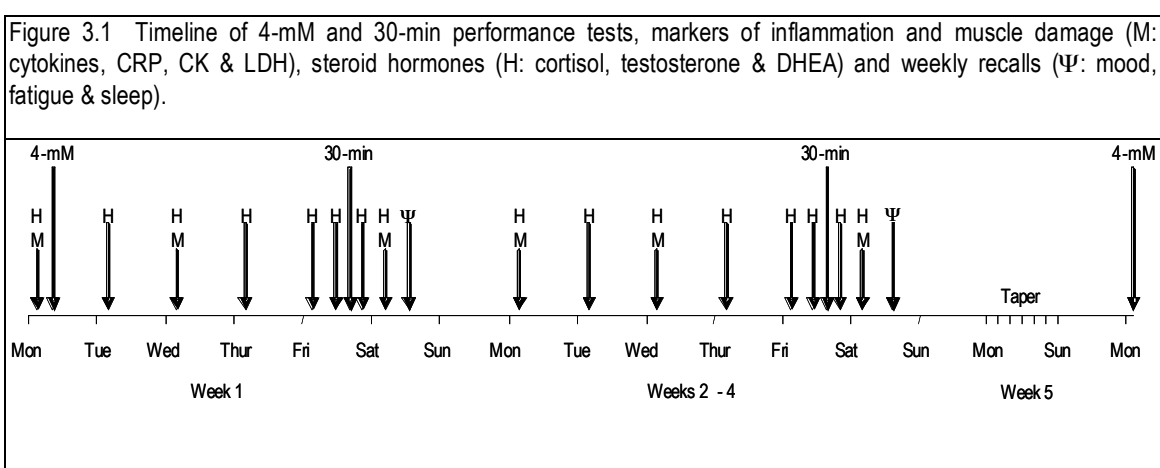
Saliva Sampling

Saliva sampling was undertaken at ~0700 each morning Monday to Saturday and immediately before and after the 30-min maximal rowing ergometer test undertaken at midday on Friday. Saliva production was stimulated by giving the rowers Wrigley's sugar-free gum to chew. To ensure adequate flow and minimize contamination the gum was chewed for ~30 s and this initial saliva swallowed, then with continued chewing 3-5 ml of saliva was collected into a labeled 10 polyethylene centrifuge tube. The saliva samples were frozen at -20°C until analysis. On the first Monday and all subsequent Saturdays a sub-sample was stored at -80°C for cytokine analysis.

Blood Sampling

A capillary blood sample (0.12-0.25 ml) was taken from the ear lobe at ~0700 three times a week. Ear lobes were sampled using a standard lancet to prick the ear lobe, followed by collection of the blood into heparinized capillary tubes. The tubes were immediately centrifuged and the plasma separated and stored at 4-8°C, and then analyzed within 24 hours for lactate dehydrogenase and creatine kinase activity. On the first Monday and all subsequent Saturdays a sub-sample was stored at -80°C for cytokine analysis.

Performance Tests

All rowers were familiar with both performance tests that had been implemented for at least six years prior to the study and each rower had at least four years experience with each. The stepwise lactate-threshold tests were conducted on a Concept IIb rowing ergometer (Concept2, Morrisville, VT) on the first day of the study (Monday, Week 1) and on the last day of the training block (Monday, Week 6). All rowers had at least four days rest from all rowing prior to completing the first test and the second test was completed after a 5.5-d taper followed by a 1.5-d rest. This test involves a 6-min step followed by a 1-min rest, during which blood lactate was sampled, with each subsequent step increasing by 15 W until the lactate concentration was above 6 mM. All rowers were provided with recommendations on appropriate pre-test nutrition and blood glucose concentration was also determined with the final blood lactate sample to examine whether any rower fell outside the normal blood glucose concentrations of 4-8 mM. The power corresponding to 4 mM lactate (4-mM power) was determined visually from a scatter chart generated in Microsoft Excel with the data points connected by smoothed lines.

A 30-min maximal rowing ergometer test for which the rating was restricted to 18 strokes per minute was conducted at midday every Friday during the 4-wk overload period. The results for these team sessions were used by the coaches to rank all rowers, which engendered a highly competitive environment and motivated the rowers to give near-maximal performances in all these tests.

Psychometrics

At the end of each training week (Saturday afternoon) after the final training session the rowers were asked to recall the number of hours of sleep and number of times they awoke each night over the previous 7 d. For the same time frame they were also asked to recall their perceived daily fatigue levels when awaking and prior to bed using a 5-point scale (1, not at all; 2, a little; 3, moderate; 4, quite a lot; 5, very much). Owing to concerns over the length and effectiveness of the Profile of Mood States in predicting performance maladaptation,[49] we used the Positive and Negative Affect Score (PANAS)[52] to assess mood state at the time of administration on the Saturday afternoon. A recall was used because pilot work demonstrated unacceptably poor compliance with a diary; also the team physician was already using a similar recall to assess for excessive chronic fatigue.

Saliva and Blood Assays

Saliva samples were analyzed in triplicate for cortisol, testosterone and DHEA using radioimmunoassay (RIA). The methods were modified from those described by Granger at al.[137, 138] and Morelius et al.[139] Plasma samples were analyzed for creatine kinase using the IFCC primary reference procedure.[140] Lactate dehydrogenase was determined by the method described by Howell et al.[141] Plasma and saliva samples were collected for measurement of 14 cytokine/chemokines (IFN-g, IFN-a, IL-1b, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, IL-18, TNF-a, TNF-b and MCP-1) and CRP levels. They were determined simultaneously using a multiplexing bead assay from a standardized kit (BMS810FF, FlowCytomix, Bender MedSystems, Vienna, Austria) consisting of a set array of cytokines that included those proposed to relate to non-functional overreaching and overtraining syndrome.[36, 136]

Statistical Analyses

We log-transformed performance and used back-transformation to express the changes in percent units.[71] Change in performance in the stepwise lactate test was the change in 4-mM power between the beginning of the study and after the taper. The 30-

min test was performed before the taper so changes were better estimated by a measure of linearized mean change over the four weekly tests. The mean change in 30-min performance was derived by fitting a straight line to the log of the weekly values, then back-transforming the difference between the predicted values for the fourth and first test. Similar linearized changes were derived from the weekly values of the markers, via log transformation for those representing concentrations. For markers assayed on multiple occasions each week, mean values for each week and linearized changes for each week were derived before further analysis of the linearized change (in the mean and in the weekly change) over the four weeks. These analyses were performed using the Statistical Analysis System (Version 9.1, SAS Institute, Cary, NC).

The relationships between changes in performance and changes in markers were assessed first by plotting change scores for performance in percent units plotted on a linear scale vs those for markers expressed as factors plotted on a log scale. The plots were examined for any indication of an inverted-U relationship or an association between extreme values of a marker and impairment or poor improvement in performance. We also assessed the overall direction of the relationship between changes in a marker and performance by fitting straight lines to the change scores. The magnitude of the linear effect of the marker on performance was evaluated as the difference in performance change associated with a difference of two standard deviations (2 SD) of change in the marker.[71] Change in performance was expressed with 90% confidence limits. In keeping with the exploratory nature of this study we did not adjust for inflation of error when declaring effects to be clear. Inferences about the true (large-sample) values of effects were based on interpreting the magnitudes of observed value, the lower confidence limit and the upper confidence limit in relation to thresholds for small, moderate and large effects, assumed to be 0.3, 0.9 and 1.6 of within-athlete variability in performance between competitions.[71] In the absence of published values for rowers, we assumed the variability to be similar to that of kayakers[66] and cyclists:[68] ~3% for mean power. The thresholds for small, moderate and large effects on mean power were therefore 1%, 3% and 5.3% .

Error of measurement in 4-mM lactate power was estimated with a spreadsheet (available at newstats.org/xrely.xls) by treating the tests before and after the overload as trials in a reliability study; the error is then given by the standard deviation of the change scores divided by $\sqrt{2}$. This error analysis was performed for the monitored overload (March-April) and the following overload (April-May). The additional individual differences in the response to the first overload relative to the second

overload were estimated as the square root of the differences in the squares of the change scores.

Error in a single measurement of 30-min mean power was estimated by treating the four tests as trials in a reliability study and averaging the consecutive pairwise estimates of error of measurement. The contribution of this error to the error in the linearized change in 30-min mean power was estimated by simulation with a spreadsheet as follows: four random, normally distributed test scores with this error were generated for each of 1000 imaginary rowers, the linearized change score was calculated for each rower, then the SD of these change scores was divided by $\sqrt{2}$. Individual differences in the observed linearized change in the overload were estimated as the square root of the differences in the squares of the observed and simulated change scores.

Results

Performance

Baseline values and changes in the performance tests are shown in Table 3.1. There were small overall improvements in 30-min mean power in the overload for females and males. The 1-wk error of measurement in 30-min mean power derived from the reliability analysis of the four tests was 2.0% and 2.5% for females and males respectively (90% confidence limits, $\times\div1.30$). In the simulations with these errors, the SD of linearized change over the four tests were 2.7% for females and 3.3% for males. The corresponding observed SD for females in Table 3.1 is slightly smaller (2.4%), indicating no real individual differences in the changes, but the observed SD for males (4.7%) implies an SD representing individual differences of $\sqrt{(4.7^2-3.3^2)} = 3.3\%$.

Blood-glucose concentration for all rowers in the 4-mM tests was within the normal range of 4-8 mM. The improvements in 4-mM power were much larger than those for 30-min power (Table 1). Improvements in the next overload phase (April-May) were more modest (percent mean $\pm$ SD: 1.5 $\pm$ 2.0 and 0.7 $\pm$ 3.1 for females and males respectively). There was a similar pattern of improvement for the two overload phases averaged over the previous seven years (percent mean $\pm$ SD: 6.3 $\pm$ 5.2 and 5.5 $\pm$ 5.0 for March-April vs 2.2 $\pm$ 3.5 and 1.7 $\pm$ 3.8 for April-May). Errors of measurement estimated from the SD for April-May 2007 were 1.4% for females 2.2% for males (90% confidence limits, $\times\div1.50$). When these errors were removed from the SD for the changes in the March-April overload shown in Table 3.1, individual differences in the changes were represented by SD of 3.9% for females and 3.7% for males.

Table 3.1 Initial values (Week 1) and changes in performance in the time trials (30-min mean power) and incremental tests (4-mM lactate power).

	Initial (W)	Change (%)	
	Mean ± SD%	Mean ± SD; ±CL	Inference
Females (n=10)			
30-min mean power	212 ± 6%	1.1 ± 2.4; ±1.4	Small ↑
4-mM lactate power	234 ± 10%	8.3 ± 4.3; ±2.5	Large ↑
Males (n=10)			
30-min mean power	296 ± 13%	1.9 ± 4.7; ±2.8	Small ↑
4-mM lactate power	318 ± 14%	8.4 ± 4.5; 2.6	Large ↑

Change scores for 30-min mean power are linearized estimates of change in the four test of the overload; those for 4-mM lactate power are Week 6 minus Week 1.

SD%, standard deviation express as percent of the mean.

CL, 90% confidence limits for the mean, expressed in "±" form.

Markers and Relationship with Performance

Only three inflammatory markers showed a hint of an inverted-U relationship with performance, and only for male rowers in the incremental test (Figure 3.2). This relationship arose because of one rower who consistently appeared in the bottom right quadrant. Despite this rower's relative underperformance and large increases in these inflammatory markers, he maintained the normal training regime and went on to record a series of career-best results for that season. Another rower was diagnosed as overtrained by the team physician in response to the rower's reports of sleep and psychological state. However this rower was apparently suffering only from acute fatigue, because s/he showed above average gains in both performance tests, no abnormal measures in any of the markers, and some months later performed successfully in the world rowing championships.

Initial values of and changes in markers during the overload training are presented in Tables 3.2 and 3.3. Statistics representing the linear relationships between changes in performance and markers are presented in Table 3.4. In summary, change in 30-min performance for females had substantial positive linear relationships with change in post-pre 30-min testosterone and DHEA, morning and evening fatigue, weekly change in sleep, IL5 saliva, MCP1 plasma, MCP1 saliva, TNFalpha plasma and TNFbeta plasma, while sleep and weekly change in morning fatigue had negative relationships; for males the markers with substantial positive relationships were post-pre 30-min

testosterone, weekly change in LDH, PANAS negative, IL4 plasma, IL5 saliva and IL8 plasma, while MCP1 saliva had a substantial negative relationship. Changes in 4-mM power for females had positive relationships with changes in morning CK, weekly change in wakeups, IL5 Plasma and TNFbeta plasma and negative relationships with weekly change in testosterone and IL8 plasma; for males, only morning CK displayed a positive relationship, whereas negative relationships were observed for post-pre 30-min cortisol, evening fatigue, weekly change in sleep, CRP saliva, IFNalpha saliva, IFNgamma saliva, IL10 saliva, IL18 plasma, IL2 saliva and IL4 saliva.

Figure 3.2   Change in 4-mM power vs change in inflammatory markers that illustrate a possible inverted-U relationship for males (•) but not females (o). Arrows indicate the one rower who had the largest changes in the markers.
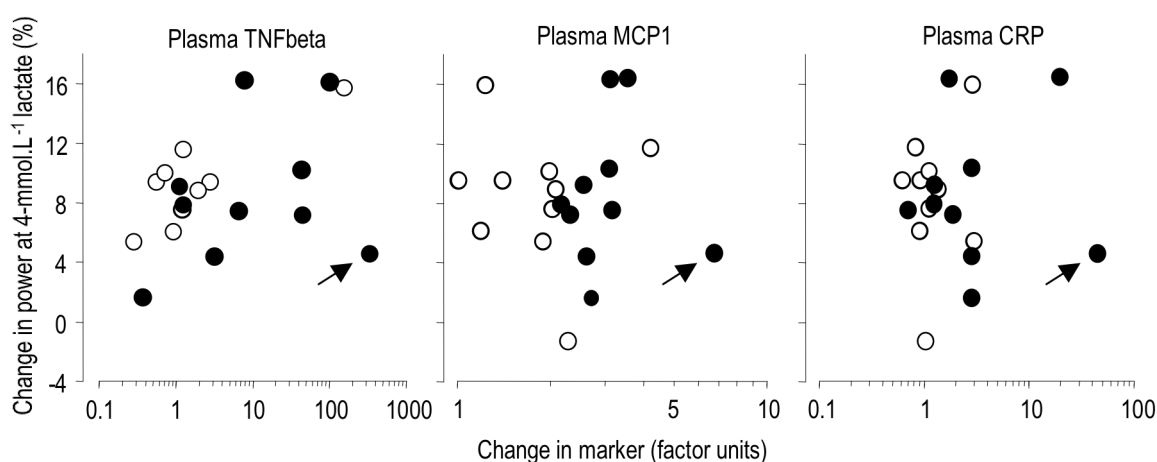
Discussion

In the present study the few instances of a possible inverted-U relationship were due to one under-performing rower with large increases in inflammatory markers who subsequently recorded his best-ever race performances. There was little evidence of a consistent association between relatively poor performance and extreme values of physiological and psychological variables that are regarded as potential markers of maladaptation. The only measure with a clear association for both males and females was the positive relationship for creatine kinase, suggesting that targeting training to produce increases in this marker will produce greater performance enhancements. Amongst the remaining markers we expected fatigue-related declines in performance associated with hard training to be accompanied by extreme values of some stress markers, as we have seen for example in the blunting of the exercise-associated increase in cortisol with males (but not females). Our failure to consistently observe such

negative relationships has several possible explanations and implications for the monitoring of elite athletes training hard before a competition.

First, it is possible that some rowers exceeded their optimum training load and had large increases in values of some markers, but differences between individuals in what constitutes a large increase meant that an inverted U would not be observed in our plots of change scores. In this scenario, monitoring of one or more of the markers we investigated might still be useful for detection and possibly prevention of non-functional overreaching or overtraining syndrome, but each athlete's usual range of values of the marker and the association between the marker and concurrent performance would have to be established during a longer period of monitoring that included at least one occurrence of non-functional overreaching and/or overtraining syndrome. It might be possible to obtain such data over a period of several years of intensive monitoring. Other markers we did not investigate, such as heart-rate variability[61, 142] and heart-rate recovery,[143] might show less individual differences in their relationship with performance and thereby demonstrate their potential for monitoring overload training and detecting non-functional overreaching or overtraining syndrome in a study similar to this one.

Secondly, errors in measurement of a marker might have masked the underlying relationship between the marker and performance. This problem may well apply to the salivary cytokines, which we included along with the salivary steroid hormones for their potential as non-invasive physiological markers of stress. The other markers were less problematic: salivary concentrations of steroid hormones are known to reflect blood concentrations; we used standard assays for cytokines and muscle-damage markers in blood; the PANAS is an acceptable instrument for measuring mood state; and our simple questionnaire for assessing sleep and fatigue ought to be suitable for tracking substantial changes during the overload. If all of these measures nevertheless had measurement problems, we would have to conclude that current approaches to practical monitoring of these markers of stress in athletes are unlikely to provide useful information.

Table 3.2 Initial values and changes in physiological markers.

	FEMALE						MALE					
	Initial		4-wk change (%)				Initial		4-wk change (%)			
	Mean	SD (%)	Mean	SD	CL	Infer-ence	Mean	SD (%)	Mean	SD	CL	Infer-ence
STEROID HORMONES												
Testosterone: (pg/ml)	24	82	-6.7	18	9	trv ↓	83	34	1.3	11	6	trv ↑
Cortisol: (ng/ml)	5.6	44	1.5	17	9	trv ↑	4.4	34	8.2	22	12	sm ↑
DHEA: (ng/ml)	3.2	93	-2.1	25	12	trv						
Test: within-wk chg	0.92	23	12	27	13	sm ↑	1	7.8	4.8	7.2	4.1	sm ↑
Cort: within-wk chg	1.1	31	1.4	31	15		1.3	55	-11	43	23	sm ↓
DHEA: within-wk chg	0.97	19	2.1	23	11							
Test: Pre 30-min (pg/ml)	19	89	-7.2	35	17	trv ↓	69	30	2.4	20	12	
Cort: Pre 30-min (ng/ml)	1.9	79	-8.6	69	31		1.5	44	-3.4	86	47	
DHEA: Pre 30-min (ng/ml)	3.0	107	6.0	55	25							
Test: post/pre 30-min	1.7	24	-18	33	16	mod ↓	1.6	40	-12	44	25	
Cort: post/pre 30-min	1.9	66	0.8	63	29		2.1	62	35	82	45	sm ↑
DHEA: post/pre 30-min)	1.9	36	-1.6	55	26							
MARKERS OF MUSCLE DAMAGE												
CK: (U/L)	181	63	-29	53	24	mod ↓	354	42	-14	46	24	mod ↓
LDH: (U/L)	295	17	-23	25	12	lrg ↓	274	14	-18	14	8.0	lrg ↓
CK: within-wk chg	1.4	20	-31	36	17	lrg ↓	1.4	19	-28	17	9.3	
LDH: within-wk chg	0.94	16	10	24	12	mod ↑	0.9	10	13	25	14	mod ↑
CYTOKINES AND CRP												
CRP: plasma (ng/ml)	15163	131	28	70	32	sm ↑	11552	72	205	265	112	mod ↑
CRP: saliva (ng/ml)	42	182	57	297	104		40	33	172	187	92	mod ↑
IFNalpha: plasma (pg/ml)	4.3	193	-36	146	59	sm ↓	2	530	149	377	147	mod ↑
IFNalpha: saliva (pg/ml)	1.4	129	89	161	65	mod ↑	1.5	259	184	197	97	lrg ↑
IFNgamma: plasma (pg/ml)	17	125	-27	81	36	sm ↓	9.1	379	111	127	61	mod ↑
IFNgamma: saliva (pg/ml)	6	200	53	242	89	sm ↑	8.4	237	36	184	91	
IL1beta: plasma (pg/ml)	17	78	-41	99	43	mod ↓	10	153	134	194	87	lrg ↑
IL1beta: saliva (pg/ml)	25	73	-3.4	72	32		36	77	41	66	37	mod ↑
IL2: plasma (pg/ml)	13	563	-71	1511	322	mod ↑	5.2	16	301	706	235	mod ↑
IL2: saliva (pg/ml)	11	428	21	823	217		6.1	15	144	471	195	mod ↑
IL4: plasma (pg/ml)	352	353	-75	1400	307	mod ↓	379	32	302	628	216	mod ↑
IL4: saliva (pg/ml)	12	345	-0.5	743	202		6	39	140	494	202	mod ↑
IL5: plasma (pg/ml)	123	108	-9.0	138	57		65	254	212	334	134	mod ↑
IL5: saliva (pg/ml)	69	198	12.2	182	71		60	153	70	116	61	mod ↑
IL6: plasma (pg/ml)	136	12	-7.1	23	11		127	215	31	23	12	lrg ↑
IL6: saliva (pg/ml)	135	12	-1.0	15	7.4		148	89	6.0	19	12	
IL8: plasma (pg/ml)	21	66	6.6	86	38		17	92	73	52	27	lrg ↑
IL8: saliva (pg/ml)	38	54	0.1	78	35		56	121	40	63	36	mod ↑
IL10: plasma (pg/ml)	5.3	229	-43	268	96		3.1	50	268	468	174	mod ↑

IL10: saliva (pg/ml)	6.2	182	1.4	341	116		5.9	77	108	167	84	mod ↑
IL12p70: plasma (pg/ml)	7	189	-40	481	149		2.5	51	318	604	210	mod ↑
IL12p70: saliva (pg/ml)	8.6	218	3.8	251	92		7.8	55	142	151	71	lrg ↑
IL18: plasma (pg/ml)	264	51	14	67	30		175	726	196	102	50	lrg ↑
IL18: saliva (pg/ml)	120	127	-41	123	51	sm ↓	141	335	-5.8	120	63	
MCP1: plasma (pg/ml)	147	55	69	53	25	mod ↑	143	72	204	38	20	vlrg ↑
MCP1: saliva (pg/ml)	44	57	-18	68	31	sm ↓	68	33	52	101	54	mod ↑
TNFalpha: plasma (pg/ml)	12	61	-26	82	36	sm ↓	7.8	530	110	151	70	lrg ↑
TNFalpha: saliva (pg/ml)	10	46	3.5	55	25		11	259	50	54	31	mod ↑
TNFbeta: plasma (pg/ml)	20	919	62	459	144		9	379	845	805	258	mod ↑
TNFbeta: saliva (pg/ml)	5	469	-19	607	176		4	237	-64	2,148	589	

Initial data are the Week 1 mean and SD(%). Change scores are linearized change over four tests (Weeks 1 to 4). The unit of measure for the mean, SD and 90% confidence limits (CL) for the change scores are percentages.

Test, Testosterone; Cort, Cortisol; within-wk chg, the linearized factor change from Monday to Friday determined from each individual's line of best fit for each day of the week; Pre 30-min, mean value for markers collected immediately before the 30-min test; Post/Pre 30-min, difference between the mean value for the marker collected immediately post the 30-min test and the mean value collected immediately pre the 30-min test.

Inference is the qualitative assessment of the magnitude of any clear effect.[132] Descriptors for the inferences: trv, trivial; sm, small; mod, moderate; lrg, large; vlrg, very large. Blank inferences are unclear.

Thirdly, there may have been underlying inverted-U relationships between markers and performance, but we failed to observe them because of problems with our measures of performance. Competitive performance is the criterion measure for assessing overreaching and overtraining, but monitoring of on-water performance even in staged time-trials is currently not practical with rowers in crewed boats. Ergometer performance, especially in a simulated 2000-m time trial with a Concept II, is considered a reliable and valid measure of on-water race performance.[78, 80, 144] Unfortunately the coaches in this study would not allow this form of assessment, owing to concerns that the stress of repeated maximal sessions might hinder aerobic development in this phase of training. They use instead the 30-min test, which in their view not only assesses but also develops aerobic power. The 4-mM test is considered an effective method of measuring changes in aerobic conditioning,[145, 146] and coaches include it because the measure of performance it provides does not depend on motivation or pacing. Both tests had been used in the program for at least six years, and during that time correlations between performance in these tests and in a 2000-m ergometer time-trial conducted within a week or two have ranged from 0.88 to 0.93 for males and females analyzed separately (TB Smith, unpublished observations).

Table 3.3 Initial values and changes in psychological markers.

| | FEMALE | | | | | | MALE | | | | | |
| | Initial | | 4-wk change | | | | Initial | | 4-wk change | | | |
	Mean	SD	Mean	SD	CL	Infer-ence	Mean	SD	Mean	SD	CL	Infer-ence
PSYCHOMETRIC MEASURES												
PANAS: positive (10-50)	35	6.0	-0.6	11.0	5.6		35	6.5	-8.7	6.1	6.5	lrg ↓
PANAS: negative (10-50)	18	6.3	0.3	7.7	3.9		16	2.8	4.6	4.9	6.6	lrg ↑
Sleep (h)	8.2	0.7	-0.8	1.1	0.6	lrg ↓	7.9	0.9	0.2	0.7	9.6	
Wakeups (per night)	1.1	1.1	0.5	1.6	0.9		0.9	0.9	0.0	1.1	8.3	
Fatigue: morning (1-5)	2.5	0.5	2.2	1.1	0.6	vlrg ↑	2.4	0.7	1.0	0.9	9.4	lrg ↑
Fatigue: evening (1-5)	3.4	0.6	1.7	1.2	0.7	lrg ↑	3.1	1.2	1.2	1.1	5.3	mod ↑
Sleep: within-wk chg (h)	-0.24	1.4	0.2	2.3	1.3		-0.81	0.6	0.4	1.2	6.1	
Wakeups: within-wk chg (per night)	0.54	1.0	-0.4	1.1	0.6	sm ↓	0.56	1.2	0.0	1.5	8.1	
Fatigue: within-wk chg morning (±4)	0.71	0.8	0.1	0.6	0.3		0.31	1.2	0.0	1.6	12	
Fatigue: within-wk chg evening (±4)	0.57	0.8	0.3	1.4	0.8		0.80	0.5	-0.1	1.5	9.3	

Initial data are the Week 1 mean and SD. Change scores are linearized change over four tests (Weeks 1 to 4). All measures are displayed as raw values.

Within-wk chg, the linearized change from Monday to Friday determined from each individuals line of best fit for each day of the week.

Inference is the qualitative assessment of the magnitude of any clear effect.[71] Descriptors for the inferences: trv, trivial; sm, small; mod, moderate; lrg, large; vlrg, very large. Blank inferences are unclear.

The large changes in 4-mM power for some rowers (up to 16%) raise the issue of whether lactate production was suppressed in the post-test and/or elevated in the pre-test. Normal blood glucose values combined with an ample pre-test rest period for both tests provide some confidence that the changes reflect changes in endurance performance rather than acute changes in glucose metabolism.[147] Examination of 4-mM test results over the previous six years suggests that changes of this magnitude are the norm, a situation probably created by the two months of intensive anaerobic training that precedes this training phase.

Taken together, these arguments lead us to the reasonably confident conclusion that our failure to observe consistently high values of markers in rowers who responded poorly to the overload was not a consequence of measurement issues with performance. We are left with the final and in our view mostly likely explanation of our results: there were no instances of non-functional overreaching or overtraining syndrome with these rowers. Performance in the 30-min test is consistent with lack of non-functional overreaching in the females, who all appeared to improve by a similar small amount during the overload. The males also improved on average by a small amount in this

test, but some of them apparently deteriorated. Their decline in performance is consistent with overreaching, but we suspect that these rowers showed a decline either because they did not train hard enough or (more likely) because they put less effort into the test towards the end of the overload. The subsequent performance history of the females and males is not consistent with any long-term under-performing that would qualify as overtraining syndrome. The team physician diagnosed one of the rowers as overtrained, but that rower subsequently performed well in competitions.

The head coach of this elite rowing squad has a reputation for setting very hard training programs, so we were expecting a reasonable proportion of the rowers to show signs of functional overreaching towards the end of the overload, with perhaps one or two rowers showing signs of non-functional overreaching after the taper. This rowing program also has a reputation for achieving outstanding performance, with nine rowers from this study winning medals at the 2007 world rowing championships. If we accept that there was little or no non-functional overreaching and no overtraining syndrome, it appears that the program is based on a gradual improvement in performance in the overload phase rather than deterioration from the effects of accumulated fatigue. These successful elite rowers may also be survivors who do not experience non-functional overreaching or overtraining syndrome, no matter how hard the training, without the additional stress of a chronic infection, overuse injury, or psychological trauma. The coaches may well be sufficiently attuned to the behavior and demeanor of their rowers to reduce the training of individual rowers who betray signs of such stress. Whether our finding of a modest positive relationship between change in performance and change in creatine kinase would be useful to such coaches is worthy of further investigation. If the increased creatine kinase was due to increased weight training, and weight training did not contribute to performance enhancement, then creatine kinase would not be a useful marker: as its increase would only be coincidental to performance improvements. Further research is required to clarify this issue, especially given the suspicion of the head coach that weight training does not benefit performance in elite rowers. In the meantime, our advice to sport scientists associated with elite athletes is that a large increase in creatine kinase and possibly other physiological and psychological stress markers, may be more indicative of an effective training overload than impending maladaptation.

Table 3.4 Percent change in performance (mean; ± 90% confidence limits) for the increment lactate test (4-mM power) and the 30-min ergometer test (mean power) associated with 2-SD difference between rowers. Data are shown only for makers with clear linear associations.

	FEMALE		MALE	
	4-mM power	30-min power	4-mM power	30-min power
STEROID HORMONES				
Testosterone: post/pre 30-min		2.3; ±2.1		5.0; ±5.9
Cortisol: post/pre 30-min			-7.0; ±7.0	
DHEA: post/pre 30-min		1.3; ±2.3	-	-
Testosterone: within-week change	-6.6; ±4.2			
MARKERS OF MUSCLE DAMAGE				
CK	5.5; ±4.5		6.4; ±4.3	
LDH: within week-change				7.2; ±4.6
CYTOKINES AND CRP				
CRP: saliva			-5.3; ±4.3	
IFNalpha: saliva			-6.5; ±5.4	
IFNgamma: saliva			-5.3; ±4.3	
IL2: saliva			-5.0; ±4.5	
IL4: plasma				4.7; ±5.5
IL4: saliva			-5.7; ±4.1	
IL5: plasma	4.5; ±4.9			
IL5: saliva		2.4; ±2.4		5.1; ±5.4
IL8: plasma	-5.3; ±4.0			6.5; ±4.7
IL10: saliva			-5.7; ±4.1	
IL18: plasma			-4.5; ±4.7	
MCP1: plasma		2.1; ±2.5		
MCP1: saliva		1.7; ±2.6		-4.4; ±5.1
TNFalpha: plasma		1.6; ±2.6		
TNFbeta: plasma	4.8; ±2.4			
TNFbeta: saliva		1.8; ±2.8		
PSYCHOMETRICS				
PANAS: negative				5.0; ±5.9
Sleep		-2.3; ±2.2		
Fatigue: morning		3.6; ±1.7		
Fatigue: evening		3.1; ±2.2	-6.6; ±5.3	
Sleep: within-week change		2.2; ±2.4	-7.4; ±6.1	
Wakeups: within-week change	3.3; ±3.4			
Fatigue: within-week change		-2.2; ±2.3		

Within-week change, the linearized change from Monday to Friday determined from each individual's line of best fit for each day of the week; post/pre 30-min, ratio of the value for the marker collected immediately post the 30-min test to that immediately pre the test.

CHAPTER 4

Heart-Rate Variability and Psychological Stress in an Elite Female Rower who Developed Overtraining Syndrome: Case Report

Running Head: Overtraining and heart-rate variability

Key words: Rowing, overtraining syndrome, heart rate variability, psychological stress

Overview

Coaches of elite athletes would benefit from simple markers that predict overtraining syndrome. Despite extensive research efforts the search continues for a marker that is both sensitive and specific enough to consistently predict overtraining syndrome.[6] Amongst candidate markers are measures of heart-rate variability,[59, 61, 148, 149] psychological state,[6] and performance itself.[6] Here we describe changes in these measures during a period of overload training that resulted in overtraining syndrome of one member of a squad of elite athletes. Our findings lead us to doubt whether any of these measures alone can be trusted for decisions about early intervention to prevent overtraining syndrome, but in combination they may be useful.

Methods

Ten elite female rowers from the New Zealand Senior-A rowing squad preparing for international competition were monitored during a 4-wk overload period in accordance with ethical approval of the Auckland University of Technology. Their mean height, weight and age were: 177.7 (range 175-181.5) cm, 74.9 (68.2-92.8) kg, 22 (19-25) y. The timeline for the study was 2 wk of cross-training, a 4-wk overload during which various markers were sampled, and a 1-wk taper. The 4-wk overload consisted of ~12 aerobic rowing sessions and two weight training sessions per week, and volume of rowing training increased by ~7% each week. Training volume was reduced by ~20% during the subsequent taper. Change in performance was determined by a stepwise lactate-threshold test at the beginning of the study and following the 1-wk taper, plus a 30-min rowing ergometer test at midday on Friday during the 4-wk overload. Reduced performance in both tests was used to classify athletes as non-functionally overreached,

and any rower whose performance in tests and competitions deteriorated over the subsequent seven months was deemed to have overtraining syndrome. All rowers maintained the prescribed training throughout the season.

The performance tests were conducted on a rowing ergometer (Model IIb, Concept2, Morrisville, VT). The first lactate-threshold test was preceded by at least 2-d of rest prior, and the second test was completed after a 7-d taper. The test involved 6-min steps with 15-W increments and rests of 1 min, during which blood lactate was sampled; the power corresponding to 4 $mmol.L^{-1}$ lactate was determined visually by interpolation from a scatter chart. The 30-min ergometer was a maximal test, restricted to 18 strokes per minute and conducted as a competition to optimise motivation. Both tests have very high correlations with the criterion measure of rowing performance the 2000-m rowing ergometer time trial[144] (TB Smith, unpublished observations).

The rowers were provided with a daily diary to record hours of sleep and number of times they awoke; they also recorded perceived fatigue when awakening and prior to bed using a 5-point scale (1–not at all, through 5–very much), rating of perceived exertion from the morning row on a 10-point scale, and their mood state using the Positive and Negative Affect Schedule (PANAS).[52] Immediately upon waking the rowers were requested to record their resting morning heart rate for 5 min while lying supine with the RS800 polar heart rate monitor (Polar Electro OY, Kempele, Finland). The Polar software (Version 4.03) provided five valid measures of heart-rate variability[150] (see Table 4.1).

The team undertook a medical examination approximately every 6 months, a monthly blood test and all had free access to a highly qualified sport physician and physiotherapist. We received all reports of injury, illness or disease diagnosed by these practitioners or any specialists the rowers were referred to. No rowers showed any consistent poor performances in the two months immediately preceding this study that would indicate any possibility of overreaching or overtraining.

Table 4.1 Linearized changes in measures of performance, heart-rate variability and psychological stress over the 4-wk overload. Data are means and SD of changes for the group, standardized individual change scores, and standardized 90% confidence limits for each individual's change. Boxes enclose individual rowers' clusters of substantial reductions in performance, reductions in heart-rate variability, and increases in psychological stress.

	Mean ± SD	OTS 1	Non-functional OR 2	3	4	5	6	7	8	Improved 9	10	CL
Performance												
30-min power	-0.4 ± 2.7 %	-1.4	-1.2	-0.5	-0.2	-0.1	0.1	0.2	0.2	0.6	2.2	1.4
4-mmol.L^{-1} power	2.5 ± 3.3 %	-1.1	-1.3	-0.4	-0.1	0.4	-0.7	1.0	1.7		0.5	1.5
Heart-Rate Variability												
LF/HF	-2 ± 40 %	-0.8	-0.1	-0.5	2.4		-0.4	0.6	-0.3	-0.6	-0.3	0.8
SDNN	18 ± 34 %	-1.0	-1.0	0.8	1.8		0.2	-0.7	-0.8	0.9	-0.2	0.7
RMSSD	10 ± 35 %	-0.8	-1.0	1.5	-0.3		0.3	-0.6	-0.9	1.5	0.4	0.9
MeanRR	4.0 ± 4.2 %	-1.1	-1.8	1.2	0.5		-0.4	0.9	-0.5	1.0	0.2	1.3
pNN50	10 ± 20 %	-0.2	-1.0	2.5	-0.3		-0.4	-0.3	-0.6	0.1	0.3	1.1
Psychological Stress[a]												
Morning fatigue	0.53 ± 0.66	2.0		-0.2	-0.3	0.0	-0.3	-1.1	-0.7	-0.5	1.4	0.7
Evening fatigue	0.51 ± 0.73	1.4		-0.2	-0.1	0.1	-1.0	-0.4	-1.4	-0.1	1.7	0.7
PANAS negative	2.9 ± 6.7	1.3		-0.2	2.0	-1.2	-0.1	-0.8	-0.6	-0.2	-0.2	0.4
PANAS positive[a]	0.7 ± 2.6	0.2		1.7	-0.8	-0.3	0.5	-0.5	-0.2	1.0	-1.6	1.2
Sleep duration[a]	-0.59 ± 0.44 h	0.2		-0.4	0.2	-0.6	1.0	-1.0	1.9	0.0	-1.4	1.6
Times awoken	0.08 ± 0.67	-0.1		-1.9	0.3	-0.7	0.1	1.5	-0.2	-0.2	1.2	1.0
Morning RPE	1.3 ± 1.7	0.1		0.8	0.5	1.8	-0.3	0.2	-1.1	-1.3	-0.8	0.9

OTS, overtraining syndrome; OR, overreaching; CL, 90% confidence limits for each rower's change, expressed in "±" form; LF/HF, ratio of low frequency to high frequency power; SDNN, standard deviation of all NN intervals; RMSSD, root mean square of differences; MeanRR, mean time between normal RR intervals; pNN50, proportion of differences between adjacent NN intervals of more than 50 ms.

Standardized change scores are each individual's change divided by the SD of change scores for the squad.

Rowers are ordered from worst to best change in 30-min performance.

Markers are ordered approximately best to worst for diagnosing the only case of overtraining syndrome (Rower 1).

[a]The sign has been changed on PANAS positive and sleep duration so that positive changes indicate increased stress for all psychometrics.

Measures of performance and heart-rate variability were log-transformed before analysis; changes were back transformed to percents after analysis. A measure of linearized mean change was derived by fitting a straight line through each individual's values, then calculating the difference between the predicted values for the first and last performance test or assay. We were interested a gradual changes that would be consistent with a gradual onset of overtraining fitting a simple linear model and deriving a linear change estimates such gradual changes. This approach is an appropriate parsimonious way to estimate a gradual change during the monitoring

period, which would be consistent with overtraining for noisy data of this nature. Individual changes were standardized for interpretation of magnitude.[71] The analyses were performed with the Statistical Analysis System (Version 9.1, SAS Institute, Cary, NC).

Results and Discussion

Taking the uncertainties in the performance measures into account, at least three rowers (1, 2 and 3) were candidates for non-functional overreaching (Table 1); of these, only Rower 1 suffered continuing performance decrement over the seven-month season with no indications of illness, injury or disease, which is consistent with overtraining syndrome.[6] Rowers 1 and 2 showed substantial reductions in heart-rate variability, whereas Rower 3 showed generally large *increases*. Small-moderate reductions in heart-rate variability occurred with only one other rower (Rower 8). The only rowers to experience large increases in several markers of psychological stress were the rower with overtraining syndrome and the most improved rower (Rower 10). An increase in psychological stress in combination with a decrease in heart-rate variability might nevertheless identify an overreached athlete who will eventually overtrain (Rower 1), but the failure of Rower 2 to complete her diary deprived us of additional evidence for this possibility.

Whether measures of heart-rate variability and psychological stress together would have sufficient sensitivity and specificity for early intervention to prevent overtraining syndrome is unclear. For example, Figure 4.1 shows that the rower destined to suffer overtraining syndrome (Rower 1) had a marked reduction in the LF/HF ratio in the first 2-3 wk of the overload, but so did the rower who showed the most improvement in the overload (Rower 10). Both rowers also had high levels of morning fatigue, and if anything Rower 10's morning fatigue showed a greater increase than that of Rower 1. A reduction of training aimed at limiting the decline in LF/HF and reducing morning fatigue might have prevented overtraining in Rower 1 but might also have reduced Rower 10's dramatic improvement. We are unaware of any studies providing evidence that decreases in heart-rate variability presage overtraining syndrome, but measures of psychological stress were apparently useful predictors of at least non-functional overreaching.[151]

Figure 4.1 LF/HF and morning fatigue over the 4-wk overload for Rower 1 (who became overtrained) and 10 (who was amongst the most improved in performance). Values and changes over the first 2-3 wk are similar.
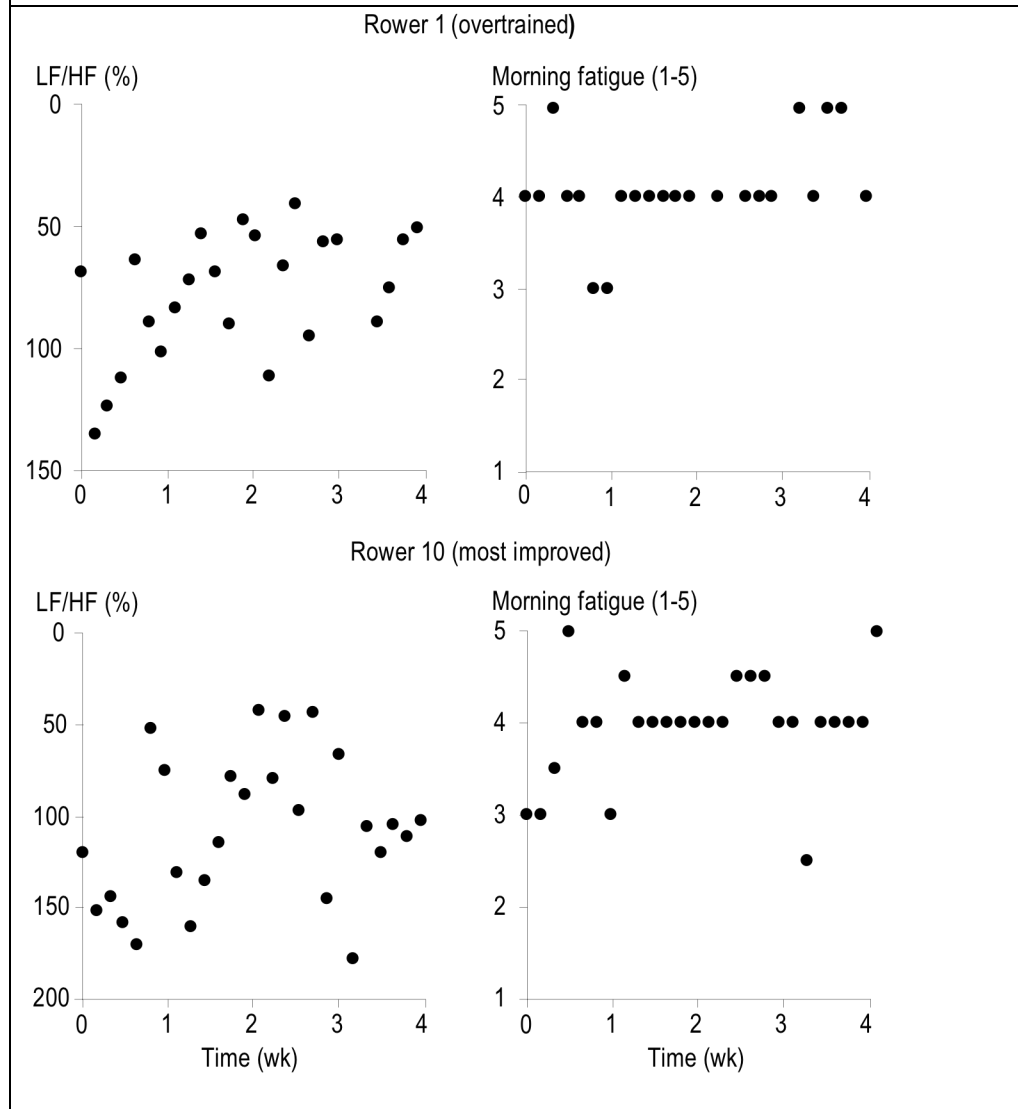
In conclusion changes in performance, heart-rate variability, and psychometrics individually appear to lack the sensitivity and specificity to prevent overtraining syndrome, at least in these elite rowers. Together these measures may be useful indicators of excessive training.

Acknowledgements

CHAPTER 5

Managing Athletes on the Edge: Overtraining, and the Complexities of Coaches' Decision Making.

Overview

In many elite sport performance contexts it is acknowledged that overtraining is a critical issue for coaches to understand, be able to recognise and respond to. While much scientific research has been conducted on overtraining, no diagnostic test exists that reliably predicts impending overtraining. The research reported in this paper anticipated that successful endurance coaches have developed effective strategies for managing overtraining amongst elite athletes. In a first of its type, the study explored in depth the practices and beliefs of three highly successful professional elite rowing coaches as they managed the risk of athlete overtraining during crucial periods of intensive training. These coaches took part in a series of open-ended semi-formal interviews that explored their practices, philosophies and experiences. The paper draws on sociological perspectives pertinent to coaching pedagogy and more specifically, the tenets of naturalistic decision making (NDM) in exploring how the coaches define and make decisions about overtraining. This framework reflects recognition that coaching involves decision-making in complex and demanding situations. The coaches were found to monitor the fatigue of their rowers through observation, communication and measures of training pace. In general, the intuitions and cues for overtraining employed by them had little in common with those promoted in the sport science and medical literature. We contend that coaches' decision-making is based largely on subjective processes and influenced by various stressors unique to their positions, which may explain the differences between their definitions and practices and those of the sport science and medical fraternity. We suggest that successful coaches have unique insights into this topical area that is worthy of further exploration.

Key words: overtraining; monitoring; decision making; qualitative; coach; rowing

Introduction

In many elite sport performances contexts it is acknowledged that overtraining is a critical issue for coaches to understand, be able to recognise, and respond to. From a simple training and conditioning perspective, overtraining can be described as performance decrement associated with excessive workload and inadequate recovery. It is recognised as a particularly important issue for elite endurance athletes, as the hard training required to optimise performance can create an imbalance between training load and recovery, increasing the risk of overtraining. An important challenge therefore facing the coach of endurance athlete is to determine how to maximise their athletes' training load so as to increase the chance of optimising conditioning, while also alleviating the concomitant risk of overtraining. Incidents of under-performance on a world stage due to overtraining arguably point to coaches struggling in this challenge.[152] This is despite increasing sophistication in the sports science support available in many elite performance and coaching contexts. The prevalence of athletes who have suffered overtraining is reported to range from 5% to 64%[135, 153] with rates for Olympic athletes reported at 10-28%.[4]

The research reported in this paper is notable for focusing on a context in which repeated performance successes on the world stage point to coaches being able to predict and effectively manage impending overtraining. In the space of a decade Rowing New Zealand has made a dramatic increase in its world rankings, resulting in their elevation to a position (based on Beijing Olympic results) of being New Zealand's premier Olympic sport. These successes include an approximate five fold increase in elite world championship and Olympic medals from the previous decade, being ranked in the top three rowing nations 2009-2011, and both coaches and rowers receiving numerous national and international awards (e.g. since 2000 the Head Coach has been awarded the FISA coach of the year twice and both rowers and coaches have regularly won New Zealand's top sports awards).

Drawing on interview and observation data gathered from three Rowing New Zealand coaches between 2007 and 2010 we describe how the coaches define and make decisions about overtraining. As we discuss further below, the research aligns with a

growing body of coaching literature that acknowledges decision-making as complex, inherently subjective, and central to the interactive social process of coaching.[154-156] While locating the work within the broad frame provided by sociological studies of coaching pedagogy, we also draw on the tenets of naturalistic decision making (NDM)[157, 158] in an endeavor to generate clearer insights into the decision-making processes that the coaches engage in amidst complex and clearly demanding contexts in which expectations for continued international success are explicit and from a coaching perspective 'ever present'. As an illustration of this expectation, over the period of the study two coaches with histories of international success were forced to seek employment elsewhere because of short periods of poor performances.

Necessarily, we open by providing some essential background commentary on overtraining. Subsequent sections of the paper then focus attention on how the coaches in this study define overtraining, justify intensive training regimes that could induce overtraining and recognise the signs of impending overtraining.

Optimal Conditioning, Overreaching and Overtraining

It has been estimated that training loads have increased dramatically over the last decade as elite athletes work harder to ensure optimal performance, which in turn has increased the risk of overtraining.[152] This risk is exacerbated by the often-used practice of implementing periods of excessive workloads in an attempt to maximise training overload.[6, 11] While optimal conditioning requires a healthy athlete and an effective training programme that balances intensive training with adequate recovery, arguably the athlete who can continue to respond positively to the greatest training load will develop the best physical conditioning.[159] This creates more pressure for the coach who must find a balance between the benefits and risks of intensive training.

Unfortunately research on overtraining has lead to some confusion around the definition and diagnostic criteria.[7] In response to this the European College of Sport Sciences formed a task force that has developed a position statement on the prevention, diagnosis and treatment of overtraining syndrome.[6] An important part of this statement was defining overtraining as a verb, "a process of intensified training with possible outcomes of short-term overreaching (functional overreaching), extreme overreaching (non-functional overreaching) or overtraining syndrome". These terms reflect the notion of a fatigue continuum that results in performance changes which range from increase, temporary decrease, stagnation and decrease (see table 1). Non-functional overreaching and overtraining syndrome are considered the stage on the

fatigue continuum where reduced performance is combined with mal-adaptation of various physiological mechanisms. The differences between non-functional overreaching and overtraining syndrome are often subtle and based on a retrospective diagnosis on the period of performance reduction, plus prolonged mal-adaptation of several biological, neurochemical and hormonal regulation mechanisms

Table 5.1 Possible presentation of the different stages of overreaching and overtraining

PROCESS	TRAINING (overload)	INTENSIFIED TRAINING →		
OUTCOME	Acute Fatigue	Functional overreaching	Non-functional overreaching	Overtraining syndrome
RECOVERY	Day (s)	Days - weeks	Weeks – Months	Months - …
PERFORMANCE	Increase	Acute decrease (e.g. training camp)	Chronic stagnation or decrease	Decrease

Adapted from Meeusen et al.[6]

The prevalence and detrimental effects of overtraining in endurance athletes during hard training has stimulated a search for diagnostic markers that could predict its onset[6, 11, 38] An effective marker must be sensitive enough to predict impending overtraining enabling remedial reductions in training, and also be specific enough so that training is not reduced for an athlete who would benefit from the increased training overload.

Fry et al.[9] identified 84 major symptoms of overtraining as indicated by the prevalence in the literature, with various subsequent publications promoting a wide range of signs and symptoms of overtraining.[3, 6, 11, 38, 55, 132, 160] Kentta and Hassmen[161, 162] proposed a novel athlete self-monitoring system that attempted to address the physiological, psychological and social dimensions of training and recovery. This monitoring system consisted of the ratings of the perceived exertion[53] and the total quality recovery[161, 162] scales.

Despite the large volume of research in this area and the regular promotion of these various monitoring tools, there is negligible evidence for a diagnostic marker that has the sensitivity and specificity to accurately predict impending overtraining.[5, 159] Furthermore, proper diagnosis of overtraining itself can only be made retrospectively, after all other possible influences on change in performance have been excluded.[6]

Despite this, various markers are routinely promoted as effective monitoring tools.[6, 20, 36, 38, 55, 130-132, 135, 163]

During periods of training when overload is maximised, the coach must constantly monitor the athlete so as to judge if recovery periods and/or training intensity need to be modified in response to signs of impending overtraining. The markers proposed to predict impending overtraining are predominantly objective and scientific, yet by its very nature coaching is arguably more of an art than a science,[164] more subjective than objective.[165] Therefore, not only is the effectiveness of the markers of overtraining questionable,[5] but also the ability of the coach to utilise them effectively seems likely to be limited.

Making Meaning and Making Decisions as a Coach: the Scientific and the Subjective

The challenge for the coach, positioned at the nexus of a vast array of information when managing their athletes' training, is to decide what to act on and when. Do they rely on the plethora of imprecise scientific and psychological tools? Moreover, how do coaches select effective monitoring strategies that complement their skill sets, knowledge, personality, environment and the athletes they coach? The multitude of such considerations infers there is no simple strategy for monitoring overtraining and that methods employed by successful coaches working in applied settings are quite possibly *unique, changeable and unorthodox*. Collectively these considerations would indicate that not everything is clear-cut and there will always be some element of subjectivity in answering such questions. As indicated above, such a view aligns with conceptions of coaching as more of an art than science,[164] 'as pedagogy' and as first and foremost, a social process.[154] Lyle[166] advocates that many coaches must make decisions often based on experience and the fickle behaviours of athletes. Such factors are not situated within the scientific world, but rather the social, and they reflect the need to acknowledge coaches as working with and amidst a dynamic between the coaching context and the wider world within which it is set and to which it always relates. Coaches learning, decisions and actions are then, always socially situated and as such, will necessarily reflect a dynamic between 'the scientific' (knowledge, understandings, principles and data) and 'the social', with the latter acknowledging knowledge as grounded in personal experiences (of coaches and athletes), and as developed in and

through pedagogical relations that are inherently social relations (see for example, Jones[154].

This study responds to a relative lack of research engaging with this dynamic in relation to coaching pedagogy, and more specifically, coaches' decision-making relating to specific issues or aspects of training. As indicated above, overtraining offers a potentially very productive focus via which to gain further insight into the complexities of coaches' decision making and particularly, explore how coaches variously turn to and draw on scientific and subjective knowledge in their decision making.

While publications on strategies to detect and manage overtraining from the sport science and medicine fraternity abound[3, 6, 11, 38] very little attention has been accorded to the practices of successful coaches. We were intrigued by the potential discrepancy between the strategies promoted by sport scientists such as[6] for predicting impending overtraining and those employed by successful coaches , and saw the prospective worth in looking more closely at coaches' decision-making associated with overtraining. We acknowledged in particular, that because of the limitations and pressures mentioned above, many coaches might have developed their own unique strategies for identifying and managing overtraining in an elite performance environment.

Naturalistic Decision Making (NDM) as a Framework for Exploring Situated Decisions

In essence we wanted to explore the transition from 'concept to context' and learn how these coaches actually made decisions about overtraining in a real-world or naturalistic setting, that is characterised by the need to make difficult and potentially 'high stakes' decisions often under demanding conditions. Naturalistic decision making (NDM)[157, 158] emerged in the 1980s and was adopted to explore the role of experience in enabling people to rapidly categorize situations to execute effective decisions in real world settings. It is a framework used by researchers in medicine[167, 168] and business settings[169] where the stakes can be high, decisions often have to be made in conditions of limited time, and where decision-makers are forced to make a judgment call that may have very significant consequences for themselves and others. Field research into the

notion of decision-making by Klein and colleagues[170] was conducted to explore the strategies people used. This was an alternative to systematic evaluation techniques or optimal performance principles that they discovered were rarely employed. Their research highlighted that many formal standards of decision-making employed in training programmes did not improve decision quality and were basically discarded as cumbersome and irrelevant in field settings.

As the NDM movement evolved attention turned to the role of experience in decision-making, as opposed to individuals making choices from available options. Focus was accorded to prior perception and recognition of situations. Klein's work has emphasised the idea that "when people need to make a decision they quickly match the situation to patterns they have learned"[157, p457] and if a clear match is made, they can follow a typical course of action. In sum NDM has presented an option of intuitive decision-making.[170-172]

Arguably the most celebrated work that has examined intuitive judgement and decision-making is that of Nobel Prize winner Daniel Kahneman.[173] His collaborative work with Amos Tversky[174] is couched through three pertinent topics: heuristics and judgements, risky choice, and framing effects. All three topics focused heavily on intuitions based on thoughts and preferences that come to mind quickly and with little reflection.

Two generic modes; an intuitive mode (thoughts and preferences that come to mind quickly and without significant reflection) and a controlled mode that is deliberate and slower. These proposed topics later became accepted as a two-system view that distinguishes intuition from reasoning. System 1 is "fast, automatic, effortless, associative, implicit (not available to introspection), and often emotionally charged; they are also governed by habit and are therefore difficult to control or modify".[173, p698] System 2 operations are " "slower, serial, effortful, more likely to be consciously monitored and deliberately controlled: they are relatively flexible and potentially rule-governed.[173, p698] Kahnerman explains that system 1 generates impressions of the attributes of objects of perception and thought. The label *intuitive* is applied to pronouncements that directly reflect impressions - they are not modified by system 2. By contrast *judgements* are always intentional and explicit whether they stem from impressions or reasoning.

Klein[170] has argued that skilled decision makers often are more effective when they trust their intuitions than when they engage in detailed and systematic analysis. In

contemporary contexts of elite performance coaching, it seems highly pertinent to explore this proposition and in so doing, extend insights into the ways in which the dynamic between scientific and subjective knowledge plays out in coaching pedagogy. Following Cassidy et al.[175] we recognise the prospective significance of the reflective process as a means of coaches extending their awareness of experience and meaning. While subjective knowledge connotes knowledge gained through personal experience, experience alone will not necessarily create knowledge. Experience must be reflected upon for it to become meaningful.[176] From this perspective, through a reflective process, coaches generate new subjective knowledge and thus, from a personal perspective, simultaneously develop enhanced understandings of aspects of training and performance, such as overtraining. This qualitative study provides insight into the subjective knowledge and subsequent actions of three elite rowing coaches who have shared their situated experiences with overtraining.

The Research Context

Three experienced successful professional coaches from the Rowing New Zealand elite team agreed to participate in this research. They all had at least eight years experience coaching national representatives and between them had coached five Olympic medalists and 38 crews who had won medals at either Senior A or Senior B World Rowing Championships. Semi-structured, in-depth interviews were used to explore how the coaches' defined overtraining, why they considered overtraining important and the cues they used to determine whether a rower was at risk of impending overtraining. This research method was deemed appropriate as it promoted the flexibility to probe and explore the multiple levels of discourse used by these respondents, allowing a focus not only on the words spoken but also on the meanings intended.[177] The project received ethical approval from a New Zealand University. As the Rowing New Zealand community is very small, as part of the ethics requirements we agreed to take all possible measures to protect the identities of the three coaches. We have therefore adopted a style of writing that balances the requirements of presenting the coaches stories and views while trying to reduce the speculation within or beyond the Rowing New Zealand community about 'who said what'. It is for this reason we have not allocated pseudonyms or assigned any contextual detail.

As well as extensive coaching experience the coaches had also undertaken a number of coaching and sport science courses, with two having completed tertiary qualifications that included papers in sport science, sport psychology and coaching. The coaches also had extensive open access to a network of qualified and respected sport science and medicine specialists.

The first author, who performed all the interviews as part of a doctoral thesis, was the lead sport physiologist for this rowing programme. His research focused on examining a wide range of strategies for optimising performance and predicting overtraining for elite rowers as they prepared for international competition. As part of his doctoral research he was imbedded fulltime with the New Zealand elite rowing team from 2006 to 2010. In this role the first author lived and traveled with the team at every training venue and international competition from the preparation for the 2006 world rowing championships in Eton up until and including the 2008 Olympic Games in Beijing. The interviewer's standing and trust through the years of shared experiences enabled unlimited access to the coaches and facilitated in-depth exploration of the motives and reasoning underpinning the coaches various practices.

All of the interviews were digitally audio-recorded and transcribed verbatim. Each coach was interviewed at least three times, with each interview lasting between ~20 minutes and ~2 hours. The first interview was conducted at the team's training base at Lake Karapiro at a date and time of the participant coaches' determination. Follow-up interviews were conducted at various times and places where the team was located over the subsequent three years.

Following transcription all interview data was inductively analysed using ethnographic content analysis.[178, 179] This procedure facilitates categories to emerge from the raw data whilst conserving a grounded status rather than binding the data to theoretically informed codes. The transcripts were analysed using open coding[180] where the descriptive data were ordered for salient theoretical themes or issues that lay behind the text. Each transcript was read reflectively several times to promote and confirm the identification of key themes.[181] The coaches were provided with copies of their completed transcripts, which they signed to verify that they were accurate renderings of their various interviews. They were also consulted regularly during the preparation of this manuscript and any issues they had were factored into the final article. Themes that highlighted salient considerations and informed how the decision-making process evolved were thereby explored.

Results and Discussion

Qualitative research involves a close relationship between description and interpretation that we have sought to facilitate by integrating the results and discussion and focusing on three themes that were developed from the analysis process described above. The three themes were: personal definition and description of overtraining; risk and justification of intensive training regimes; and recognising the signs of overtraining.

Theme 1: Personal Definition and Description of Overtraining

While reducing the incidences of overtraining is probably the most important goal for these elite coaches, before we addressed their various strategies we felt it important to first examine how these three coaches defined overtraining. The aim in this section is to therefore examine the similarities and differences between the three coaches definitions and if possible compare their definitions to that presented by the European College of Sport Sciences Task Force on Overtraining.[6]

The coaches had spent many years working together which had probably resulted in a shared philosophy on many aspects of coaching including training and overtraining. It was therefore not surprising that their definitions for overtraining had a common theme, namely: training that caused excessive fatigue would lead to reduced performance. From the coaches' perspective, reduced performance was determined by an inability to maintain "normal" training workloads for an "extended" period of time despite the best intention of the rowers.

> They want to do it but they just can't go fast enough, They may talk about wanting to do well, they may keep training, but they are just well off the pace, they just fall behind, they can't do it.

The coaches believed that as the rowers neared the end of an intensive period of training they should experience reductions in training performance, a condition they all considered normal. However these reductions should be acute, that is training and

competition performance should return to normal after the programmed period of recovery, which typically lasts three to seven days.

> We push them hard so that some weeks they are going to get fatigued and tired but tiredness and fatigue from just training is different than overtraining... There is a difference between fatigue and overtraining ... the difference for me is when I build them through a really hard week I expect them to be tired ... I expect them to be absolutely knackered for the last week of training, but that is what the plan is, if you have worked hard enough that is what happens, I think you've got to go a little bit past what they think they can handle to get more out of them. Day to day you are assessing how they cope with the workload in training. It's the standard thing, their [boat speed].

Rowers whose performance failed to return to normal after these programmed rest periods were considered to be at risk of overtraining.

> They are coming back from their hardest week and after a break and their lightest week they are not recovering, they haven't bounced back.

This coach verified his interpretation by stating:

> They need to be out there doing the same amount of work holding the same [speeds] every day and if they can't for a long period of time ... then something's not right and we might back them off.

The risk of overtraining was seen as increasing as the period of time over which their performance had decreased extended, with the decrease generally determined from their training speeds. Furthermore these performance reductions were not the normal fade experienced during period of acute fatigue. They are very obvious and severe reductions.

Overtraining in the crudest sense of the word is when they have consistent [poor] performance … noticeable decrements in performance ... we are not talking about the gradual fade that you might get across a training week or a training block, it's quite gross and obvious.

The coaches' views of overtraining were clarified after a series of follow up interviews culminating in the following consensus "*overtraining is a level of chronic fatigue that reduces training to the point where performance at upcoming international competitions is impaired*". International competition performance could be impaired due to undue fatigue effecting race performance and / or reduced quality of training resulting in suboptimal conditioning. This definition is in essence subjective and individualised from a coaching perspective, as the level of performance decrease that signifies overtraining differs for each rower. The salience of subjectivity is perhaps best reflected in the view of one coach who argued:

I have had two athletes that were [theoretically] overtrained, They were continually tracking backwards for [2-3 months] but they were not properly overtrained because I don't think it ruined their [international] season.

In this instance the two rowers performance measured regularly during testing, training and competition slowly deteriorated over the 2-3 month period of training. The coach believed that while these two rowers were very fatigued during this training period this fatigue eventually dissipated and they subsequently performed well, hence they were not overtrained.

If we return to literature, it is clear that a single definition for overtraining from researchers and practitioners has proved difficult and that a wide range of definitions prevail.[5, 7] It is therefore not surprising that these coaches had developed their own definition, and that their definitions appeared to have few similarities to the European College of Sport Science'[6] position statement on overtraining (see table 1). Meeusen et al.[6] and the coaches viewed reduced performance as the criterion determinant of overtraining, although the coaches discussed measuring performance during both training and competition. By comparison Meeusen et al.[6] did not address how performance stagnation or reductions were determined. Meeusen et al. [6] also proposed an overtraining continuum (see Table 1) with the diagnosis of overtraining dependent

on the timeframe of performance reduction and the degree of physiological maladaptation and/or psychological disturbance. The coaches disagreed with both a set timeframe of performance reduction plus physiological and psychological factors being used to define overtraining.

The aforementioned disparity between the coaches definitions of overtraining and those presented by Meeusen et al.[6] may relate to the different ways in which coaches and sport science or medical practitioners respectively conceptualise overtraining. While the coaches appear to view overtraining through a subjective lens, the sport science and medical practitioners are more likely to employ a scientific perspective that has thus far failed to determine a common definition and effective diagnostic tools.[5, 182] This incongruence has led these three coaches to ignore the lexicon of scientific terminology. The coaches were unaware of the terms non-functional overreaching and overtraining syndrome. The disparity between the subjective and the scientific definitions was perhaps best articulated by one coach who insisted:

> I am not a subscriber to overtraining as it's laid out in the textbooks.

While many of the published definitions of overtraining focus on reduced performance the coaches believed that sport science and medical practitioners defined and determined overtraining from tests of physiological maladaption and/or psychological disorders. This may reflect that the various sport scientists and sport physicians they have come in contact with have been seen to use solely physiological and/or psychological assessments to diagnose overtraining. One coach described an incident when one of his better athletes had been diagnosed as overtrained by a sport physician.

> [Athlete] had gone to the [Sport Physician] because he said he wasn't feeling well. [The Sport Physician] told him he was overtrained and that he needed time off. [Athlete] was rowing well and doing good numbers (training speeds)… We carried on as normal and he went on to [perform with distinction].

In this instance there were no abnormalities in the blood tests or physical assessments, the diagnosis was made solely on the basis of psychological

assessments of mood, quality of sleep and perceptions of fatigue. The coach expressed his displeasure at this process and believed that he had information relevant to the diagnosis. Another coach described an incident where a medical practitioner gave an athlete that he considered as overtrained a clean bill of health;

> I remember one guy who had a blood test... there was nothing wrong with his blood test, but he was just down and he never came back up, he was overtrained and never recovered.

The inclusion of physiological maladaptation and psychological disturbance in the definition of overtraining is common and these measures are commonly referred to as signs and symptoms of overtraining.[3, 6, 11, 38, 132] We believe that logistical and technical difficulties in accurately assessing change in performance[10] has forced many sport scientists and physicians to utilise physiological and psychological signs and symptoms of overtraining as proxies for both defining and detecting overtraining. There is however little evidence to support the notion that abnormal changes or levels of one or more physiological and/or psychological marker, reliably predicts chronic reductions in sporting performance.[5]

There are many possible reasons for the differences in the definition employed by the coaches compared to that of Meeusen et al.[6] We believe one of the major reasons is that the coaches definition is based on performance, as this is the area they understand, have some control over and is very important since the fastest endurance athlete is most likely to be successful. Hence their definition is pragmatic and focuses on subjective interpretation of measures of performance. On the other hand the definition proposed by Meeusen et al.[6] focuses not only on the objective measures of performance but also those causal mechanisms that can be objectively quantified, which is a key tenet of the scientific method. The problem with this approach is the relationship between the criterion measure (performance) and the various proposed causal mechanisms determined by various physiological and/or psychological measures, is usually poor[5], which should invalidate their inclusion in the definition and description of overtraining.

Theme 2: Risk and Justification of Intensive Training Regimes

Overtraining is considered to be more prevalent when endurance athletes are involved in periods of intensive training with reduced opportunity for effective recovery[3]. All the coaches held strong beliefs on the importance of training beyond the athlete's tolerance for prolonged periods, as a means of extending their physiological and psychological capacities. They believed that what differentiated success from failure in international competition came down to small fractions of a percent advantage in speed over their competitors. Therefore, optimising their athletes' performance capabilities was vital for success and this was only achieved through periods of intensive training. Hence all three coaches considered overtraining to be an ever-present and legitimate risk, due to the intensity of their respective programmes. One coach exclaimed:

> You've got to train hard, harder than anyone else and there are so many examples out there that you have to train harder than the opposition [to beat them], if you don't train as hard as, or harder, then your not going to beat them. There are always going to be crews that are going to beat you that are exceptional, but on a level playing field if you don't train as hard or harder than them, you are not going to win.

Another coach who described his international competition further qualified this perspective:

> They're [successful because they] have enough physiological specimens who can race bloody well, but aren't necessarily that fit and trained that hard, if we can train harder, we can beat a lot of those countries.

Advocacy for the emphasis on the importance of hard training was reflected in another coach stating:

> I expect them to be absolutely knackered for the last week [of the block of training], … that is what the plan is, if you have worked hard enough that is what happens.

> I think you've got to be careful not to cut back [the training load] too early… you've got to go a little bit past what they think they can handle to get more out of them.

These quotes reflect that the coaches' focus was on trying to train athletes beyond what they believe they are capable of and in the process, reset the athlete's perceptions of what their bodies can handle. All the coaches spoke about gradually increasing the workload over a period of 3-5 weeks so that at the end of this period the athlete could not adequately recover, resulting in severe fatigue and performance reductions. This was followed by a period of reduced workload and rest that should stimulate a super-compensatory response thereby creating enhanced physical adaptations. This process of intensified training overload creating acute fatigue is an often-used method for stimulating adaptations in high performance endurance athletes.[1, 2] The coaches all considered this process to be important for the development of their athletes but were however, also cognisant that it dramatically increased the chances of the athlete not recovering in the available timeframe, thus increasing the risk of overtraining:

> When you are working with elites and you're trying to push, really push them along and get those fractions of a percent, you are on the knife edge it would be very easy for it to go either way…. it's a fine line probably between training in what they would call an overtrained state and actually overtraining and tipping them over.

> If they're not slightly over the edge then perhaps they're not pushing the boundary of their limits enough". "You almost have to have someone [almost overtrained] or else you're not quite pushing them hard enough. This is what they've got to do as an elite rower, … if they don't do it they're not going to race to the best of their ability anyway.

> You've got to push it to the limit, you don't know what the limit is so you've got to push, its better to push up there and take a day off and recover than never to get there at all.

Although there were only a few incidences of overtraining suffered by rowers under their care, these three coaches described numerous incidences when their athletes had been on a "knife edge" and were acknowledged as very close to becoming overtrained. All of these incidences occurred during blocks of intensive training, but the coaches were unable to shed light on the possible reasons why at these specific times, the rowers fatigue had developed to the point that they were on the cusp of overtraining. The first incident occurred early in the training year, during a period of intensive aerobic conditioning. The coach was alerted initially by the crews' training pace becoming erratic and from changes in their body language, and later from communication with the rowers.

> They'd got to the level where they were just overworking … we were overworking them … we cut the miles down because [the athlete] was starting to row badly.

The next two incidences occurred near the end of an intensive period of anaerobic conditioning just prior to competition at the world championships.

> We were lucky that year we probably pushed them over the edge, fortunately the (weather) kept us off the water for a while and it allowed them to come back up again … it was very close.

> Oh yeah, we've pushed them pretty hard ... I remember [Athlete] saying 'I don't think we're going to survive this' ... it was pretty close to tipping them over.

Once again, coaches became aware of the danger by monitoring training performance, interpreting body language and from communication with the rowers. The ways in which each of these factors informed coaches' thinking and decision-making is discussed further in the next section).

The following case was not a specific incident, but relates to a rower who was recognised as having a habit of training extremely hard all of the time, and was

perceived to be constantly in danger of overtraining and/or injury. The coach explained how constant monitoring of the body language and performance of this rower often revealed episodes of extreme fatigue, which needed to be managed to reduce the risk of overtraining.

> [Athlete] works bloody hard just to hang on and you could easily tip [them] over, [they've] just got to keep on going, yet [Athlete] managed to hold on for the last 3 to 4 years and [has] improved because of it.

The coaches were aware of times when they had pushed the boundaries and their rowers required extended periods of recovery to halt the deterioration of their performance. A good example of this is described thus:

> We had to back him right off and just really keep him out of the boat for three to four weeks, and if we hadn't … he wouldn't have bounced back at all … but it was always touch and go.

Collectively the coaches had achieved numerous international successes that they believed confirmed the effectiveness of their intensive training regimes, despite the increased risk of overtraining. Their view was that if an athlete could not follow their respective programmes in the early stages of the training year they would not be competitive internationally. Therefore, any athlete who struggled to handle the training load, especially during the early "easier" portions of the programme was probably unlikely to succeed and retain their position in the team:

> They're coming into the elite programme so they're expected to be able to do it. It's the old story, if they can't do it perhaps they shouldn't be there anyway.

The coaches made the rowers aware that they required a large volume of intensive endurance training in an extremely competitive environment. The challenge for each coach was to ensure that the rowers completed the required periods of intensive training while at the same time managing the concomitant risk of overtraining. This challenge was exacerbated by the fact that the coaches would often have to monitor the response

of the individual within a crew of two, four or eight rowers. It is therefore possible that the same training programme can result in a mixture of under, optimal and over training for the individual rowers within the same crew.

> [Athlete] is a genuine case of overtraining ... which goes to the individuality of training doesn't it, where as we train everyone the same, someone like [Athlete] obviously can't, physically probably handle that volume of work, somehow you've got to do something different with [Athlete], which is hard in a crew.

Further factors to consider here are that the coaches are working with a very limited pool of athletes and that the continued funding of the programme, and their ongoing employment as coaches was reliant upon maintaining success at the international level. The coaches were also aware, however, that arguably 'risky practices' in terms of the level of risk of overtraining, were an integral component of Rowing New Zealand's international successes.

> You can't hold the whole crew back for one person, you can't hold the whole programme back for one person, you've got to push …. so if we wrote off the whole programme [almost all became overtrained] to get that one person who could do the work, that would give us a gold medal, whereas if we focused on the bottom level we'd never get a gold medal.

While this "survival of the fittest" approach to training can sound harsh, it has reaped many rewards. Nevertheless, despite this success and the associated promotion of intensive training, the coaches believed that *constant vigilance for overtraining cues was an important aspect of their coaching roles*.

Although we have considered if these practices are the best for optimising athlete performance or whether there are better, possibly more "humane" ways of developing a successful international rower, again, we have been promoted to acknowledge some unique characteristics of this coaching context. The athletes are volunteers with generally limited financial support who stay in the programme for a number of reasons that include the achievement of their goals. The numbers of elite New Zealand rowers

has risen approximately four-fold from 2000-2010 compared to the previous decade, with the numbers of rowers winning world championship and Olympic medals increasing approximately ten-fold for the same period. Arguably, if the rowers did not support the coaches' intensive approach to training, their numbers and successes would have declined during the tenure of these coaches.

As indicated above, the most important criterion for the on-going employment of these coaches is continued international success. Their success is evident in their previously mentioned medal counts and the fact that the RNZ elite programme has enjoyed a world ranking of 2nd to 3rd over the last four years while their ranking was only ~25th in the late 1990's. Therefore, while there may be more effective training strategies than those employed by these coaches, their methods have arguably been very successful.

The coaches all cared for their rowers' wellbeing, but at the same time they were aware that implementing intensive training regimes that increased the risk of overtraining also increased the chance of the rower realising their goal of success. The coaches' sentiments are borne out in the following coaches comment:

> I definitely have a genuine concern for my rowers, I think it's probably a flaw and I had to be a little more ruthless but I was just of the opinion that you know we're trying to get a [boat] to go fast and we need the psychologically, physiologically, whatever, strongest people sitting in that boat

The coaches were therefore, committed to maintaining a programme that they acknowledged, carried inherent risks and challenges in terms of managing athletes during periods where they are on the knife-edge of optimal training versus overtraining. Being effective in this context centres on coaches' ability to recognise and respond to signs of overtraining.

Theme 3: Recognising the Signs of Overtraining

The challenge for the coaches was to determine what useful tools and strategies could help them to manage the heightened risk of overtraining. As indicated previously, this challenge was accentuated by the dominance of objective scientific methodologies that have thus far failed to produce reliable and valid predictors of overtraining. Hence, as

we now illustrate, the coaches' focus is 'performance reduction' and they have developed *subjective strategies to predict impending overtraining.*

The most important marker or cue of increased risk of overtraining for all three coaches was *performance reduction*, recognition of which was reliant on *performance monitoring*. While these terms may be seen to imply or align with scientific discourses, this research revealed the extent to which in this elite and applied context, subjective discourses necessarily came to the fore in coaches' descriptions of the ways in which they gain an understanding of the rowers' performance, responses to training, and risks of overtraining. -

Competition is often regarded as the best monitoring tool for determining changes in performance. It is, however, rare for elite rowers to have regular competition, with consistent environmental conditions and the same competitors competing maximally. Training workload is another common performance monitoring tool, however logistical and technical difficulties have resulted in a lack of valid and reliable measures of training performance for most endurance sports.[183] In monitoring performance the coaches measured boat speed for most training sessions, implemented weekly competition sessions over various distances and raced in a regatta at least once a month. These performance-monitoring strategies only measured the speed of the boat and not the individual members of the crew, unless the rower was a single sculler. For this reason the coaches regularly added competitive rowing ergometry sessions and/or training in the smaller boats (singles and coxless pairs) to enable more effective monitoring of the individual rower. The importance and complexity of performance monitoring is reflected in the following coaches' comments:

> It comes down to performance. If they had come and said to me oh you know look I'm [really tired] I'd go back and say look at this, you're ok, the numbers are still good, ... the boat speed's still there, we're at a point in the training programme when you should be feeling like this.

> You're always watching your rowers, just general mood and things like that, you're always keeping an eye on them but it's all quantified off the [boat speed and rowing ergometer] numbers that you collect.

Well actually the [unit for measuring boat speed] is a good one because I mean that's an invaluable tool really. I think that's just the best thing that we've had access to. If the numbers on the U2 (aerobic base) you know [boat speed] are not where they're supposed to be, you know, and they've been holding the right [boat speeds] for a long time and then suddenly it just goes down for a while then I have to look at them and see what's going on physically.

Performance monitoring in rowing is not a simplistic process because changing environmental conditions between and within a session can dramatically modify the boat speed. For this reason the coaches spoke about not just concentrating on absolute speed but also examining their crews' speed relative to the remainder of their squad and the team.

If you have two crews going against each other then it becomes pretty obvious one crew suddenly really drops its bundle. That's pretty easy to do.

If you've got two crews it's easier you know if the boats are neck and neck and then one just drops right off the back.

The challenge was determining the degree and timeframe of absolute and/or relative boat speed reductions and/or reductions in rowing ergometry that predicted upcoming overtraining. All three coaches expected some performance drop off during the week and possibly near the end of the 3-5 week training block. Their problem was determining what level of training performance drop off is considered abnormal and therefore a cue for impending overtraining. While they all made a number of comments on this issue, none were able to offer any heuristic for recognising impending overtraining. Such a drop off was acknowledged in one coach's description of an early cue for possible overtraining;

They have consistent [poor] performance over 2 or 3 sessions, a noticeable decrement in performance over 2 or 3 sessions, we are not talking that gradual fade that you might get across a training week or even across a training block, its quite gross and obvious.

While dramatic reductions in training speed is an important cue for impending overtraining the more important consideration was the potential underpinning reasons for such a reduction. For example, a possible reason could be nutrition:

> But what would be the level of falling off? You know a few metres, fifty metres, 100 metres if they dropped off, 500 you know it could be just purely a lack of energy and food you know on the day.

Another consideration could be the difference in how the individual rowers approached training and competition. While the coaches considered most of their rowers to be honest hard workers who followed the training programme to the best of their ability and competed maximally in all competitions, they described others as lazy trainers or inconsistent competitors.

> You need to be careful because some people don't put the effort in earlier on, they'll save themselves for the last one then you've got to juggle you know you might have to do another couple or something or you don't tell them what you're doing and try and get the work out of them ... Knowing what they're like, whether they're hard workers or a bit scared at the beginning or saving themselves for the last one you can understand those people.

> [Athlete] is not competitive enough day in day out [Athlete] would just do the same, same, same not worry about [trying to do faster speeds] every day ... and then believe that [they] could [race successively] on the day.

Thus, it was important for the coach to develop an understanding of how individual athletes approached training to help them understand the possible reasons for poor performances. Whether *performance reduction* was a potential sign of impending overtraining, rather than being due to the considerations mentioned above, was achieved primarily through the coaches' experience, intuition, communication and observation of a rower.

They're your battle-hardened sort of warrior-type athletes, and so when they started falling over it was like you know maybe we need to change.

Knowing what they're like, whether they're hard workers or a bit scared at the beginning or saving themselves for the last one you can understand those people.

Communication was also considered an important cue for impending overtraining and was usually assessed in combination with the performance measures and observational strategies. Two of the coaches spoke at length about the importance of communication to determine the rowers' level of fatigue and performance reduction. Both sought regular feedback and created an environment where rowers and coaches communicated openly.

I see myself more as being like a facilitator than a coach, so yeah I reckon athlete feedback is crucial, it's how I coach, it's just me, I like to know, I like to get the athletes' input.

Well I just talk to them, I say how did that piece feel or how did that session feel … I know how it looked, how did it feel … tell me how you thought [it went] and I can tell you what I saw, and so we have interaction in that way.

Communication with the athlete focused on a variety of issues that included whether fatigue was "excessive" and if so, their perceptions of why this was so. The coaches spoke of asking directly about the rowers' fatigue but also seeking to determine the their mood to gain a better insight into the level of fatigue.

Mood affects the way you act so I think it's a hard one to pick really, you know sometimes you can see it, sometimes you can't, but hopefully if you can see it you try and help them change.

If the rowers' fatigue was "excessive" the coaches spoke of trying to determine whether they were following proper nutrition, hydration, preparatory and recovery processes plus whether they were feeling healthy.

One coach was keen for the athletes to initiate communication with him, but did not proactively communicate with the athletes to determine the level of fatigue they were experiencing during training. He believed that the reason for this stance was that the intensive approach to training meant that that the athletes were often expected to be fatigued. He went as far as to suggest that if they were not heavily fatigued then they were probably not training properly. Hence when asked about fatigue a rower would usually reply that they were fatigued, which in his view provided few useful insights into how to manage the programme. Moreover he didn't think it productive to have the athletes engage in discussions about their level of fatigue.

> The trouble is if you [initiate communication with the rowers] it can start influencing what you do. You know, they're always going to be tired and if you really ask them then you start reducing your training all the time and your programme wouldn't get any momentum.

> If you ask them, they're always going to be tired, they always have to be tired. If you've got someone who's got an injury if you go and ask them they'll tell you about the injury you know, you almost want them to forget about it and move on. Don't think of the pink elephant you know and you think of a pink elephant, how are you feeling, well I'm not feeling good now you come to mention it.

All coaches acknowledged that opening the lines of communication with their rowers could increase their vulnerability to athlete manipulation. Their major concern was a rower accentuating their fatigue as a questionable means of reducing training load.

> I'm probably a bit more savvy now, I know, I can tell if a rower is having me on.

The coaches described the need to balance what they heard versus their observations of the athletes and their performance measures to determine whether the athletes were being honest.

> If someone says they're feeling a little bit [tired], if the speed [has not decreased below normal], well that's bad luck, we'll keep going, but if the speed's really starting to get affected then yeah you button off. So it's a combination of [communication and performance] and you get to know your athletes as a consequence of that. You can tell they'll test you out and when you've got the numbers there to back it up it's pretty easy to tell when … they're trying to have you on.

Observation of the athletes was considered important for determining whether a rower was suffering severe fatigue and at risk of overtraining. The most important observational tool identified by the coaches was the athlete's body language, viewed prior to, during and after training. The coaches described several manifestations of body language including facial expression, how they carried themselves, their demeanour and how they sat or moved in the boat.

The coaches openly described the value of recognising the subtleties of body language. However, recognition could often be problematic and the different coaches tended to concentrate on selected aspects of body language;

> It would just be looking at them, how they're going really. Look in the eyes, the face, the demeanour, the body language. A good coach, … should be there early enough to see them when they arrive and then you get a fair idea of how tired they are, how they carry the boat down to the pontoon will tell you pretty much where you want to be at. But being a top-class athlete is about being tired all the time. You can't be fresh.

> So they're coming up and you're looking at their demeanour and they're tired, sometimes you're thinking they're tired but they can handle it. But there comes that crucial time … and you're making the decision that they can't handle it.

> You look at the person and they actually look [different] and you look at them again and think, I never knew they looked like that. They actually look different you know. I remember looking at [Athlete] and I looked at her again I thought 'oh' and the facial look is totally different and then you know [it was time to] give them three days off [because] I felt I'd driven them down to the point [of overtraining].

> You can see the way they sit in the boat some days they might be slumping in a way they never do or they might be looking lethargic or slow and that's little indicators to me physically that things maybe they're not on top of their game today.

Even though body language was considered important it was also considered important to balance what you saw with the performance numbers.

> You're always watching your rowers, just general mood and things like that you're always keeping an eye on them but it's all quantified against the [boat speeds and rowing ergometer speeds] that you collect.

Another observational strategy described by only one coach was changes in their rowing technique that occurred when they became fatigued. Prolonged periods of poor technique were described as a possible indicator of impending overtraining. The other two coaches didn't agree with this approach and proposed that as the rowers were all elite it was very difficult to see any fatigue related changes in technique. After further discussions with this coach it is our belief that he was talking about efficiency determined by a combination of body language and boat speed. Hence while there may have been some small changes in technique the major change in the fatigued rower was that they looked less efficient, that is they looked to be putting more effort into maintaining their normal boat speed.

> Sometimes you might have the speed there but if they look like they're just working real hard to get their speed so there's things like that what's going on

there it doesn't look that bad but they look like they're just working [really hard] so facial expression and that sort of thing. Then you know again it's so easy on the water because you can see the boat speed and you can tell that they've dropped off.

They're holding the same speed but they're just working harder, but technically working worse, because [they are] trying to muscle it.

Observation was considered an important tool by all the coaches with one coach going so far as to suggest that the body language of the rower is the most important cue for impending overtraining. The coaches spoke about observation being used to assess the level of their athletes' fatigue through changes in body language, movement patterns, reaction speed, posture, facial expression, demeanour, mood, concentration and stress. These largely external indicators collectively present quite a contrast to the largely internal physiological indicators espoused by Meeusen et al.[6]

Each of the coaches was effectively endeavouring to process and inter-relate information from these sources in order to determine when the threshold for impending overtraining had been reached. The over-riding impression from the data is that there are no straightforward rules or a simple formula to inform the decision. Rather the coaches must necessarily rely on their subjective interpretation of the information that they receive for each rower. That information comprises quantitative performance data and qualitative data from conversations and observations. Furthermore, the timeframe that the coach has to make a decision about whether to maintain or reduce the training load compounds the difficulty that coaches face in assessing the risk of overtraining

The coaches acknowledged the difficulty in assessing the point at which the rower was in danger of impending overtraining. This difficulty is highlighted by one of the coaches who stated:

Well I wouldn't just go training for the sake it. If we went out on the water and I figured it just wasn't going to go then no we wouldn't do it. If it's got to the stage where they were too tired then it's no use flogging a dead horse. *But picking that time especially at our level is almost nearly impossible.*

Conclusion

These three coaches believed intensive training was essential for success and that this practice increased the risk of overtraining. They relied on changes in performance, observation and often communication together with knowledge of the rower to determine cues for overtraining. No defined measurement tools were employed rather they favoured subjective appraisal to ascertain the point at which overtraining was imminent and an appropriate decision would be made. Each coach had his own unique approach depending on the rower, the environment and the stage of the training programme. Each approach required the processing of numerous sources of information and any proposed changes weighed up the risks versus benefits of maintaining or reducing the training overload.

Another factor that may also impact the coaches' decision-making was the pressure for continued success created from their achievements in elevating rowing to New Zealand's top Olympic sport. This success led to an increased public profile and greater funding, which in turn resulted in greater numbers of rowers, coaches and support staff all requiring continued success to maintain their positions. The coaches' decisions on how to manage their athletes during the periods of intensive training were often made against a backdrop of athlete, sport code and national expectation, thereby exacerbating an already high-pressure situation. We believe this investigation highlights how a successful coach requires a vast amount of experience and confidence to support what would arguably be a highly intuitive decision.

A comparison of the strategies used by these coaches with the signs and symptoms of overtraining[6, 9, 162] show more differences than similarities or agreement. Agreement existed around decreased performance being the criterion determinant for overtraining. The area of similarity is that some of the psychological variables mentioned in the literature such as demeanour, mood, lethargy, fatigue, recovery, concentration and stress[9, 55, 162, 184] are similar to those mentioned by the coaches. However, while various inventories are recommended to assess these psychological variables, the coaches clearly relied on relatively unstructured techniques namely verbal communication and/or observation. There were acknowledged disparities in the terminologies to describe overtraining, the methods for determining decreases in performance and the degree of performance decrease that defines overtraining. Another area of difference is the many physiological, immunological, biochemical and

psychological signs and symptoms of overtraining, which were largely disregarded by these coaches.

Kennta and Hassmen[162] discussed the importance of monitoring perceptions of both fatigue and recovery as a method of understanding the rowers psychological, social and physiological stress levels. Despite much probing the coaches were reticent about discussing issues related to psychosocial stressors, especially those that originated outside the training or competition environs. The reasons for this remain unclear but we speculate that it is a deliberate tactic, as the coaches do not wish to be perceived as encroaching into their rower's personal lives. Yet throughout the interviews the coaches showed that they were often aware that issues in the rower's personal lives exacerbated their stress levels and effected performance. We believe this important and often-overlooked area requires further examination.

We would argue that the apparent lack of any objective diagnostic marker that accurately predicts impending overtraining[5, 6] has promoted the adoption of various subjective strategies by these coaches. It is more likely that the coaches have selected strategies they understand, are comfortable with, fit their coaching personalities and most importantly, have control over. Furthermore, predicting the point where normal acute fatigue becomes chronic excessive fatigue or overtraining is possibly too subtle and individualistic to ever be determined using scientific measurements alone. Although two of the coaches have tertiary qualifications that include sport science, all three suggested that *the management of athletes during hard training is more of an art than a science*. This perspective is best illustrated through the following reflective comments:

> There is a real art in managing it as opposed to being completely reliant on the science, I think the science substantiates it and gives you a lot of what you need but in terms of really managing it and getting the most out of your athletes … the art side of it is more important.

> Yeah well you're either a scientific coach or you're an intuitive coach, an arty-farty type coach, science, art, one way or another. either way … I guess I'm more that way ... intuitive, arty.

Despite these assertions there is negligible literature on the art of managing athletes against overtraining while the scientific and medical literature abounds. We can only speculate on the possible reason for this state of affairs with possible explanations including successful coaches' unwillingness to divulge their secrets, scientific hegemony in the area of overtraining and the perceived need by all involved for simple objective tools. This investigation highlights how coach decision-making is strongly influenced by a subjective process, a process that to this point has sat in the too hard basket for too long.[185] We believe the cues for overtraining used by these successful elite endurance coaches during their management of training overload provide valuable insights that should be further explored. Whilst this investigation has revealed subjective coach strategies for managing overtraining that are arguably open to scrutiny, the continued international success of these three coaches cannot be ignored.

Lyle,[166] requested that "[f]ar greater attention is required to be paid to the coach's cognitive processes. ...Performance enhancement research cannot be complete without attention to the coach's application, integration and delivery of knowledge" (p. 303-304). In response to this request, we suggest that these coaches' decisions about overtraining were based heavily on their subjective knowledge but the decision-making process they employed included multiple markers to support a perspective of self-consistency. Through the accumulation and processing of these markers each coach was able to express their confidence in their decisions. There was not a propensity toward a specific marker, rather each coach participated in a search for supporting and inter-related elements. While the presentation of these findings should not favour subjective knowledge as an absolute replacement for its objective equivalent, we would argue that subjective knowledge provides a better understanding about coaches' decision-making.

Decision-making is but one component of a complex web of thoughts and actions that interplay within natural settings indicating there is scope for related frameworks to be applied to enhance our understanding of the shifts from concept to context that play out in the coaching world. This research has highlighted the need for further conceptual and empirical investigation in a range of coaching contexts.

CHAPTER 6

General Discussion

This PhD consists of five research projects that were performed to investigate various measures for monitoring performance in elite rowers. Here I will summarise the outcomes of our research and make some recommendations for best practice in monitoring rowing performance in applied and research settings.

In Study 1 the race-to-race variability of boat times was ~1.0% which was used to estimate the smallest worthwhile effect size in performance time of ~0.3%. For any intervention this effect size signifies the smallest worthwhile enhancement in performance that would effect a change in medal prospects. This smallest worthwhile effect represents a desirable target for the error of measurement in rowing performance tests that would be sufficiently sensitive to quantify trivial-small changes in performance. These findings provide important benchmarks for the accurate monitoring of rowers in applied and research settings.

Another interesting finding from this study was the extremely large variability in competition times created by environmental conditions, which has the potential to create an advantage for some crews over others. After discussion of this study with Mike Tanner (International Rowing Federation technical delegate responsible for the operation of International rowing events), I was asked to submit a written report on the issues of fairness during the 2010 World Rowing Championships. Armed with the results from Study 1, wind data from the 2010 championships and analysis of the 2010 the world cup race results, I was able to show that the environmental conditions during the Friday's championship finals created unfair race results. Furthermore, with data I collected during the 2006 World Rowing Championships, which was held on the London Olympic rowing course, I projected that similar unfair conditions could occur during the London Olympics. I therefore recommended that under certain environmental conditions the International Rowing Federation should change their seeding system so the better-seeded crews are assigned lanes with more favorable environmental conditions. The International Rowing Federation received this report positively.

Armed with the error benchmarks established in Study 1 the reliability and validity of the measures of rowing performance reported in the literature was investigated (Study 2). Few studies have examined on-water measures, which is probably due various factors that include changing environmental conditions and logistical difficulties. We therefore performed our own analysis of an NK impeller and a GPS device for inclusion in the review. The GPS unit has negligible measurement error over 2000 m but is only useful for monitoring an individual's performance when rowing in a single scull under ideal environmental conditions. Mobile ergometry via instrumented oars or rowlocks could overcome the environmental effects and measure individual performance but the accuracy of these devices are yet to be reported.

Given the limitations to measuring on-water performance it is not surprising that the most widely used measure of rowing performance is the 2000-m time-trial on the Concept II rowing ergometer   It was determined from the literature review that this test has a typical error suitable for tracking changes in physiological performance and factors affecting it. Other measures that under ideal testing conditions have typical error between repeated tests on a Concept II suitably low for tracking changes include peak power output in an incremental test, some measures of lactate threshold, and measures of 30-s all-out power. However, the standard error of the estimate of on-water 2000-m time predicted by most of these Concept II measures was very large, probably reflecting different effects of skill, body mass and environment in on-water vs ergometer performance. The poor reported validity of these Concept II measures is likely the result of environmental variability. Further studies with an appropriate sample size of top rowers under ideal environmental conditions should confirm the validity of these measures. In the last few years of my work with Rowing New Zealand these findings contributed to the reshaping of the Rowing New Zealand athlete-testing program, in particular the inclusion of GPS monitoring for all rowing training and the removal of a number of the inaccurate laboratory tests.

Many researchers have proposed that increases in stress assayed by changes in various physiological and psychological markers can be used to predict non-functional overreaching and overtraining syndrome. Indeed, the conventional belief amongst sport scientists is that a large increase in stress, measured by the change in one or more stress markers during a period of intensive training, will predict impending non-functional overreaching or overtraining syndrome. This belief has led to the regular measurement of various physiological and psychological stress markers in elite athletes to predict whether training load is excessive. Despite this practice, it is generally accepted that no

marker is a consistently reliable and valid predictor of non-functional overreaching or overtraining syndrome.[5, 6]

An important issue with research examining the relationship between biomarkers and performance is error of measurement, which contributes noise to the real changes in a marker. If the error is small relative to the standard deviation of the real changes in the marker, then obviously the error will have little or no impact on the relationship between changes in the marker and changes in performance. On the other hand, if the error of measurement is large compared with the real change scores, the observed change scores will be almost entirely noise, so there can be little or no relationship between the observed change scores and performance regardless of any true underlying relationship. This effect of error of measurement is an issue that researchers habitually overlook. As a case in point, none of the reviewers of the study published in Chapter 3 of this thesis queried the lack of any reported measurement errors.

Error of measurement for biomarkers is usually reported as the intra-assay coefficients of variation, which represents the typical error or variation in a measurement when the sample is reanalysed. As such, it can also be called the technical error of measurement, because it is the noise contributed by the measurement process. For Study 3 the intra-assay coefficients of variation were 3.8% for CK, 3.0% for LDH, 10.7% for cortisol, 4.0% for testosterone, 10.2% for DHEA, and 8.7% for the cytokines, which are on par with those reported in the literature.[136, 186-195] To compare these errors with the change scores over the 4 weeks of the study, they need to be multiplied by $\sqrt{2}$ to produce the standard deviation of a change score that would be due entirely to error of measurement.[196] The resulting errors and the standard deviation of the change scores are shown in Appendix M.

The change score for the majority of the markers were well in excess of the error of measurement, so the effects involving these markers have probably not been attenuated substantially. However, the observed change scores for morning cortisol in both genders, for within week change in testosterone in men, and for salivary IL6 in both genders, consist almost entirely of error of measurement. Unsurprisingly, the uncertainty in the effect of these markers allowed for the effects to be trivial. These markers may have substantial effects on performance, but to observe them would require assays with smaller errors of measurement.

From a practical perspective there is interest in the utility of markers for predicting change in performance. The goal is to explore which markers are effective predictors of

change using available assays. Some markers will fail to predict substantial change either because of error of measurement or because they do not relate to changes in performance; either way, the marker is not useful. The search for useful markers is important for practitioners aiming to optimize performance and reduce the risk of overtraining syndrome

In Chapters 3 and 4 there were many instances where large increases in a stress marker during the period of overload signalled effective training rather than overtraining. These findings are contrary to the previously mentioned conventional belief, and I propose the following reason for this disparity. Large training stress has two possible outcomes: the first is a greater risk of non-functional overreaching or overtraining syndrome, and the second is based on the general adaptation syndrome, which describes how greater stress can also lead to greater adaptation[1, 197]. Therefore, an excessive negative change in one or more stress markers may have a range of responses that include optimal conditioning or non-functional overreaching/overtraining syndrome. What distinguishes between positive or negative adaptation may not necessarily be the change in the stress marker, but the ability of the athlete to continue to adapt positively in response to these large increases in stress.

The only method of predicting non-functional overreaching or overtraining syndrome is probably via measurements of performance. Even if an association between the change in a marker and overtraining syndrome could be established, it is likely the level of change in the marker that previously predicted overtraining syndrome, would not do so in the future. This could be due to a number of factors that include changes in other contributing factors, improved "toughness" of the athlete, receptor-site down-regulation, and metabolic changes.

It is likely that changes in one or more stress markers lack the sensitivity and specificity to predict non-functional overreaching or overtraining syndrome. These markers may therefore be more useful for adjusting training load to ensure appropriate increases in stress occur. This is an important issue, which I hope to explore further. It is possible that these findings will not apply to some sub-elite athletes, for whom large increases in some marker or markers of stress will predict overtraining syndrome.

During the final stages of this PhD I had the opportunity to put a number of these findings into practice with an elite athlete who suffered non-functional overreaching. A number of the markers described in this thesis were monitored, but only a few of the psychological measures showed abnormal changes. However, these psychological

measures only became abnormal once the rower's performance had reduced to the level that resulted in loss of their position in the team. I subsequently monitored their progress solely through the use of performance measures. After a 7-month intervention, the rower regained a position on the New Zealand elite rowing team and went on to record a career best result at a world rowing championship. There was a similar scenario with another rower, but for reasons of confidentiality I cannot record details here.

It is my experience that the incidences of non-functional overreaching and overtraining syndrome are low in elite rowers. However, quantification of the rates of non-functional overreaching and overtraining syndrome in elite rowers and the possible causative factors are yet to be published. The few incidences that I have witnessed have occurred in the better, highly motivated, younger rowers during preparations for big events such as the Olympics. The loss of a good athlete in any sports program is difficult to accept, so it is important to develop effective monitoring strategies to prevent such loss.

Despite the inability of the various markers to predict overtraining syndrome, most successful elite rowing coaches are able to manage the balance between intense training that improves conditioning versus that which decreases it. This ability was the incentive for the final study, which explored the practices and beliefs of three elite rowing coaches in the management of their athletes training. These coaches believed that intensive training was essential for success and that this practice increased the risk of performance reductions associated with overtraining syndrome.

The coaches all discussed the importance of constantly monitoring their athletes, especially during intensive training. They monitored their rowers through observation, communication and measures of training pace. In general, the intuitions and cues for overtraining they employed had little in common with those usually reported in academic print media. We suggest that these differences are due to the coaches' decision-making being based largely on subjective processes and influenced by various stressors unique to their positions.

Despite the successes and experience of these three coaches, they were all aware that their monitoring strategies were not foolproof, and that occasionally a rower suffered a bout of chronic fatigue. While the choice of these subjective strategies may be challenged on the grounds of their reliability and validity, there seem to be no alternatives. Further research is required to determine whether these strategies are

employed by other coaches in other sports, or whether it is an isolated phenomenon. If the use of these strategies were widespread amongst successful endurance coaches, then teaching these skills in coach-education programs would be an important development.

Near the end of this thesis I refocused the New Zealand Rowing sport-science program more towards the development of reliable and valid tools and processes for monitoring individual and crew performance. In particular I contributed to the development and implementation of performance-monitoring tools (GPS, impellers and force gates) and interpretative models (models of training speed and stroke rate, effect of environmental variations and normal reductions in workload due to acute fatigue). These tools and models are under development, but they are already an integral part of the Rowing New Zealand elite program. Their continued development provides an important opportunity for sport scientists to add value to elite rowing programs, specifically in monitoring for overtraining syndrome and providing objective feedback to help optimise the training program. Accurate measures of training performance are also important for the coach as an objective measure to compare against their subjective assessments of body language and communication.

Future Directions

We encourage rowers, sports scientists and coaches to estimate errors of measurement and magnitudes of effects of treatments and individual responses to rowing interventions using the methods outlined in Studies 1 to 4.

Specific research projects in the near future include:

- The accuracy of on-water ergometry for measuring individual rowers' power output and how the measured power represents power propelling the boat forward;

- The validity of 2000-m time-trials on a Concept II rowing ergometer for predicting 2000-m on-water time-trials in a single scull with an appropriate sample size of top rowers under ideal environmental conditions;

- Whether large increase in creatine kinase and possibly other physiological and psychological stress markers may be more indicative of an effective training overload than of impending maladaptation.

- Examine the incidence rates of non-functional overreaching and overtraining syndrome in elite rowers globally, and investigate the causative factors.

- How succesful coaches from various endurance sports monitor elite endurance athletes during intensive training.

COMPILED REFERENCE LIST

1. Bompa TO. Periodisation: theory and methodology of training. 4th ed. Champaign, IL: Human Kinetics, 1999

2. Bompa TO, Carrera M. Periodization training for sports: Human Kinetics Champaign, IL, 1999

3. O'Toole ML. Overreaching and overtraining in endurance athletes. In: Kreider RB, Fry AC, O'Toole ML, editors. Overtraining in sport. Champaign, IL: Human Kinetics; 1998: 3-17

4. Gould D, Greenleaf C, Chung Y, Guinan D. A survey of US Atlanta and Nagano Olympians: Variables perceived to influence performance. Res Q Exerc Sport 2002; 73: 175-86

5. Halson SL, Jeukendrup AE. Does overtraining exist? An analysis of overreaching and overtraining research. Sports Med 2004; 34: 967-81

6. Meeusen R, Duclos M, Gleeson M, Rietjens G, Steinacker J, Urhausen A. Prevention, diagnosis and treatment of the overtraining syndrome. Eur J Sport Sci 2006; 6: 1-14

7. Kreider RB, Fry AC, O'Toole ML. Overtraining in sport: Terms definitions and prevalence. In: Kreider RB, Fry AC, O'Toole ML, editors. Overtraining in sport. Champaign, IL: Human Kinetics; 1998: vii-x

8. Kuipers H, Keizer HA. Overtraining in elite athletes: review and directions for the future. Sports Med 1988; 6: 79-92

9. Fry RW, Morton AR, Keast D. Overtraining in athletes. An update. Sports Med 1991; 12: 32-65

10. Hopkins WG. Quantification of training in competitive sports. Methods and applications. Sports Med 1991; 12: 161-83

11. Urhausen A, Kindermann W. Diagnosis of overtraining: what tools do we have? Sports Med 2002; 32: 95-102

12. Blomstrand E, Celsing F, Newsholme EA. Changes in plasma concentrations of aromatic and branched-chain amino acids during sustained exercise in man and their possible role in fatigue. Acta Physiol Scand 1988; 133: 115-21

13. Opstad P. The hypothalamo-pituitary regulation of androgen secretion in young men after prolonged physical stress combined with energy and sleep deprivation. Acta Endocrinol (Copenh) 1992; 127: 231-6

14. Snyder A. Overtraining and glycogen depletion hypothesis. Med Sci Sports Exerc 1998; 30: 1146-50

15. Urhausen A, Gabriel HH, Kindermann W. Impaired pituitary hormonal response to exhaustive exercise in overtrained endurance athletes. Med Sci Sports Exerc 1998; 30: 407-14

16. Koutedakis Y, Frischknecht R, VrbovÁ G, Sharp NCC, Budgett R. Maximal voluntary quadriceps strength patterns in Olympic overtrained athletes. Med Sci Sports Exerc 1995; 27: 566

17. Billat VL, Flechet B, Petit B, Muriaux G, Koralsztein JP. Interval training at VO2max: effects on aerobic performance and overtraining markers. Med Sci Sports Exerc 1999; 31: 156

18. Smith LL. Tissue trauma: the underlying cause of overtraining syndrome. J Str Cond Res 2004 2004; 18: 185-93

19. Vervoorn C, Quist AM, Vermulst LJ, Erich WB, de Vries WR, Thijssen JH. The behaviour of the plasma free testosterone/cortisol ratio during a season of elite rowing training. Int J Sports Med 1991; 12: 257-63

20. Barron JL. Hypothalamic dysfunction in overtrained athletes. J Clin Endocrinol Metab 1985; 60: 803-6

21. Uusitalo ALT, Huttunen P, Hanin Y, Uusitalo AJ, Rusko HK. Hormonal responses to endurance training and overtraining in female athletes. Clin J Sport Med 1998; 8: 178-86

22. Keizer HA. Neuroendocrine aspects of overtraining. In: Kreider RB, Fry AC, O'Toole ML, editors. Overtraining in sport. Champaign, Il: Human Kinetics; 1998

23. Lehmann M, Foster C, Keul J. Overtraining in endurance athletes: a brief review. Med Sci Sports Exerc 1993; 25: 854-62

24. Meeusen R, Piacentini MF, Busschaert B, Buyse L, Schutter GD, Stray-Gundersen J. Hormonal responses in athletes: the use of a two bout exercise protocol to detect subtle differences in (over) training status. Eur J Appl Physiol 2004; 91: 140-6

25.   Steinacker JM, Lormes W, Kellmann M, Liu Y, Reissnecker S, Opitz-Gress A, et al. Training of junior rowers before world championships. Effects on performance, mood state and selected hormonal and metabolic responses. J Sports Med Phys Fit 2000; 40: 327-35

26.   Steinacker JM, Lormes W, Reissnecker S, Liu Y. New aspects of the hormone and cytokine response to training. Eur J Appl Physiol 2004; 91: 382-91

27.   Urhausen A, Gabriel H, Kindermann W. Blood hormones as markers of training stress and overtraining. Sports Med 1995; 20: 251-76

28.   Maestu J, Jurimae J, Jurimae T. Hormonal responses to maximal rowing before and after heavy increase in training volume in highly trained male rowers. J Sports Med Phys Fit 2005; 45: 121-6

29.   Steinacker JM. Physiological aspects of training in rowing. Int J Sports Med 1993; 14: 3-10

30.   Lucia A, Diaz B, Hoyos J, Fernandez C, Villa G, Bandres F, et al. Hormone levels of world class cyclists during the Tour of Spain stage race. Br J Sports Med 2001; 35: 424-30

31.   Kraemer WJ, Loebel CC, Volek JS, Ratamess NA, Newton RU, Wickham RB, et al. The effect of heavy resistance exercise on the circadian rhythm of salivary testosterone in men. Eur J Appl Physiol 2001; 84: 13-8

32.   Vining RF, McGinley RA. The measurement of hormones in saliva: possibilities and pitfalls. J Steroid Biochem 1987; 27: 81-94

33.   Mackinnon LT. Overtraining effects on immunity and performance in athletes. Immunol Cell Biol 2000; 78: 502-9

34.   Nieman DC, Nehlsen-Cannarella SL, Fagoaga OR, Henson DA, Shannon M, Hjertman JME, et al. Immune function in female elite rowers and non-athletes. Brit J Sports Med 2000; 34: 181-7

35.   Sharp NCC, Koutedakis Y. Sport and the overtraining syndrome: immunological aspects. Br Med Bull 1992; 48: 518-33

36.   Smith LL. Cytokine hypothesis of overtraining: a physiological adaptation to excessive stress? Med Sci Sports Exerc 2000; 32: 317-31

37.   Smith LL. Overtraining, excessive exercise, and altered immunity: is this a T helper-1 versus T helper-2 lymphocyte response? Sports Med 2003; 33: 347-64

38. Gleeson M. Biochemical and immunological markers of overtraining. J Sport Sci Med 2002; 1: 31-41

39. Jürimäe J, Mäestu J, Jürimäe T. Leptin as a marker of training stress in highly trained male rowers? Eur J Appl Physiol 2003; 90: 533-8

40. Yuan Y, Kwong AWK, Kaptein WA, C. F, Tse M, C. GJF, et al. The response of fatty acid-binding protein and creatine kinase to acute and chronic exercise in junior rowers. Res Q Exerc Sport 2003; 74: 277-83

41. Brancaccio P, Maffulli N, Limongelli FM. Creatine kinase monitoring in sport medicine. Br Med Bull 2007:

42. Rowbottom DG, Keast D, Goodman C, Morton AR. The haematological, biochemical and immunological profile of athletes suffering from the overtraining syndrome. Eur J Appl Physiol 1995; 70: 502-9

43. Mäestu J, Jürimäe J, Jürimäe T. Monitoring of performance and training in rowing. Sports Med 2005; 35: 597-617

44. Meyers AW, Whelan JP, editors. A systemic model for understanding psychosocial influences in overtraining. Champaign, IL: Human Kinetics; 1998

45. Prapavessis H. The POMS and sports performance: A review. J Appl Sport Psychol 2000; 12: 34-48

46. McNair D, Lorr M, Droppleman L. Profile of mood states (POMS). San Diego: Educational and Industrial Testing Service; 1989

47. Hawley CJ, Schoene RB. Overtraining syndrome: a guide to diagnosis, treatment, and prevention. Physician Sportsmed 2003; 31: 25-31

48. Purge P, Ülikool T. Performance, mood state and selected hormonal parameters during the rowing season in elite male rowers: Tartu University Press, 2006

49. Beedie CJ, Terry PC, Lane AM. The profile of mood states and athletic performance: two meta-analyses. J Appl Sport Psychol 2000; 12: 49-68

50. LeUnes A, Burger J. Bibliography on the profile of mood states in sport and exercise, 1971-1995. J Sport Behav 1998; 21: 53-70

51. Crawford JR, Henry JD. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol 2004; 43: 245

52.   Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988; 54: 1063-70

53.   Borg G. Borg's perceived exertion and pain scales. Champaign, Il: Human Kinetics, 1998

54.   Noble BJ, Noble JM. Perceived exertion: the measurement. Champaign, Il: Human Kinetics Publishers, 1998

55.   Hooper SL, Mackinnon LT, Howard ALF, Gordon RD, Bachmann AW. Markers for monitoring overtraining and recovery. Med Sci Sports Exerc 1995; 27: 106-12

56.   Malik M, Bigger JT, Camm AJ, Kleiger RE, Malliani A, Moss AJ, et al. Guidelines: heart rate variability standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Eur Heart J 1996; 17: 354-81

57.   Hedelin R, Wiklund U, Bjerle PER, Henriksson-Lars…N K. Cardiac autonomic imbalance in an overtrained athlete. Med Sci Sports Exerc 2000; 32: 1531-3

58.   Lehmann M, Foster C, Dickhuth HH, Gastmann UWE. Autonomic imbalance hypothesis and overtraining syndrome. Med Sci Sports Exerc 1998; 30: 1140-5

59.   Hedelin R, Wiklund U, Bjerle P, Henriksson-Larsen K. Pre-and post-season heart rate variability in adolescent cross-country skiers. Scand J Med Sci Sport 2000; 10: 298-303

60.   Hynynen E, Uusitalo A, Konttinen N, Rusko H. Heart rate variability during night sleep and after awakening in overtrained athletes. Med Sci Sports Exerc 2006; 38: 313-7

61.   Mourot L, Bouhaddi M, Perrey S, Cappelle S, Henriet MT, Wolf JP, et al. Decrease in heart rate variability with overtraining: assessment by the Poincare plot analysis. Clin Physiol 2004; 24: 10-8

62.   Achten J, Jeukendrup AE. Heart rate monitoring: applications and limitations. Sports Med 2003; 33: 517

63.   Aubert AE, Seps B, Beckers F. Heart rate variability in athletes. Sports Med 2003; 33: 889-919

64. Baumert M, Brechtel L, Lock J, Hermsdorf M, Wolff R, Baier V, et al. Heart rate variability, blood pressure variability, and baroreflex sensitivity in overtrained athletes. Clin J Sport Med 2006; 16: 412

65. Hopkins WG, Hawley JA, Burke LM. Design and analysis of research on sport performance enhancement. Med Sci Sports Exerc 1999; 31: 472-85

66. Bonetti DL, Hopkins WG. Variation in performance times of elite flat-water canoeists from race to race. Int J Sports Physiol Perform 2010; 5: 210-7

67. Pyne DB, Trewin CB, Hopkins WG. Progression and variability of competitive performance of Olympic swimmers. J Sports Sci 2004; 22: 613-20

68. Paton CD, Hopkins WG. Variation in performance of elite cyclists from race to race. Eur J Sport Sci 2006; 6: 25-31

69. Bullock N, Hopkins WG, Martin DT, Marino FE. Characteristics of performance in skeleton World Cup races. J Sports Sci 2009; 27: 367-72

70. Nibali ML, Hopkins WG, Clark SA, Drinkwater EJ. Variability and predictability of elite competitive slalom canoe-kayak performance. Eur J Sport Sci 2011; 11: 125-30

71. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc 2009; 41: 3-13

72. Hopkins WG. A spreadsheet for combining outcomes from several subject groups. Sportscience 2006; 10: 51-3

73. Drinkwater EJ, Hopkins WG, McKenna MJ, Hunt PH, Pyne DB. Modelling age and secular differences in fitness between basketball players. J Sports Sci 2007; 25: 869-78

74. Hopkins WG. Competitive performance of elite track-and-field athletes: variability and smallest worthwhile enhancements. Sportscience 2005; 9: 17-20

75. Hopkins WG, Schabort EJ, Hawley JA. Reliability of power in physical performance tests. Sports Med 2001; 31: 211-34

76. Hopkins WG. How to interpret changes in an athletic performance test. Sportscience 2004; 8: 1-7

77. Soper C, Hume PA. Reliability of power output during rowing changes with ergometer type and race distance. Sports Biomech 2004; 3: 237-48

78.   Schabort EJ, Hawley JA, Hopkins WG, Blum H. High reliability of performance of well-trained rowers on a rowing ergometer. J Sports Sci 1999; 17: 627-32

79.   Hopkins WG. Estimating sample size for magnitude-based inferences. Sportscience 2006; 10: 63-70

80.   Hagerman FC, Hagerman GR, Mickelson TC. Physiological profiles of elite rowers. Physician Sportsmed 1979; 7: 74-81

81.   De Campos Mello F, De Moraes Bertuzzi RC, Grangeiro PM, Franchini E. Energy systems contributions in 2,000 m race simulation: a comparison among rowing ergometers and water. Eur J Appl Physiol 2009; 107: 615-9

82.   Hagerman FC, Connors MC, Gault JA, Hagerman GR, Polinski WJ. Energy expenditure during simulated rowing. J Appl Physiol 1978; 45: 87-93

83.   Secher NH. The physiology of rowing. J Sports Sci 1983; 1: 23-53

84.   Mikulic P. Anthropometric and physiological profiles of rowers of varying ages and ranks. Kinesiology 2008; 40: 80-8

85.   Shephard RJ. Science and medicine of rowing: a review. J Sports Sci 1998; 16: 603-20

86.   Smith TB, Hopkins WG. Variability and predictability of finals times of elite rowers. Med Sci Sports Exerc 2011; 43: 2155-60

87.   Kleshnev V. Estimation of biomechanical parameters and propulsive efficiency of rowing. 1998 [cited; 1-17]. Available from: http://www.biorow.com/Papers.htm

88.   Kleshnev V. Propulsive efficiency of rowing In: Sanders RH, Gibson NR, editors. Proceedings of the XVII International Symposium on Biomechanics in Sports; 1999: School of Biomedical and Sport Science, Edith Cowan University, Perth, Australia; 1999: 224-8

89.   McArthur J. High performance rowing. Wiltshire, UK: The Crowood press, 1997

90.   Kleshnev V. Prognostic times. Rowing Biomechanics Newsletter; 2005. 1

91.   Patterson A. Review of rowing technology. World Rowing Coaches Conference. Windsor, UK: International Rowing Federation; 2011. 1-30

92.   Langmuir I. Surface motion of water induced by wind. Science 1938; 87: 119-23

93.   Wu J. Wind-induced drift currents. J Fluid Mech 1975; 68: 49-70

94. Kleshnev V. Weather and boat speed. Rowing Biomechanics Newsletter; 2009. 1-3

95. Pulman C. The physics of rowing: University of Cambridge: Ithaca, 2005

96. Witte TH, Wilson AM. Accuracy of non-differential GPS for the determination of speed over ground. J Biomech 2004; 37: 1891-8

97. Duffield R, Reid M, Baker J, Spratford W. Accuracy and reliability of GPS devices for measurement of movement patterns in confined spaces for court-based sports. J Sci Med Sport 2010; 13: 523-5

98. Jennings D, Cormack S, Coutts AJ, Boyd L, Aughey RJ. The validity and reliability of GPS units for measuring distance in team sport specific running patterns. Int J Sports Physiol Perform 2010; 5: 328-41

99. Petersen C, Pyne D, Portus M, Dawson B. Validity and reliability of GPS units to monitor cricket-specific movement patterns. Int J Sports Physiol Perform 2009; 4: 381-93

100. Pyne DB, Petersen C, Higham DG, Cramer MN. Comparison of 5-And 10-hz gps technology for team sport analysis. Med Sci Sports Exerc 2010; 42: 78

101. Caplan N, Gardner T. Modeling the influence of crew movement on boat velocity fluctuations during the rowing stroke. Int J Sport Sci Eng 2007; 1: 165-76

102. Hill H, Fahrig S. The impact of fluctuations in boat velocity during the rowing cycle on race time. Scand J Med Sci Sport 2009; 19: 585-94

103. Vogler A, Lindg A, Rice A. Accuracy and reliability of Minimaxx GPS technology in rowing [abstract]. In: Cabri J, Alves F, Araújo D, Barreiros J, Diniz J, Veloso A, editors. Proceedings of the 13th annual Congress of the European College of Sport Science; 2008; Estoril, Portugal: Gamlebyen grafiske; 2008: 413

104. Kleshnev V. Comparison of measurements of the force at the oar handle and at the gate or pin. Rowing Biomechanics Newsletter; 2010. 1-3

105. Baudouin A, Hawkins D. Investigation of biomechanical factors affecting rowing performance. J Biomech 2004; 37: 969-76

106. Smith R, Galloway M, Patton R, Spinks W. Analysing on-water rowing performance. Sports Coach 1994; 17: 37-40

107. Lamb DH. A kinematic comparison of ergometer and on-water rowing. Am J Sports Med 1989; 17: 367-73

108. Kleshnev V. Comparison of on-water rowing with its simulation on Concept2 and Rowperfect machines In: Wang Q, editor. International Symposium on Biomechanics in Sports; 2005; Beijing; 2005: 130-3

109. Dawson RG, Lockwood RJ, Wilson JD, Freeman G. The rowing cycle: sources of variance and invariance in ergometer and on-the-water performance. J Motor Behav 1998; 30: 33-43

110. Hahn A, Bourdon P, Tanner R. Protocols for the physiological assessment of rowers. In: Gore CJ, editor. Physiological tests for elite athletes. Champaign, Il: Human Kinetics; 2000: 311-26

111. Vogler AJ, Rice AJ, Withers RT. Physiological responses to exercise on different models of the concept II rowing ergometer. Int J Sports Physiol Perform 2007; 2: 360-70

112. Benson A, Abendroth J, King D, Swensen T. Comparison of rowing on a Concept 2 stationary and dynamic ergometer. J Sport Sci Med 2011; 10: 267-73

113. Holsgaard-Larsen A, Jensen K. Ergometer Rowing With and Without Slides. Int J Sports Med 2010; 31: 870-4

114. Colloud F, Bahuaud P, Doriot N, Champely S, Cheze L. Fixed versus free-floating stretcher mechanism in rowing ergometers: mechanical aspects. J Sports Sci 2006; 24: 479-93

115. Elliott E, Lyttle A, Birkett O. The RowPerfect Ergometer: a training aid for on-water single scull rowing. Sports Biomech 2002; 1: 123-34

116. Mikulic P, Smoljanovi T, Bojani I, Hannafin JA, Matkovi BR. Relationship between 2000-m rowing ergometer performance times and World Rowing Championships rankings in elite-standard rowers. J Sports Sci 2009; 27: 907-13

117. Mikulic P, Smoljanovic T, Bojanic I, Hannafin J, Pedisic Z. Does 2000-m rowing ergometer performance time correlate with final rankings at the World Junior Rowing Championship? A case study of 398 elite junior rowers. J Sports Sci 2009; 27: 361-6

118. Jurimae J, Maestu J, Jurimae T, Pihl E. Prediction of rowing performance on single sculls from metabolic and anthropometric variables. J Hum Movement Stud 2000; 38: 123-36

119. Nevill AM, Beech C, Holder RL, Wyon M. Scaling concept II rowing ergometer performance for differences in body mass to better reflect rowing in water. Scand J Med Sci Sport 2010; 20: 122-7

120. Nevill AM, Allen SV, Ingham SA. Modelling the determinants of 2000 m rowing ergometer performance: a proportional, curvilinear allometric approach. Scand J Med Sci Sport 2011; 21: 73-8

121. Bourdin M, Messonnier L, Hager JP, Lacour JR. Peak power output predicts rowing ergometer performance in elite male rowers. Int J Sports Med 2004; 25: 368-73

122. Cosgrove MJ, Wilson J, Watt D, Grant SF. The relationship between selected physiological variables of rowers and rowing performance as determined by a 2000 m ergometer test. J Sports Sci 1999; 17: 845-52

123. Faff J, Bienko A, Burkhard-Jagodzinska K, Borkowski L. Diagnostic value of indices derived from the critical power test in assessing the anaerobic work capacity of rowers. Biol Sport 1993; 10: 9-14

124. Gillies EM, Bell GJ. The relationship of physical and physiological parameters to 2000 m simulated rowing performance. Res Sports Med 2000; 9: 277-88

125. Riechman SE, Zoeller RF, Balasekaran G, Goss FL, Robertson RJ. Prediction of 2000 m indoor rowing performance using a 30 s sprint and maximal oxygen uptake. J Sports Sci 2002; 20: 681-7

126. Russell AP, Le Rossignol PF, Sparrow WA. Prediction of elite schoolboy 2000-m rowing ergometer performance from metabolic, anthropometric and strength variables. J Sports Sci 1998; 16: 749-54

127. Womack CJ, Davis SE, Wood CM, Sauer K, Alvarez J, Weltman A, et al. Effects of training on physiological correlates of rowing ergometry performance. J Str Cond Res 1996; 10: 234-8

128. Bourdon PC, David AZ, Buckley JD. A single exercise test for assessing physiological and performance parameters in elite rowers: the 2-in-1 test. J Sci Med Sport 2009; 12: 205-11

129. Hopkins WG. Adjusting regression statistics to assess validity of measures in different populations [abstract]. In: Loland S, Bø K, Fasting K, Hallén J, Ommundsen Y, Roberts G, et al., editors. Proceedings of the 14th annual

Congress of the European College of Sport Science; 2009; Oslo, Norway: Gamlebyen grafiske; 2009: 413

130. Aldercreutz H, Harkonen K, Kuoppasalmi H, Naveri H, Huhtaniemi H, Timsanen H. Effect of training on plasma anabolic and catabolic steroid hormone and their physical exercise. J Sports Med 1986; 30: 407-14

131. Urhausen A, Gabriel HH, Weiler B, Kindermann W. Ergometric and psychological findings during overtraining: a long-term follow-up study in endurance athletes. Int J Sports Med 1998; 19: 114-20

132. Hartmann U, Mester J. Training and overtraining markers in selected sport events. Med Sci Sports Exerc 2000; 32: 209

133. Halson SL, Bridge MW, Meeusen R, Busschaert B, Gleeson M, Jones DA, et al. Time course of performance changes and fatigue markers during intensified training in trained cyclists. J Appl Physiol 2002; 93: 947-56

134. Hooper SL, Mackinnon LT, Hanrahan S. Mood state as an indication of staleness and recovery. Int J Sports Psychol 1997; 20: 321-7

135. Morgan WP, Brown DR, Raglin JS, O'Connor PJ, Ellickson KA. Psychological monitoring of overtraining and staleness. Br J Sports Med 1987; 21: 107-14

136. Main LC, Dawson B, Heel K, Grove JR, Landers GJ, Goodman C. Relationship between inflammatory cytokines and self-report measures of training overload. Res Sports Med 2010; 18: 127-39

137. Granger DA, Schwartz EB, Booth A, Arentz M. Salivary testosterone determination in studies of child health and development. Horm Behav 1999; 35: 18-27

138. Granger DA, Schwartz EB, Booth A, Curran M, Zakaria D. Assessing dehydroepiandrosterone in saliva: a simple radioimmunoassay for use in studies of children, adolescents and adults. Psychoneuroendocrinology 1999; 24: 567-79

139. Morelius E, Nelson N, Theodorsson E. Salivary cortisol and administration of concentrated oral glucose in newborn infants: improved detection limit and smaller sample volumes without glucose interference. Scand J Clin Lab Invest 2004; 64: 113-8

140. Schumann G, Aoki R, Ferrero CA, Ehlers G, Ferard G, Gella FJ, et al. IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 °C. Clin Chem Lab Med 2006; 44: 1146-55

141. Howell BF, McCune S, Schaffer R. Lactate-to-pyruvate or pyruvate-to-lactate assay for lactate dehydrogenase: a re-examination. Clin Chem 1979; 25: 269-72

142. Uusitalo AL, Uusitalo AJ, Rusko HK. Heart rate and blood pressure variability during heavy training and overtraining in the female athlete. Int J Sports Med 2000; 21: 45-53

143. Lamberts RP, Swart J, Capostagno B, Noakes TD, Lambert MI. Heart rate recovery as a guide to monitor fatigue and predict changes in performance parameters. Scand J Med Sci Sport 2010; 20: 449-57

144. Barrett RS, Manning JM. Relationships between rigging set-up, anthropometry, physical capacity, rowing kinematics and rowing performance. Sports Biomech 2004; 3: 221-35

145. Heck H, Mader A, Hess G, Mücke S, Müller R, Hollmann W. Justification of the 4-mmol/L lactate threshold. Int J Sports Med 1985; 6: 117-30

146. Mader A, Heck H. A theory of the metabolic origin of anaerobic threshold. Int J Sports Med 1986; 7: S45-S65

147. Maassen N, Busse MW. The relationship between lactic acid and work load: a measure for endurance capacity or an indicator of carbohydrate deficiency? Eur J Appl Physiol 1989; 58: 728-37

148. Hynynen E, Uusitalo A, Konttinen N, Rusko H. Heart rate variability during night sleep and after awakening in overtrained athletes. Med Sci Sports Exerc 2006; 38: 313-7

149. Hedelin R, Kentt¡ G, Wiklund U, Bjerle PER, Henriksson-Lars NK. Short-term overtraining: effects on performance, circulatory responses, and heart rate variability. Med Sci Sports Exerc 2000; 32: 1480-4

150. Gamelin FX, Berthoin S, Bosquet L. Validity of the polar S810 heart rate monitor to measure RR intervals at rest. Med Sci Sports Exerc 2006; 38: 887

151. Belcastro AN, Dallaire J, McKenzie DC, Schutz R, Smith M, Wenger HA. CASS Overstress Study. Glouchester, Ontario: Canadian Association of Sport Sciences, 1993

152. Gould D, Dieffenbach K. Overtraining, underrecovery, and burnout in sport. In: Kellmann M, editor. Enhancing recovery: Preventing underperformance in athletes; 2002: 25-35

153. Morgan WP, Costill DL, Flynn MG, Raglin JS, O'Connor PJ. Mood disturbance following increased training in swimmers. Med Sci Sports Exerc 1988; 20: 408-14

154. Jones RL. The sports coach as educator: reconceptualising sports coaching. Int J Sport Sci Coach 2006; 1: 405-12

155. Jones RL, Ronglan LT, Potrac P, Cushion C, editors. The Sociology of Sports Coaching. London: Routledge; 2011

156. Kidman L, Thorpe R, Hadfield D. Athlete-centred coaching: Developing inspired and inspiring people. Philadelphia: Innovative Print Communications, 2005

157. Klein G. Naturalistic decision making. Human Factors: Hum Factors 2008; 50: 456

158. Lipshitz R, Klein G, Orasanu J, Salas E. Taking stock of naturalistic decision making. J Behav Dec Making 2001; 14: 331-52

159. Smith TB, Hopkins WG, Lowe TE. Are there useful physiological or psychological markers for monitoring overload training in elite rowers? Int J Sports Physiol Perform 2011; 6: 469-84

160. Lehman M, Foster C. Definition, types, symptoms, underlying mechanisms, and frequency of overtraining and overtraining syndrome. In: Lehman M GH, editor. Overload, performance incompetence, and regression in sport. New York: Kluwer Academic / Plenum Publication; 1999

161. Kentta G, Hassmen P. Overtraining and recovery. A conceptual model. Sports Med 1998; 26: 1-16

162. Kentta G, Hassmen P. Underrecovery and overtraining: A conceptual model. In: Kellmann M, editor. Enhancing recovery: Preventing underperformance in athletes. Champaign, Il: Human Kinetics Publishers; 2002: 57-79

163. Mackinnon LT. Overtraining effects on immunity and performance in athletes. Immunol Cell Biol 2000; 78: 502-9

164. Woodman L. Coaching: A science, an art, an emerging profession. Sport Sci Rev 1993; 2: 1-13

165. Nash C, Collins D. Tacit knowledge in expert coaching: science or art. Quest 2006; 58: 465-77

166. Lyle J. Sports coaching concepts: a framework for coaches' behaviour. London: Routledge, 2002

167. Elstein AS, Shulman, L.S. and Sprafka, S. . Medical problem solving: An analysis of clinical reasoning. Cambridge, Massachusetts: Harvard University Press, 1978

168. Gawande A. Complications: A surgeon's notes on an imperfect science: Picador USA, 2003

169. Isenbcrg D. How senior managers think. Harvard Bus Rev, 62 (6) 1984: 80-90

170. Klein GA. Intuition at work: Why developing your gut instincts will make you better at what you do. New York: Currency/Doubleday, 2003

171. Klein G. Naturalistic decision making by experienced persons. Sources of power: How people make decisions. Cambridge, Massachusetts: The MIT Press, 1998

172. Myers DG. Intuition: Its powers and perils: Yale Univ Pr, 2004

173. Kahneman D. A perspective on judgment and choice: Mapping bounded rationality. Am Psychol 2003; 58: 697

174. Kahneman D, Tversky A. Choices, values, and frames. Am Psychol 1984; 39: 341

175. Cassidy T, Jones RL, Potrac P. Understanding sports coaching: The social, cultural and pedagogical foundations of coaching practice. Oxford: Routledge, 2008

176. Arnold PJ. Meaning in movement, sport, and physical education. London: Heinemann, 1979

177. Wolcott HF. Transforming qualitative data: Description, analysis and interpretation. Thousand Oaks: Sage Publications, 1994

178. Altheide DL. Qualitative media analysis. London: Sage Publications, Inc, 1996

179. Saldana J. The coding manual for qualitative researchers. Los Angeles: Sage, 2009

180. Gibbs GR. Analyzing qualitative data (Book 6 of the SAGE Qualitative Research Kit). London: Sage, 2007

181. Coffey A, Atkinson P. Making sense of qualitative data: Complementary research strategies. London: Sage Publications, Inc, 1996

182. Budgett R, Newsholme E, Lehmann M, Sharp C, Jones D, Jones T, et al. Redefining the overtraining syndrome as the unexplained underperformance syndrome. Br J Sports Med 2000; 34: 67-8

183. Hopkins WG. Quantification of training in competitive sports. Methods and applications. Sports Med 1991; 12: 161-83

184. Kellmann M. Enhancing recovery: preventing underperformance in athletes. Champaign, Il: Human Kinetics Publishers, 2002

185. Polanyi M. Personal knowledge. Towards a post-critical philosophy. Chicago/London: Chicago Press, 1962

186. Brownlee KK, Moore AW, Hackney AC. Relationship between circulating cortisol and testosterone: influence of physical exercise. J Sports Sci Med 2005; 4: 76-83

187. Cox AJ, Pyne DB, Gleson M, Callister R. Resting plasma and salivary IL-6 concentrations are not correlated in distance runners. Eur J Appl Physiol 2008; 103: 477-9

188. Daly W, Seegers CA, Rubin DA, Dobridge JD, Hackney AC. Relationship between stress hormones and testosterone with prolonged endurance exercise. Eur J Appl Physiol 2005; 93: 375-80

189. Hoogeveen AR, Zonderland ML. Relationships between testosterone, cortisol and performance in professional cyclists. Int J Sports Med 1996; 17: 423-8

190. Jurimae J, Jurimae T. Responses of blood hormones to the maximal rowing ergometer test in college rowers. J Sports Med Phys Fitness 2001; 41: 73

191. Jurimae J, Jurimae T, Purge P. Plasma testosterone and cortisol responses to prolonged sculling in male competitive rowers. J Sports Sci 2001; 19: 893-8

192. Lac G, Berthon P. Changes in cortisol and testosterone levels and T/C ratio during an endurance competition and recovery. J Sports Med Phys Fit 2000; 40: 139-44

193. Minetto M, Rainoldi A, Gazzoni M, Terzolo M, Borrione P, Termine A, et al. Differential responses of serum and salivary interleukin-6 to acute strenuous exercise. Eur J Appl Physiol 2005; 93: 679-86

194. Purge P, Jürimäe J, Jürimäe T. Hormonal and psychological adaptation in elite male rowers during prolonged training. J Sports Sci 2006; 24: 1075-82

195. Tremblay MS, Copeland JL, Van Helder W. Effect of training status and exercise mode on endogenous steroid hormones in men. J Appl Physiol 2004; 96: 531-9

196. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000; 30: 1-15

197. Hofman J. Physiological aspects of sport training. Champaign, Il: Human Kinetics, 2002

APPENDIX A

Ethics Approval and Relevant Subject Information for the Study in Chapter 3.

M E M O R A N D U M

To: Will Hopkins
From: **Madeline Banda** Executive Secretary, AUTEC
Date: 20 December 2006
Subject: Ethics Application Number 06/231 **Maximising 'supercompensatory' gains from overreaching in endurance athletes.**

Dear Will

Thank you for providing written evidence as requested. I am pleased to advise that it satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC) at their meeting on 11 December 2006 and that as the Executive Secretary of AUTEC I have approved your ethics application. I have also approved the minor amendments altering the type of study to student research, notifying the protocols used for blood collection and disposal of blood products, and notifying the approval of funding. This delegated approval is made in accordance with section 5.3.2.3 of AUTEC's *Applying for Ethics Approval: Guidelines and Procedures* and is subject to endorsement at AUTEC's meeting on 16 April 2007.

I wish to thank you for the thoroughness of your response and advise that your ethics application is approved for a period of three years until 7 March 2010.

I advise that as part of the ethics approval process, you are required to submit to AUTEC the following:

- A brief annual progress report indicating compliance with the ethical approval given using form EA2, which is available online through *http://www.aut.ac.nz/research/ethics*, including when necessary a request for extension of the approval one month prior to its expiry on 7 March 2010;

- A brief report on the status of the project using form EA3, which is available online through *http://www.aut.ac.nz/research/ethics*. This report is to be submitted either when the approval expires on 7 March 2010 or on completion of the project, whichever comes sooner;

It is also a condition of approval that AUTEC is notified of any adverse events or if the research does not commence and that AUTEC approval is sought for any alteration to the research, including any alteration of or addition to the participant documents involved.

You are reminded that, as applicant, you are responsible for ensuring that any research undertaken under this approval is carried out within the parameters approved for your application. Any change to the research outside the parameters of this approval must be submitted to AUTEC for approval before that change is implemented.

Please note that AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to make the arrangements necessary to obtain this.

To enable us to provide you with efficient service, we ask that you use the application number and study title in all written and verbal correspondence with us. Should you have any further enquiries regarding this matter, you are welcome to contact Charles Grinter, Ethics Coordinator, by email at charles.grinter@aut.ac.nz or by telephone on 921 9999 at extension 8860.

On behalf of the Committee and myself, I wish you success with your research and look forward to reading about it in your reports.

Yours sincerely

Madeline Banda
Executive Secretary
Auckland University of Technology Ethics Committee
    Cc: Tim Lowe, Brett Smith, Rosemary Godbold, Peter Vickers

Participant Information Sheet

AUT UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Date Information Sheet Produced: 31 Jan 2007

Project Title: Maximising 'supercompensatory' gains from overreaching in endurance athletes

An Invitation

You are invited to take part in the research project stated above. This project is a collaborative research project involving; Auckland University of Technology, Rowing New Zealand (RNZ), Sport and Recreation New Zealand and HortResearch. Your participation in this study is voluntary, and you may withdraw from the project at any time, without reason, and without any disadvantage.

What is the purpose of this research?

Elite endurance athletes are required to continually push their bodies to their physiological limit so as to ensure optimal conditioning and therefore run a high risk of overtraining. Overtraining can result in failure to fully recover from exercise leading to a reduction in conditioning and performance which in extreme cases can persist for weeks or months. Another indication of overtraining can be a compromised immune system and increased susceptibility to infections especially upper respiratory infections.

The aim of this study is to measure various markers of both endocrine and immune function to determine if these are useful as markers and predictors of overtraining.

How was I chosen for this invitation?

Effective measures of overtraining in elite athletes cannot be extrapolated very easily from non elite athletes who often don't consistently undertake the workloads necessary to provide the level of stress that can lead to overtraining. As an elite endurance athlete you are ideally suited to participate in this study and potentially benefit from the results.

What will happen in this research?

This project will be timetabled during a normal training month and will necessitate no changes to your training or testing regimes. The additional requirements placed upon you for this study will be up to 10 saliva and 3 capillary blood samples per week. On Monday, Wednesday and Saturday saliva will be collected before and after morning training and before afternoon training with an additional sample being collected before training on Saturday morning. Collection will involve excreting 3 millilitres of saliva into a vial which then be analysed to determine cortisol and testosterone concentrations as well as DHEA concentrations for female Participants. The three capillary samples will be collected prior to training on Monday, Wednesday and Friday and will involve identical earlobe sampling techniques to that used during regular blood lactate testing. These samples will analysed for creatine kinase and all the sampling will continue for one month. Samples from your monthly blood tests will also be analysed to determine immune function. All participants will be asked to record their rate of perceived exertion after each training session to help assess perceptions of exertion so as to examine any relationship between them and the markers of overtraining being examined in this study.

The results of the previous measures will be compared to the comprehensive monitoring of training already in place in RNZ's elite program to help analyse whether any of the markers of training correlate with the various performance monitors. While one could argue that reduction in performance capacity or physiological conditioning is the ideal measure of overtraining there is some evidence that recovery from this state is very difficult and a predictive tool that could allow modification of training before this stage would be useful

What are the discomforts and risks?

The normal training and testing regime undertaken in the RNZ elite program involves a number of discomforts and some potential risks which RNZ seeks to minimise through a comprehensive safety plan and associated monitoring procedures. Saliva sampling is the only measure that has been added to those normally undertaken during training.

Saliva sampling consists of chewing gum for 30 seconds (which improves your ability to get the saliva flowing), while swallowing the saliva and then continuing to chew and excreting 3-4 mls of saliva into a 10ml vial. The vial is sterile and is placed in a clean area until used by you so as to reduce any risks of any cross infection.

Capillary blood will be collected from the ear using the same sampling strategy as that used for blood lacatate testing and the normal safety strategies will be applied. During your normal monthly blood test one more test tube of blood will be collected so that specific immune system markers can be examined however this will not cause any more discomfort than normal for this measure.

How will these discomforts and risks be alleviated?

You are asked to inform the researchers prior to providing any blood capillary or saliva samples or completing the rate of perceived exertion charts if you are uncomfortable with the procedures and they will attempt to find an option you are comfortable with. If during or after providing one of these samples you don't feel comfortable please tell the researchers and they will attempt to find a solution. Participants who are unhappy with the discomforts and risks can pull out of the study at any time without having to provide a reason and with no negative impacts.

What are the benefits?

The aim of this project is to develop non-invasive (saliva) measures of overtraining that will help to maximise performance gains from training and reduce the risks of overtraining.

What compensation is available for injury or negligence?

Compensation is available through the Accident Compensation Corporation within its normal limitations.

How will my privacy be protected?

All participants will be coded with a number that will be used throughout the study. No participant names will be released in any publications arising from this study.

What are the costs of participating in this research?

The sampling will occur during normal warm up and warm down periods associated with training so as to minimise the time associated with this study. Extrapolating the time periods involved in other studies of this nature suggest the sampling may inconvenience you for 30-45 minutes per week.

What opportunity do I have to consider this invitation?

One week

How do I agree to participate in this research?

If you agree to participate in this research then please read and sign the attached consent form.

Will I receive feedback on the results of this research?

Yes. There will be a presentation to all participants at the end of the trial. Anonymity will be strictly adhered to.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Either, Brett Smith, brett@waikato.ac.nz, 021 627863, or Tim Lowe, tlowe@hortresearch.co.nz 07 8584821

Concerns regarding the conduct of the research should be notified to the Executive Secretary, AUTEC, Madeline Banda, madeline.banda@aut.ac.nz , 921 9999 ext 8044.

Whom do I contact for further information about this research?

Either, Brett Smith, brett@waikato.ac.nz, 021 627863, or Tim Lowe, tlowe@hortresearch.co.nz 07 8584821

Approved by the Auckland University of Technology Ethics Committee on 7th March 2007, AUTEC Reference number 06/231

Consent Form

AUT
UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project title: Maximising 'supercompensatory' gains from overreaching in endurance athletes

Project Supervisor: Professor Will Hopkins

Researchers: Brett Smith, Tim Lowe

○ I have read and understood the information provided about this research project in the Information Sheet dated 6 Feb 2007.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ I am not suffering from heart disease, high blood pressure, any respiratory condition (mild asthma excluded), any illness or injury that impairs my physical performance.

○ I agree to provide two 20 ml standard venepuncture blood samples. One at the beginning and end of the trial. I also agree to approximately 12 capillary blood samples from the ear (identical to the standard method currently used for lactate sampling). In addition, I agree to provide approximately 40 saliva samples.

○ I agree to take part in this research.

○ I wish to receive a copy of the report from the research (tick one): Yes ○ No ○

○ I wish all my samples (blood and saliva) returned at the end of this project

(please tick one): Yes ○ No ○

Participant's signature: ....……………………………………

Participant's name: ....…………………………………………

Participants ethnicity : ....……………………………………

Participant's Contact Details (if appropriate):

………………………………………………………………………………..

………...……...

Date: …………………………………………………………………………...

Approved by the Auckland University of Technology Ethics Committee on 5th March 2007 AUTEC Reference number 06/231

Note: The Participant should retain a copy of this form.

APPENDIX B

Salivary Hormone and Blood Markers of Performance in Elite Rowers

Report to New Zealand Rowing Association Incorporated and Sport and Recreation New Zealand (SPARC)

HortResearch Client Report No. 22183
HortResearch Contract No. 21618

Lowe TE, Smith TB, Hopkins WG. June
2007

Lowe TE
The Horticulture and Food Research Institute of New Zealand Ltd
HortResearch Ruakura Private Bag 3123, Waikato Mail Centre Hamilton 3240
NEW ZEALAND
Tel: +64-7-858 4650
Fax: +64-7-858 4700

Smith TB
Department of Sport and Leisure Studies
University of Waikato
Hamilton
NEW ZEALAND

Hopkins WG
Institute of Sport and Recreation Research NZ AUT University
Auckland
NEW ZEALAND

DISCLAIMER

Unless agreed otherwise, HortResearch does not give any prediction, warranty or assurance in relation to the accuracy of or fitness for any particular use or application of, any information or scientific or other result contained in this report. Neither HortResearch nor any of its employees shall be liable for any cost (including legal costs), claim, liability, loss, damage, injury or the like, which may be suffered or incurred as a direct or indirect result of the reliance by any person on any information contained in this report.

LIMITED PROTECTION
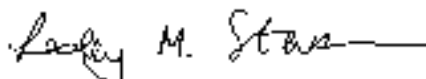
This report may be reproduced in full, but not in part, without prior consent of the author or of the Chief Executive Officer, The Horticulture and Food Research Institute of New Zealand Ltd, Private Bag 92 169, Auckland Mail Centre, Auckland 1142, New Zealand.

This report has been prepared by The Horticulture and Food Research Institute of New Zealand Ltd (HortResearch), which has its Head Office at 120 Mt Albert Rd, Mt Albert, AUCKLAND. This report has been approved by:

Research Scientist   Group Leader, Health and Food

Date: 25 June 2007   Date: 25 June 2007

CONTENTS

EXECUTIVE SUMMARY

Salivary Hormone and Blood Markers of Performance in Elite Rowers

Report to New Zealand Rowing Association Incorporated and Sport and Recreation New Zealand

Lowe TE, Smith TB, Hopkins WG.   June 2007

- The aim of this study was to identify salivary hormonal and blood markers of performance during a period of intense training in a group of elite athletes. Such markers may eventually provide more information than performance monitoring alone for monitoring of over-reaching and prevention of over-training.

- The athletes were 13 female and 10 male elite rowers. The study was conducted from 5 March to 16 April 2007.

- This report details the preliminary findings for the salivary steroid hormones (cortisol and testosterone) for all the subjects plus an androgenic hormone (dehydroepiandrosterone, DHEA) for the female subjects. The report also includes creatine kinase and lactate dehydrogenase, which were measured as markers of muscle inflammation/damage.

- Results from a weekly 30-minute ergometer session indicated performance increases of approximately 1% for the females and 2% for the males. On average, the athletes were therefore not over-reached, but approximately 30% of females and males failed to improve and were therefore candidates for over-reaching.

- An incremental lactate test was also performed at the beginning of the study and at the end of the study following a day and a half of rest. Performance improved dramatically by approximately 8% in females and 9% in males. The rest period for this test probably accounts for the better performance compared with the 30-minute test. Apparently most if not all of the squad had adapted to the intense training.

- There was a small decline (11%) in morning testosterone over the course of the study for the females but the other hormones showed no substantial trends for males or females.

- Post exercise testosterone declined by 28% in females and 13% in males.

- Improvement in the male athletes' performance of the 30-minute test was associated with an increase in morning cortisol.

- For the 30-minute test, increases in performance were associated with increases in cortisol and DHEA for the females, and testosterone for the males.

- Creatine kinase increased in relation to the work load as expected and tended to decrease after the rest day on Sundays. Lactate dehydrogenase had a different profile and appeared to follow an adaptive response to the increased workload. These enzymes will be reviewed in the context of other measures in a future report.

- DHEA responded in some situations with similarities to cortisol release and in others more like testosterone. We conclude that DHEA is not an equivalent measure to testosterone in females although absolute concentrations probably reflect general androgen status.

- For females, testosterone and DHEA concentrations predict performance changes for the 2 and 4 mmol lactate test. That is, the higher the resting value and the smaller the change in concentration over the 30-minute exercise test, the greater the improvement in performance.

- A large number of measures outside the scope of the present report were also collected and there is ongoing analysis with additional reports to be presented as they are completed. These reports will investigate relationships between the measures mentioned above and factors such as heart rates, behavioural variables, cytokines and on water performance. Therefore, conclusions relating to value of these steroidal hormones for measuring over-reaching/over-training in elite rowers can only be made once all these variables have been considered.

For further information, please contact: Dr Tim Lowe

The Horticulture and Food Research Institute of New Zealand Ltd

HortResearch Ruakura Private Bag 3123, Waikato Mail Centre Hamilton 3240

NEW ZEALAND Tel: +64-7-858 4650

Fax: +64-7-858 4700

Email: tlowe@hortresearch.co.nz

INTRODUCTION

Objectives
Monitor changes in salivary cortisol, testosterone, and dehydroepiandrosterone (DHEA) during one month of intensive training by elite rowers, to determine their potential as physiological indicators of over-reaching and possibly over-training

Background
Successful training requires a period of overload to stimulate adaptive responses followed by an appropriate recovery period so that supercompensation can occur. If an imbalance between excessive overload and inadequate recovery occurs, then a performance decrement can occur (Uusitalo et al. 2001, Meeusen et al. 2006).

Over-reaching is the condition where insufficient recovery inhibits the supercompensatory response but the process remains readily reversible. Over-training occurs when the imbalance is sustained to the point where it causes long-term physiological disorders with the concomitant performance decrements. Some coaches believe that over-reaching is an important component of training elite athletes, as the subsequent supercompensatory response can lead to large performance gains. Unfortunately the danger of over-training increases dramatically with this approach.

The issue of over-reaching/over-training is controversial and there is no consensus in the literature on how to measure over-training or if it even exists. This is covered in reviews by Kuipers & Keizer (1988), Budgett (1998), Wilmore & Costill (1999), Hug et al. (2003), Halson & Jeukendrup (2004) and Meeusen et al. (2006). Leading theories that emerge from these reviews, about over-reaching/over-training, centre on altered functioning of the endocrine system, suppression of the immune system, and altered functioning of the autonomic nervous system.

This report focuses on the endocrine responses of athletes during high volumes of intensive aerobic training. What is unique about this study in comparison with other studies is that we have employed the use of extensive salivary sampling to gain a comprehensive profile of the endocrine responses of the athletes during the experimental period. This is generally not feasible using more traditional blood sampling methods. Cortisol, testosterone and DHEA are controlled by the hypothalamus and pituitary glands and changes in these hormones reflect an integrated response to stress and training. There is also growing support for modulation of the immune system in response to over-reaching, and this is reported to increase susceptibility to infection and disease. The most commonly proposed model is that moderate exercise enhances the immune system, but with very high levels of exercise the immune system becomes compromised and an increased incidence of upper respiratory tract infections occurs (Smith 2004).

Scope of this report
The scope of this report is to present the salivary steroid hormone results along with the measures of muscle damage of creatinine kinase, and lactate dehydrogenase. These measures are compared with performance measures determined from ergometer sessions made during the study, and the results from incremental lactate tests.

The complete analysis of all the data collected throughout this study is much broader than this report, and will take some time to complete. The collective aim is for Rowing New Zealand and SPARC is to have reliable measures to evaluate the physiological responses of athletes to over-reaching and potentially over-training. This would potentially allow effective monitoring of endurance athletes to help to optimise their training programmes as well as to enable the exploration of effective interventions to help to promote the same outcome. Many of the variables considered in the reviews of over-reaching and over-training referenced above have been measured in this study. HortResearch also included the additional measurements of selected cytokines, which will be analysed from plasma and saliva samples taken during the study and these results will be made available to all parties named in the contract once they are available. There are also large amounts of heart-rate data, behavioural variables, and on-water performance measures to be analysed. The majority of these data will contribute to the PhD of Brett Smith and those results will be presented in report form to all parties as soon as they are completed.

The measurement of salivary hormones

The measurement of steroid hormones in this study is primarily to determine if the current training programme used by Rowing New Zealand alters neuroendocrine responses to exercise, as determined by the measurement of salivary cortisol, testosterone and DHEA.

Cortisol is released by the adrenal glands, is essential for life and is responsible for about 95% of all glucocorticoid activity in the body. Often cortisol is portrayed purely as a catabolic hormone, which may be viewed as negative for muscle growth. Cortisol is a non- specific stress indicator, meaning that a wide range of factors can stimulate cortisol release.

Testosterone is an androgen and is found in both men and woman. It is an anabolic hormone and promotes muscle growth and secondary sexual development. Female testosterone levels are 5-10% of male testosterone levels.

Dehydroepiandrosterone (DHEA) is also an adrogenic hormone produced in similar amounts by both males and females. Levels tend to be slightly higher in females and more importantly the DHEA that is synthesised in the adrenal gland and the ovaries is the main precursor for testosterone in females.   Technically, DHEA is easier to measure than testosterone in females, and this is one of the main reasons for including its measurement in this study. The objective of measuring DHEA was to determine if it is a more appropriate or equivocal measure of testosterone in females.

The fate of cortisol, testosterone and DHEA is complex and in the blood there is the added complication that these steroids are bound to carrier proteins. Aside from the non-invasive nature of saliva collection compared with blood collection, saliva levels of these hormones are thought to reflect the "free" or "bioactive fraction" of these hormones, thus affording a more sensitive and relevant measure than blood. Steroid hormones are lipophillic in nature and the unbound free fraction can readily passively pass through cell membranes that would otherwise require active transport mechanisms. This means that steroids can be reliably measured in saliva.

The study

Elite rowers were monitored for one month during ultra endurance training in March – April 2007. This allowed monitoring during a scheduled training period of sustained very high work loads. Twenty three athletes, 13 females and 10 males were monitored. During this type of training there are significant individual differences in the ability of the athletes to maintain a consistent performance over the weekly and monthly training periods.   Natural variation in performance has allowed the over-reaching indicators to be evaluated against performance both within and between athletes.

MATERIAL AND METHODS

Experimental design

Elite rowers from the New Zealand national rowing squad were monitored for one month during ultra endurance training in March - April 2007 at Lake Karapiro. This allowed monitoring during a scheduled training period when the intensity of training was such that rowing performance progressively increased for one month through weekly microcycles. During this type of training, there are significant individual differences in the ability of the athletes to maintain a consistent performance over the weekly and monthly training periods. This natural variation in performance allows the over-reaching indicators to be evaluated against performance both within and between athletes.

This report details the saliva hormones cortisol, testosterone and DHEA, blood plasma creatine kinase and lactate dehydrogenase, and performance measures from stepwise lactate testing and a 30-min ergometer testing. The training schedule is shown in Table 1. A total of approximately 44 saliva and 16 blood samples were collected from each individual during the study. The full experimental design includes a number of measures not included in this report and analysis is ongoing for heart rates, behavioural variables, cytokines and on-water performance.

Subjects

Twenty-three elite rowers from the New Zealand national rowing squad took part in the study. All subjects were fully informed of the nature and possible risks of the study before giving written consent. The protocol was approved by the Auckland University of Technology Ethics Committee (Ethics Application Number 06/231). All subjects were informed that they could cease their participation in the trial at any time without giving a reason, with no repercussions. The sex, height, weight, and age of the participants were: 12 females, (178.0 (176-181.5) cm (mean, range), 74.2 (69.0-93.4) kg, 23(19-31) years; and 10 males, (190.5 (181-200) cm, 89.8 (71.8-100.5) kg, 24(21-30) years, respectively.

Saliva sampling

Saliva production was stimulated by giving the subjects sugar-free gum to chew. The gum was chewed for approximately 30 seconds and this initial saliva swallowed, then with continued chewing 3-5 ml of saliva was collected into a labelled 10 polyethylene centrifuge tube. The saliva samples were frozen at -20°C until analysis. On the first Monday and all subsequent Saturdays a sub-sample was stored at -80°C for cytokine analysis.

Blood sampling

In addition to routine monthly venepuncture blood samples, 0.12-0.25 ml of capillary blood was taken from the ear lobe three times a week. Ear lobes were sampled using a standard lancet to prick the ear lobe, followed by collection of the blood into heparinised capillary tubes. The tubes were immediately centrifuged and the plasma separated and stored at 4-8°C, and then analysed within 24 hours for lactate dehydrogenase and creatine kinase activity. On the first Monday and all subsequent Saturdays a sub-sample was stored at -80°C for cytokine analysis.

Performance tests

The discontinuous stepwise submaximal ergometer test was conducted on the Concept IIb rowing ergometer. This test involves 6 minutes of work followed by 1 minute of rest when the blood lactate is sampled; the workload is then increased and the process repeated until the lactate concentration is above 6 mmol/L. The workload for the 2 and 4 mmol/L is then calculated.

A 30-minute rating restricted rowing ergometer session was conducted every Friday.   The results for all these sessions were normally published to enable comparisons between all crews and rowers to ensure a highly competitive environment, and thereby supporting maximal performance in all these events. Saliva samples were taken both before and after these sessions.

Table 1. Summary of the exercise sessions, and the saliva and blood sampling times for rowers during the study. N.B. There was no sampling on Week 5, but the exercise sessions were identical to those in Weeks 2-4.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Week 0	5 Mar 07	6 Mar 07	7 Mar 07	8 Mar 07	9 Mar 07	10 Mar 07
Week 1 7:30am	12 Mar 07 Saliva sample 1	13 Mar 07 Saliva sample	14 Mar 07 Saliva sample 4	15 Mar 07 Saliva sample	16 Mar 07 Saliva sample	17 Mar 07 Saliva sample 10
7:30am	Capillary blood 1		Capillary blood 4			Capillary blood 10
	Off water Ln test Saliva sample 2	U2	AT Saliva sample 5	U2	U2	AT Saliva sample 11
10:00am 10:00am		Venous blood			Saliva sample 30 min erg test @ 18	
	Resistance		Resistance			AR
12:00pm 12:30pm	U2	U2		U2	Saliva sample AR	
Week 2 7:30am	19 Mar 07 Saliva sample 1	20 Mar 07 Saliva sample	21 Mar 07 Saliva sample 4	22 Mar 07 Saliva sample	23 Mar 07 Saliva sample	24 Mar 07 Saliva sample 10
7:30am	Capillary blood 1	U2	Capillary blood 4 AT	U2	U2	Capillary blood 10 AT
	Saliva sample 2		Saliva sample 5			Saliva sample 11
10:00am 10:00am 12:00pm	Resistance		Resistance		Saliva sample 30 min erg test @ 18	AR
12:30pm	U2	U2		U2	Saliva sample AR	
Week 3 7:30am	26 Mar 07 Saliva sample 1	27 Mar 07 Saliva sample	28 Mar 07 Saliva sample 4	29 Mar 07 Saliva sample	30 Mar 07 Saliva sample	31 Mar 07 Saliva sample 10
7:30am	Capillary blood 1	U2	Capillary blood 4 AT	U2	U2	Capillary blood 10 AT
	Saliva sample 2		Saliva sample 5			Saliva sample 11
10:00am 10:00am 12:00pm	Resistance		Resistance		Saliva sample 30 min erg test @ 18	AR
12:30pm	U2	U2		U2	Saliva sample AR	
Week 4 7:30am	2 Apr 07 Saliva sample 1	3 Apr 07 Saliva sample	4 Apr 07 Saliva sample 4	5 Apr 07 Saliva sample	6 Apr 07 Saliva sample	7 Apr 07 Saliva sample 10
7:30am	Capillary blood 1	U2	Capillary blood 4 AT	U2	U2	Capillary blood 10 AT
	Saliva sample 2		Saliva sample 5			Saliva sample 11
10:00am 10:00am 12:00pm	Resistance		Resistance		Saliva sample 30 min erg test @ 18	AR
12:30pm	U2	U2		U2	Saliva sample AR	
Week 5	9 Apr 07	10 Apr 07	11 Apr 07 Venous blood	12 Apr 07	13 Apr 07	14 Apr 07
Week 6	16 Apr 07 Off water Ln test	17 Apr 07	18 Apr 07	19 Apr 07	20 Apr 07 2k timetrial	21 Apr 07 2km race

Saliva analysis.

Saliva samples were analysed in triplicate for cortisol, testosterone and DHEA using radioimmunoassay (RIA). The methods were modified from those described by Granger et al. (1999 a, b) and Morelius et al. (2004). Briefly, standards from serum diagnostic kits (Diagnostic Systems Laboratories, USA) were diluted in phosphate buffer saline (Sigma P4417) to cover the ranges of 0-500, 0-51.2, and 0-12 ng ml-1, for cortisol, testosterone, and DHEA respectively. Saliva sample sizes of 50, 100-200, and 75 µ l were used for cortisol, testosterone, and DHEA respectively. The antibodies were diluted in a phosphate buffered saline solution containing 0.05% bovine serum albumin. Kit standards were diluted so that approximately 50% binding was achieved compared with the total counts.   The intra- and inter-assay coefficients of variation were 10.7% and 4.6% for cortisol, 4.0% and 6.4% for testosterone, and 10.2% and 7.7% for DHEA.

Creatine kinase analysis.

Plasma samples were analysed for creatine kinase using the International Federation of Clinical Chemistry IFCC primary reference procedure as described by Schumann et al. 2006. The assay was adapted for use on plate reader by linear downscaling of the assay. The plate reader was a FLUOstar manufacturer? with thermostatic control. Assays were conducted at 30°C and standard correction factors were applied to represent the results at a 37°C equivalent rate.

Lactate dehydrogenase.

Lactate dehydrogenase was determined by the method described by Howell et al. (1979). The method was adapted for use on a FLUOstar plate reader. The decrease in NAD was followed using fluorescence rather than changes in absorbance. This provided increased sensitivity over absorbance in the plate reader format.

Statistical Analyses.

All dependent variables were log transformed for analysis of effects as percentage or factor differences and changes. Mixed modelling was used to fit linear trends to the various measures with either assay day or a hormone concentration as the predictor.   With the latter, the magnitude of its effect was estimated and presented as the change in performance accompanying two standard deviations of difference (between-subjects) or change (within-subjects) of the predictor. Where possible, a random effect was included in the model to estimate individual differences in effects as standard deviations. Outcomes were expressed as 90% confidence limits, and inferences about the true (large-sample) values of effects were based on interpreting the magnitude of observed value, the lower confidence limit and the upper confidence limit in relation to the smallest worthwhile effect. For performance in the 30-min tests, the smallest worthwhile effect was assumed to be ~1%. Because of uncertainty in the relationship between on-water competitive performance and changes in 2 mM and 4 mM lactate power, we assumed a smallest effect of ~2% for these measures. Smallest effects in hormone concentrations were based on the Cohen approach: approximately 0.20 of the between-subject standard deviation (see   http://newstats.org/effectmag.html). All analyses were performed using SAS (Statistical Analysis System, Version 9.1, SAS Institute, Cary, NC).

RESULTS AND DISCUSSION

Performance

To evaluate performance two standardised tests are reported here; a weekly 30-minute maximal intensity ergometer session in which the stroke rate was restricted to 18, and an incremental discontinuous lactate test.

The ergometer session was conducted every Friday at midday during the first four weeks of the six-week training block. This training session was familiar to all the rowers, as it has been traditionally used as a maximal intensity aerobic session. The changes in the mean power for these Friday sessions in relation to the mean power in Week 1 are presented in Figure 1.
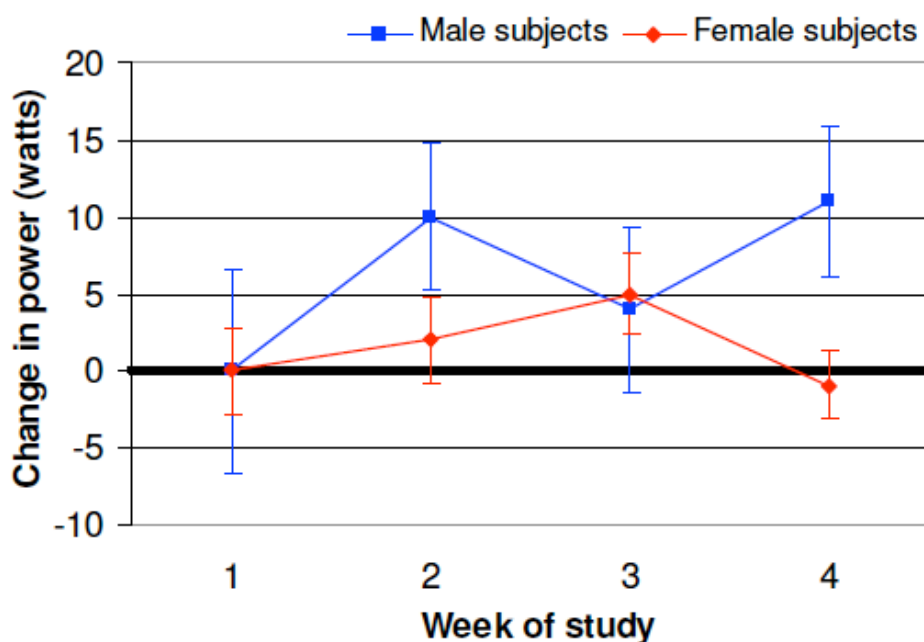
Figure 1. Change in mean power for the 30-minute ergometer test for elite rowers for each week relative to the mean power for Week 1. Mean ± standard deviation.

Motivation to perform well in the ergometer session was high as all the rowers were preparing for both the World Championships and Olympic qualifying in August. To ensure consistency in the pre-test conditioning and workload, the weekly lead up to the ergometer session was almost identical throughout the study. During each session the ergometers were programmed to record the average rating every five minutes. From these data, it was found that only two rowers did not strictly follow the recommended stroke rate, but as they both maintained consistent stroke rates of 19-20 throughout the study, their data have been included in the analysis. All these steps were taken to help to ensure consistency prior to the 30-minute ergometer test, to increase its reliability as a monitor of aerobic conditioning or fatigue.

The placement of the ergometer session on Friday with almost 80 percent of the weekly programme completed made this session particularly sensitive to fatigue and over-reaching. We suggest that a change in performance of approximately 1% is meaningful for this particular test. The percentage changes for the test are listed in Table 2 and these indicate an increase of 1% in the females and 2% in the males. In any of Weeks 2, 3 and 4, approximately 30% of the female and male rowers were unable to maintain the workloads achieved in the previous week.

The incremental discontinuous lactate tests were conducted on the first day of the study (Monday, Week 1) and on the last day of the training block (Monday, Week 6). During Week 5 monitoring had stopped but the training programme for that week was the same structure as the previous four weeks of the programme, although the workload was programmed to increase for that period. All subjects had at least one and a half days rest prior to completing both these tests. Blood glucose concentration was also determined from the final blood lactate sample and no subject recorded any abnormal values (blood glucose concentrations outside the normal 4-8 mmol/L range). The performance changes for the females and males were approximately 8-9% (Table 2).

Table 2. Changes in performance for female rowing subjects were calculated from the mean changes in the weekly 30-minute ergometer tests over the 4 weeks of the project and also the power associated with the 2 mmol and 4 mmol blood lactate concentrations determined from a discontinuous stepwise lactate test conducted at the beginning and end of the project.		
Performance measure	Mean change in	
	Power (watts) (%) Mean ± 90% CL	
	Female subjects	Male subjects
30-minute ergometer test	1.0; ±1.2 positive	1.9; ±2.2 positive
2 mmol lactate threshold test	8.1; ±2.7 positive	9.2; ±1.9 positive
4 mmol lactate threshold test	7.8; ±2.4 positive	8.4; ±3.0 positive

Over-reaching is an imbalance between workload and recovery, and performance tests are one of the criteria used for assessment of over-reaching (Halson & Jeukendrup 2004). The weekly ergometer sessions are a reliable measure of actual performance and the results suggest no meaningful improvement in the females and only a small improvement if any in the males. These results are consistent with over-reaching being achieved in some of the females and males, but in general they improved.

In contrast to the ergometer results, improvements in the incremental lactate test were very large, although these are normal for this stage of the programme. An explanation for this result is that while some of the rowers may have been over-reached during the week when they performed their ergometer session, a degree of recovery was occurring during the day and a half of rest they had at the end of each week (there is further support for this in the hormonal data that follow).

Morning hormonal results
Baseline trends during the study for salivary cortisol, testosterone and DHEA are derived from morning saliva samples typically taken between 0645-0715 h from Monday to Saturday each week. All the hormones measured in this study have daily circadian rhythms; that is, they show regular increases and decreases, throughout the day and night on an approximate 24-hour cycle (Lejeune-Lenain et al. 1987). For this reason, it is important to note the time of sampling and as much as possible to ensure that the sample timing is consistent. Sampling time was shifted by one hour from the second week of the study when New Zealand daylight saving stopped. At the present time we have not attempted to correct for this in the results presented.

Morning Cortisol
Mean morning cortisol concentrations for females and males are presented in Figure 2.
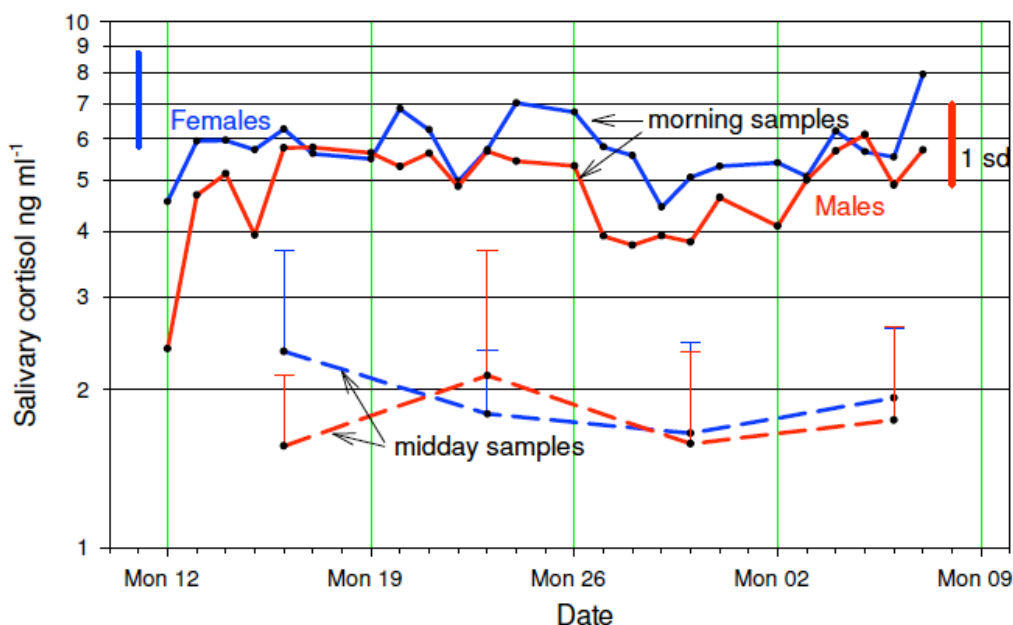
Figure 2. Salivary cortisol concentrations for female (n=12) and male (n=10) rowers during a four-week training period. The top solid lines indicate mean morning cortisol concentrations from Monday to Saturday for each week. The values shown are reverse transformed least squares means from log transformed data. The standard deviations for between-rower means are shown on the left (females) and right (males) hand sides of the figure. Saliva samples were also taken at midday on every Friday. These concentrations were lower than the morning values and are shown by dashed lines. The values are means and standard deviations.

Statistical analysis (see Appendices) suggests that there was no consistent linear trend found for mean morning cortisol over the period of the study.

The morning cortisol concentrations were much higher than we initially expected. This result is readily explained by a characteristic early morning rise in cortisol associated with waking (Pruessner et al. 1997). We had anticipated the early morning rise would have been completed approximately one hour after waking, and that the rowers would have been awake for at least this time prior to sampling. Given that the morning cortisol samples appear to have been affected by the early morning rise, we have included resting values from weekly Friday midday samples. These samples provide a second reference point in looking for a baseline shift and for this reason they have been included in Figure 1. There was no obvious mean baseline change for the midday cortisol levels.

Morning Testosterone

Over the month of monitoring, linear trend analysis of the mean morning testosterone for males showed no clear trend, but the females showed an 11% decrease (90% CL ± 9.2%) (Figure 3). The females also tended to display their highest weekly testosterones on Monday and Tuesday following the scheduled rest day on Sunday. There is, as expected, an approximate four-fold difference in concentrations of testosterone between males and females. The mean testosterone (and standard deviation) for females and males were 23 (±84%), 85 (±30%) pg ml-1 respectively.

Figure 3. Morning concentration of salivary testosterone for female (n=12) and male (n=10) rowers during a four-week training period. The values shown are reverse transformed least squares means from log transformed data. The standard deviations for between-rower means are shown on the left (females) and right (males) hand sides of the figure.

Morning Dehydroepiandrosterone
DHEA was only measured in the females, and the results are presented in Figure 4.

Over the month of monitoring, linear trend analysis of the mean morning DHEA indicated no clear trend. Mean Monday morning concentrations were always higher than the previous Saturday, which is similar to the pattern observed for testosterone, suggesting a rise in androgenic hormones following a day of rest.
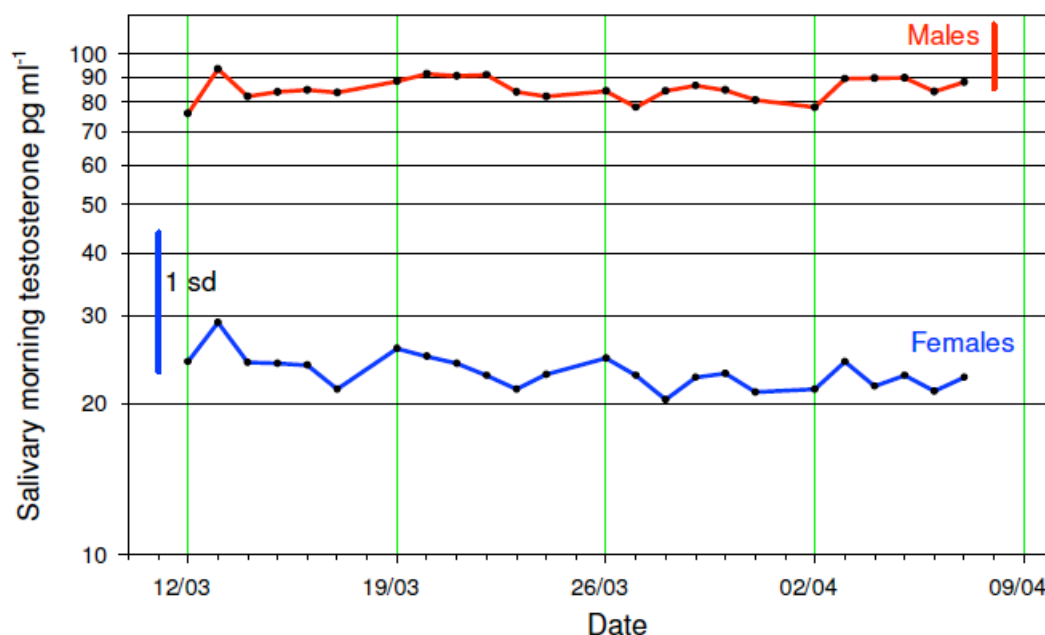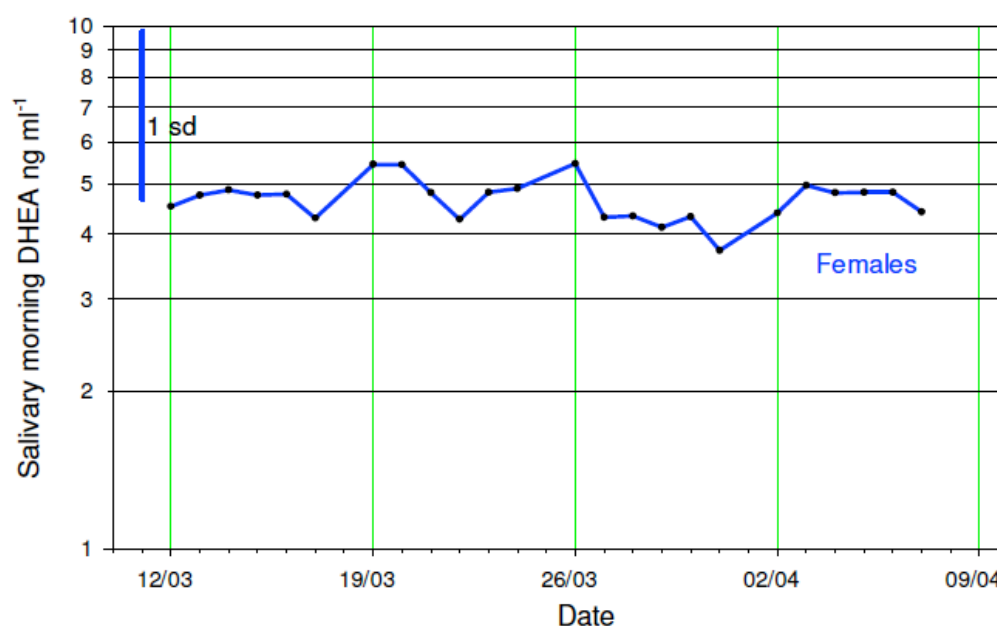

Figure 4. Morning concentration of salivary Dehydroepiandrosterone (DHEA) for female (n=12) rowers during a four-week training period. The values shown are reverse transformed least squares means from log transformed data. The standard deviations for between-rower means are shown on the left hand side of the figure.

Trends in Morning Hormones

For athletes involved in endurance training one would typically see relatively unchanged morning hormone levels unless an imbalance between training load and recovery causes excessive fatigue resulting in a suppressed endocrine response (Veroon et al. 1991). This excessive fatigue or over-reaching is typically reversible within days. However, if appropriate modifications to the athletes training are not instituted over-training may occur (Barron et al.1985).

The results of this study typically show no substantial change in mean morning hormone levels over the study period, with the exception to this pattern being the steady decrease in mean testosterone levels in the females. However, it is important to note that females did tend to increase testosterone and DHEA levels after a 1.5 day rest at the end of each training week. This result for the females suggests that as a group they were becoming over-reached during the week with partial recovery on the rest day, and as the workload increased each week, the suppression of testosterone was progressive throughout the whole study period.

To date, there are only a few studies which have examined similar parameters in elite rowers and the results have been equivocal. Steinacker et al. (1993) found that in junior rowers morning testosterone decreased over time with large volumes of training and then increased with lower volume, high intensity training. However, further research with very high volumes of training has showed no changes in either resting testosterone or cortisol (Steinacker et al. 2000). Veroon (1991) monitored the Dutch national team for nine months and found that resting testosterone and cortisol were generally unchanged. Maestu et al. (2005) found that three weeks of heavy training in junior rowers induced reductions in resting free testosterone, while cortisol showed no change.

Research in other endurance sports has also shown unclear patterns in the responses of testosterone and cortisol. Reported changes in resting cortisol with endurance training are not consistent between studies, with reports of levels going both up and down. For example, male cyclists during the Tour of Spain showed a decrease in cortisol concentrations after only one week of the tour and they continued to decrease over the remainder of the tour (Lucia et al. 2001). The exercise consisted of 21 consecutive daily stages with only one rest day. It is very difficult to compare research in different endurance sports directly because of different training regimes and different amounts of tissue trauma associated with different types of exercise.

DHEA has been studied in female athletes primarily to determine if as a measure of the adrenal androgen response to exercise it is comparable to testosterone (Filaire & Lac 2000). In the female rowers in this study DHEA did not follow the progressive decline seen for testosterone, yet levels did increase following the rest day. DHEA sulphate (sulphated form of DHEA) has been measured in male rowers doing endurance training and showed no change (Steinacker et al. 2000).

Creatine Kinase and Lactate Dehydrogenase

The mean morning creatine kinase and lactate dehydrogenase concentrations for females and males are presented in Figures 5 and 6. These measures were made in plasma, and because they needed to be determined in blood creatine kinase and lactate dehydrogenase, were only measured three times per week. Both these measures are related to muscle damage and changes in concentrations are frequently observed with changes in work load. In this study Wednesday and Saturday creatine kinase samples were consistently the highest weekly values and the Monday morning values were always lower that the previous Saturday's values. This was expected, as Sunday was a rest day. The mean creatine kinase values in the females were approximately half those of the males (Figure 5).

The detection of creatine kinase and the eventual peak values associated with an exercise event or muscle damage frequently take 24 hours to occur. In a study on junior rowers, it was determined that on any given day, the creatine kinase concentrations were dependent on the training load of the previous two days (Yuan et al. 2003).

Lactate dehydrogenase had a different response profile from that of creatine kinase (Figure 6). It was highest during the first 2 weeks of the study, and then decreased to levels just lower than those at the beginning of the study and then appeared to level off. Our interest in these measures was related to the hormones, but also to the cytokine measures that have yet to

analysed. The association with these enzymes with cytokines will be discussed in later reports.
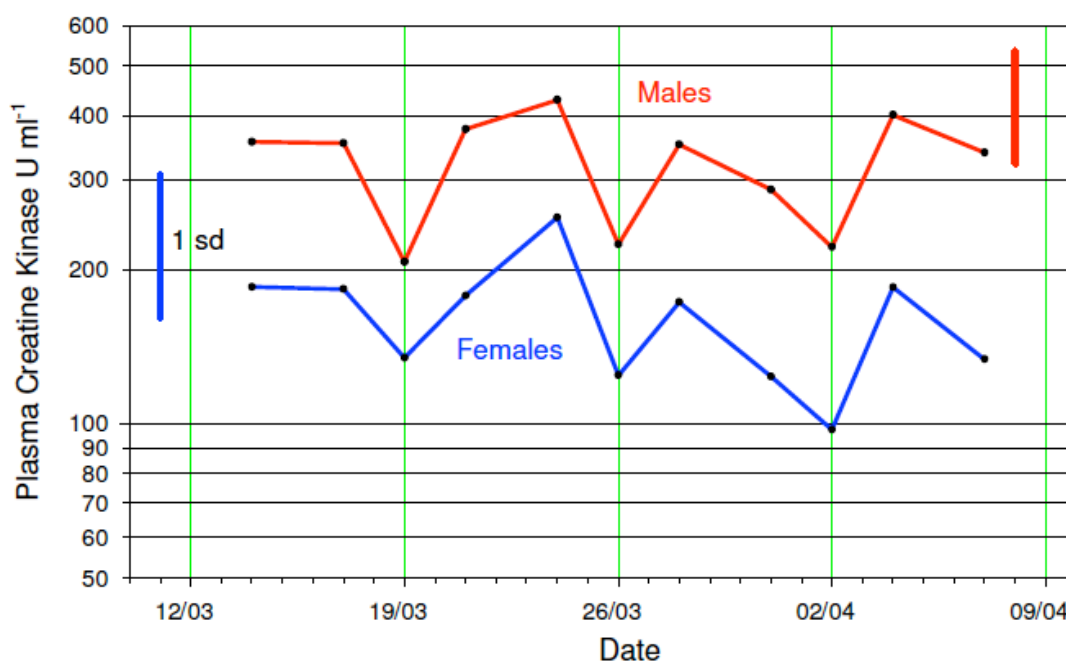
Figure 5. Morning concentration of plasma creatine kinase for female (n=12) and male (n=10) rowers during a four-week training period. The values shown are reverse transformed least squares means from log transformed data. The standard deviations for between-rower means are shown on the left (females) and right (males) hand sides of the figure.
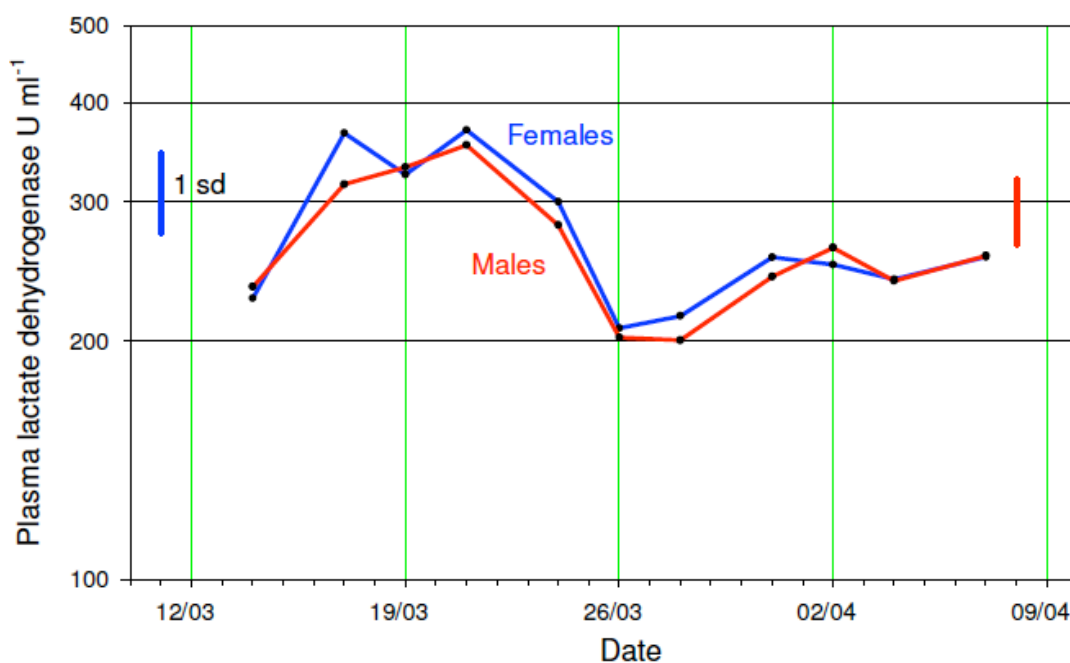
Figure 6. Morning concentration of plasma lactate dehydrogenase for female (n=12) and male (n=10) rowers during a four-week training period.   The values shown are reverse transformed least squares means from log transformed data. The standard deviations for between-rower means are shown on the left (females) and right (males) hand sides of the figure.

Performance versus morning hormones

In this section we have undertaken two separate analyses. The first involved the performance changes over the study for both the 30-min ergometer and incremental lactate tests and related these tests both to mean morning hormone concentrations (Table 3A) and the immediate pre

30-min test hormone concentrations (Table 3B).   The second type of analysis was only possible for the 30-min test, and was the effect of within-athlete changes in salivary hormone concentration on changes in performance power (%). These analyses were made for both morning (Table 4A) and the immediate pre-test hormone concentrations (Table 4B).

For the 30-min test, there was no association between an individual's absolute morning or pre 30-min test hormone concentrations and changes to aerobic performance. The exception to this result was that for the 30-min test, elevated cortisol in the females was positively associated with performance (Table 3A).   However, we suspect that measure was made during the early morning cortisol rise. The corresponding value taken pre the 30-min test was unclear (Table 3B).

Table 3A. Effect of morning salivary hormone concentration averaged over the study on changes in performance power (%) in the 30-min and incremental lactate tests. Changes in performance shown are for 2 standard deviations of difference between athletes in hormone concentration.				
		Effect of hormone concentration on change (%) in performance (mean; ±90% confidence limits)		
Sex	Test	Cortisol	Testosterone	DHEA
Females	30 min	2.4; ±2.4 positive	0.9; ±2.5 unclear	0.8; ±2.5 unclear
	2 mmol	-0.4; ±5.0 unclear	3.8; ±4.5 positive	4.1; ±4.5 positive
	4 mmol	-1.1; ±5.0 unclear	4.1; ±4.4 positive	4.6; ±4.4 positive
Males	30 min	-1.4; ±4.3 unclear	3.0; ±4.6 unclear	-
	2 mmol	-2.2; ±3.5 unclear	1.5; ±3.8 unclear	-
	4 mmol	1.6; ±5.6 unclear	1.2; ±5.6 unclear	-

Table 3B. Effect of pre-test salivary hormone concentration averaged over the study on changes in performance power (%) in the 30-min and incremental lactate tests. Changes in performance shown are for 2 standard deviations of difference between athletes in hormone concentration.				
		Effect of hormone concentration on change (%) in performance (mean; ±90% confidence limits)		
Sex	Test	Cortisol	Testosterone	DHEA
Females	30 min	-0.1; ±2.5 unclear	1.0; ±1.3 unclear	0.2; ±2.5 unclear
	2 mmol	-2.0; ±4.8 unclear	4.8; ±4.0 positive	4.9; ±3.9 Positive
	4 mmol	-2.0; ±4.9 unclear	5.1; ±3.9 positive	5.7; ±3.5 positive
Males	30 min	-0.7; ±4.6 unclear	1.7; ±2.2 unclear	-
	2 mmol	-1.6; ±3.7 unclear	1.5; ±3.8 unclear	-
	4 mmol	1.6; ±5.6 unclear	-1.5; ±5.5 unclear	-

For the incremental lactate tests, the females with higher levels of testosterone or DHEA tended to have better improvements in the lactate test, while for the males no clear

associations were evident. These relationship patterns were the same in females and males in the analysis using both morning and pre 30-min hormone concentrations. However, the statistical relationship for testosterone and DHEA in the females was stronger using the pre 30-min test hormone concentration.

We are unaware of any research examining the association between absolute levels of hormones and improvement in endurance performance in elite athletes. In this study, the standard deviations for both testosterone and DHEA were 84 % and 101%, respectively, in the females, which is very large compared with the 30% value for testosterone in the males. This suggests that the males were hormonally a much more uniform group of athletes. For the males, testosterone levels were similar to levels HortResearch has measured in other athletes, although we do not have similar data for females.

The effect of within-athlete changes in morning and pre 30-min test salivary hormones during the study, and the change in ergometer performance are shown in Tables 4A and B. In general, the association was either trivial or unclear, with the exception being the males, where increases in morning cortisol had a positive association with performance (Table 4A). However, when this male cortisol result was compared with the pre 30-min test effect, is the result was unclear (Table 4B). The only other effect was a small negative association between the pre 30-min test testosterone and performance.

Table 4A. Effect of within-athlete changes in morning salivary hormone concentration on changes in performance power (%) in the 30-min test. Changes in performance shown are for 2 standard deviations of change in an athlete's hormone concentration.			
	Effect of change in hormone concentration on change (%) in performance (mean; ±90% confidence limits)		
	Cortisol	Testosterone	DHEA
Females	-1.0±1.2 trivial	-0.1±0.7 trivial	-0.8±0.8 trivial
Males	2.6±2.1 positive	-1.7±2.4 unclear	

Table 4B. Effect of within-athlete changes in pre 30-min test salivary hormone concentration on changes in performance power (%) in the 30-min test. Changes in performance shown are for 2 standard deviations of change in an athlete's hormone concentration.			
	Effect of change in hormone concentration on change (%) in performance (mean; ±90% confidence limits)		
	Cortisol	Testosterone	DHEA
Females	-1.1±1.0 trivial	0.2±0.9 trivial	0.3±1.0 trivial
Males	0.1±3.7 unclear	-1.8±1.8 negative	

We suggest caution in interpreting the morning cortisol result for the males, as these data may have been adversely affected by the morning cortisol rise. There is a suggestion in the literature that sampling the early morning rise in cortisol may have clinical significance for assessing dynamic activity in the Hypothalamus Pituitary Adrenal HPA axis (Wirtz 2007). Thus for over-reached athletes with a blunted hormonal response, one would expect a lower early morning cortisol response. In that context, the result presented here suggests that for males with a higher early morning rise in cortisol, performance was improved. However, these are very speculative suggestions as sampling in this study was not designed to test this, and we have no data to verify the time from waking to sampling, nor was this result observed in the females. To study this would require multiple samples from the time of waking for 1-2 hours.

Generally these results suggest that during the study, neither morning nor pre 30-min test concentrations of hormone had much effect on performance in the 30-minute ergometer test.

However, it must be remembered that it is possible that some of the rowers were in an over-reached state when performing this test.

Changes in hormones in response to exercise

To test for hormonal responses to exercise, we have specifically focused on the results from saliva samples taken before and after the 30-minute ergometer sessions. Thus the session doubled as an exercise challenge test to monitor performance during the study, and enabled the measurement of hormones in response to exercise during the study. A further advantage was the midday timing of the test, which moved it well away from the early morning cortisol peak. The mean pre- and post-exercise values for cortisol, testosterone and DHEA, in females are shown in Figures 7-9.

Figure 7. Mean concentrations of cortisol pre- and post- the weekly 30-minute ergometer test for female rowers. The clear box at the bottom of the column represents the mean pre exercise concentration of cortisol. The blue box indicates the mean increase in cortisol ± standard deviation.

Figure 8. Mean concentrations of testosterone pre and post the weekly 30-minute ergometer test for female rowers. The clear box at the bottom of the column represents the mean pre exercise concentration of testosterone. The blue box indicates the mean increase in testosterone ± standard deviation.
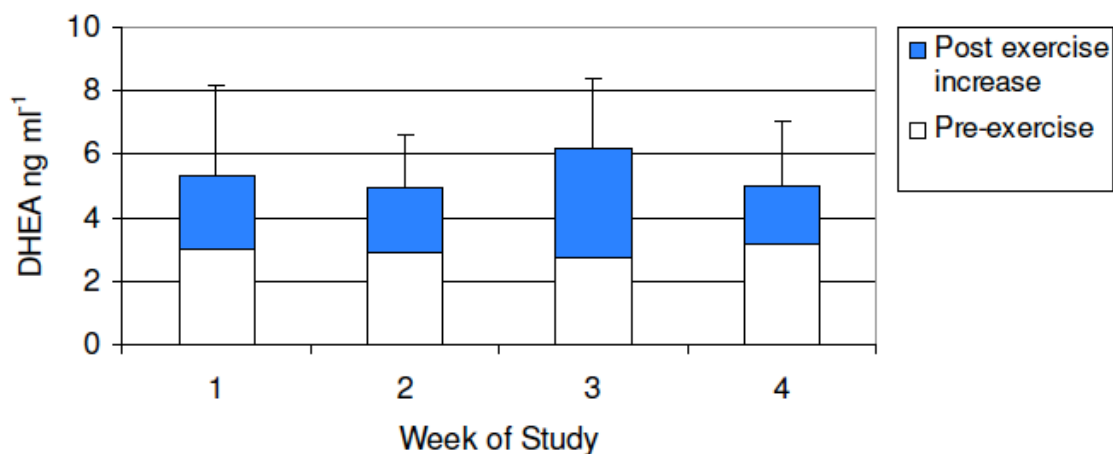
Figure 9. Mean concentrations of dehydroepiandrosterone (DHEA) and post- the weekly 30 minute ergometer test for female rowers. The clear box at the bottom of the column represents the mean pre exercise concentration of DHEA. The blue box indicates the mean increase in DHEA ± standard deviation.

In the females, testosterone showed a linear trend of -10.4% ± 10.2 90% CL for the pre-exercise levels of testosterone, while the trend for the post-exercise levels was -28% ±12.2 90% CL. Therefore the only meaningful outcomes for these measures were that pre test testosterone and the testosterone response to the exercise challenge decreased throughout the study.

The mean pre- and post-exercise values for cortisol and testosterone in males, for the ergometer sessions are shown in Figures 10 & 11.
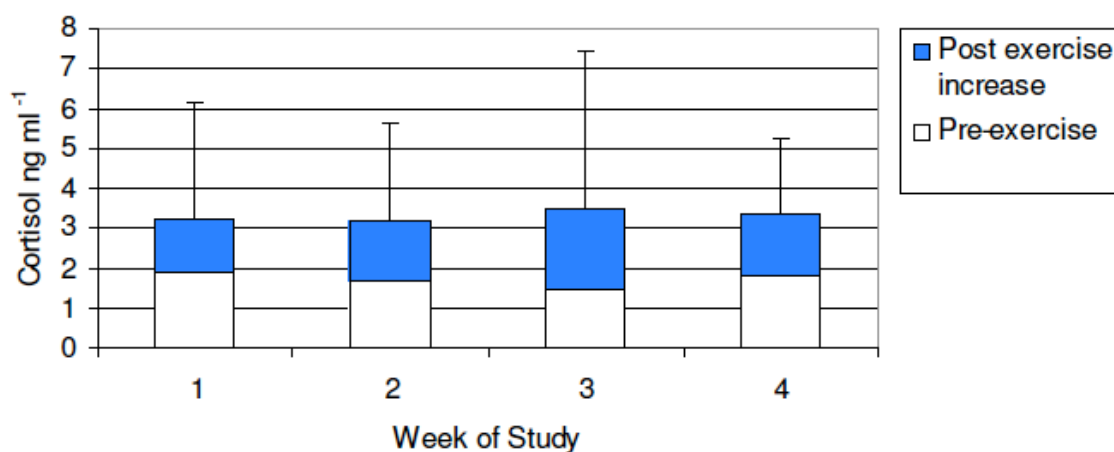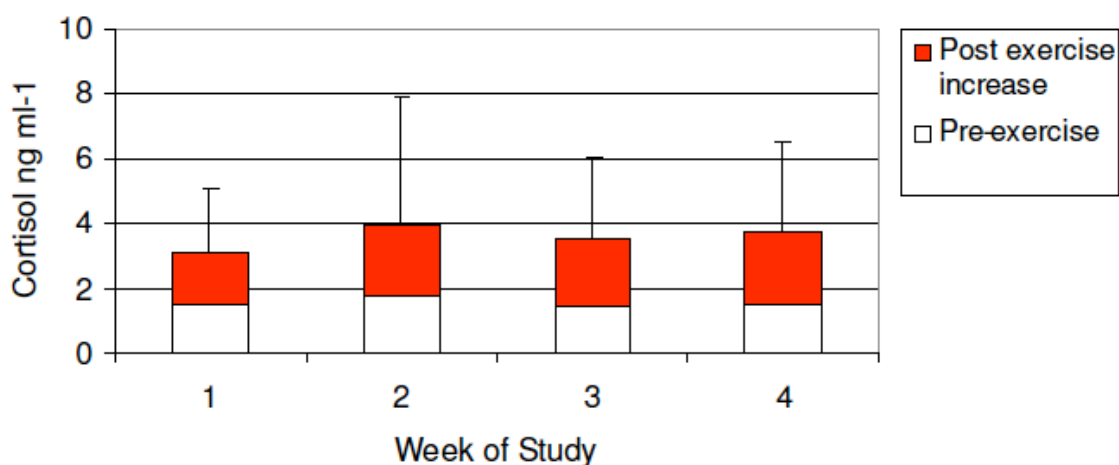

Figure 10. Mean concentrations of cortisol pre and post the weekly 30-minute ergometer test for male rowers. The clear box at the bottom of the column represents the mean pre exercise concentration of cortisol. The red box indicates the mean increase in cortisol ± standard deviation.
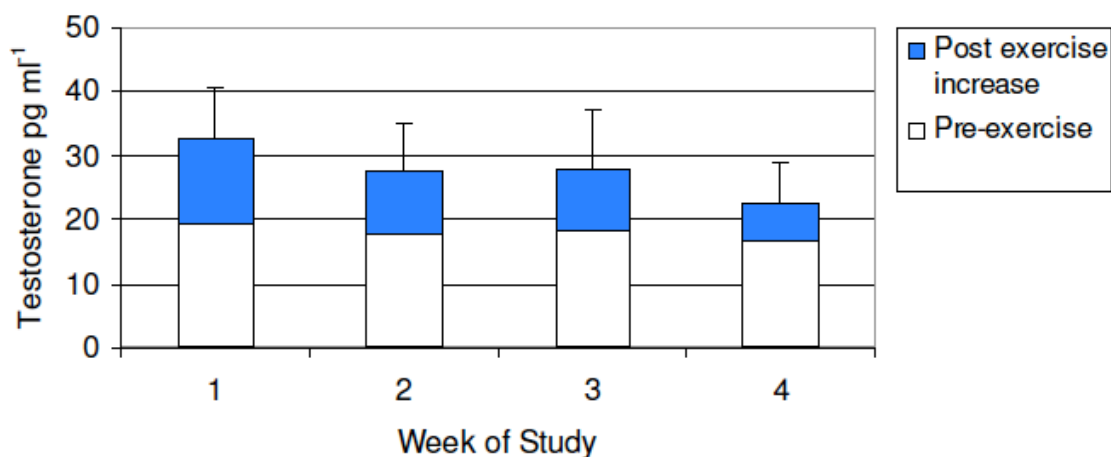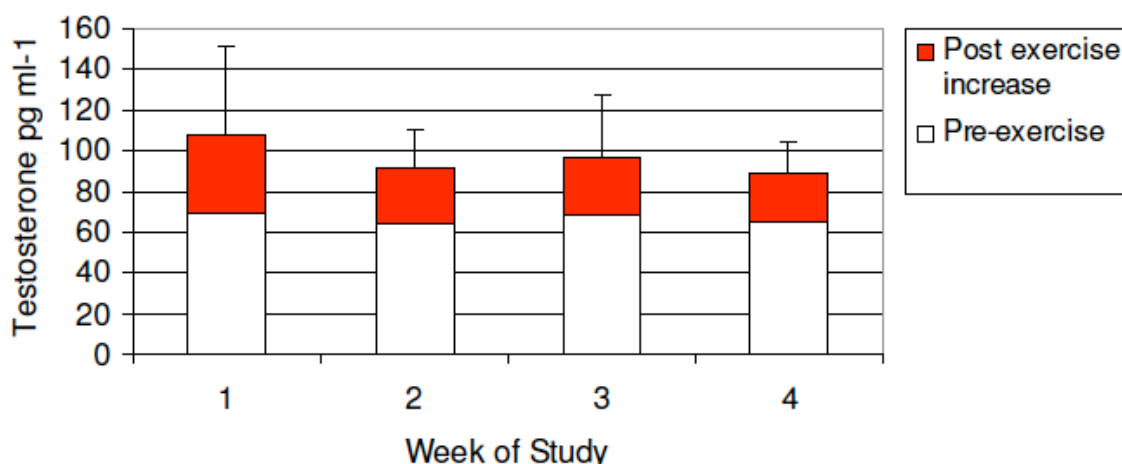
Figure 11. Mean concentrations of testosterone pre and post the weekly 30-minute ergometer test for male rowers. The clear box at the bottom of the column represents the mean pre exercise concentration of testosterone. The red box indicates the mean increase in testosterone ± standard deviation.

There were no changes in pre-exercise mean cortisol concentrations or in the post-exercise concentrations over the course of the study. Pre-exercise mean concentrations of testosterone did not change during the study. However, post-exercise testosterone concentration did decrease by -13.3% ±17 90% CL. A number of the observed hormonal responses to exercise were expected, as they play an important role in any subsequent physiological adaptations to exercise.

The sensitivity of endocrine responses to exercise increases their importance as possible markers of performance and it was our expectation that these measures would be more effective in monitoring over-reaching than changes in morning concentrations. While there is considerable support from many studies that endocrine dysfunction is a characteristic of over-reaching/over-training (Meeusen et al. 2004), the majority of studies only measure resting hormone concentrations. (Veroorn et al. 1991; Steinacker et al. 2000; Mäestu et al. 2003; Mujika et al. 1996). Where endocrine responses to exercise have been measured, they are claimed to be superior for indicating over-reaching/over-training (Uusitalo et al. 1998).

Tables 5 and 6 analyse the relationship between performance and hormonal responses to exercise. In Table 5 the change (%) in performance is compared with the mean pre-post change hormone concentrations for individuals.

Table 5. Effect of change in salivary hormone concentration in the 30-min test averaged over the study on changes in performance power (%) in the 30- min and incremental lactate tests. Changes in performance shown are for 2 standard deviations of difference between athletes in hormone concentration.

Sex	Test	Effect of hormone concentration on change (%) in performance (mean; ±90% confidence limits)		
		Cortisol	Testosterone	DHEA
Females	30 min	-1.3; ±2.5 unclear	-1.7; ±2.4 unclear	-1.3; ±2.4 unclear
	2 mmol	0.6; ±6.1 unclear	-3.6; ±4.3 negative	-6.0; ±3.8 negative
	4 mmol	2.6; ±6.4 unclear	-2.9; ±4.6 negative	-5.6; ±4.1 negative
Males	30 min	-2.5; ±4.4 unclear	-1.9; ±4.3 unclear	-
	2 mmol	-1.0; ±3.8 unclear	-3.3; ±3.1 negative	-
	4 mmol	0.4; ±5.7 unclear	1.9; ±5.6 unclear	-

For both the females and males, the relationship between the hormonal response to the 30-minute ergometer test and performance was unclear. However, there was a negative association between the testosterone and DHEA responses and performance in the 2 and 4 mmol lactate tests. As the testosterone response for males showed opposite trends between the 2 and 4 mmol lactate tests, we conclude that the trends were unclear. In this particular analysis, we compared hormonal responses made when a portion of the rowers may have been over-reached (ergometer test) with a lactate test undertaken after 1.5 days of rest.

If these results are compared with the results in Tables 3A and B, which compare changes in performance with concentrations of mean morning hormones, then for the females, high morning or pre 30-min test concentrations of both testosterone and DHEA are associated with the greater increases in performance. However, if the percentage of pre-post exercise change in testosterone and DHEA are considered, then individuals who had small changes had greater performance increases. If the corresponding cortisol and testosterone results for the males are compared, no clear trends emerge.

As the relationship between hormonal response and performance has probably never been analysed in this way before, we are cautious in our interpretation. There was only a clear negative response to exercise for the testosterone and DHEA levels in the females, whom we believe were more likely to be over-reached than the males. One suggestion is that the females who had the most suppressed hormonal responses had the biggest aerobic performance gains. From this, we conclude that the suppression of hormonal responses is important for optimising performance gains. Regardless of the potential mechanism, this is a new finding, and is especially interesting when applied to DHEA.

These results show that the best female performances were associated with their biggest (%) increases in cortisol and DHEA, but that for testosterone the relationship was unclear. For the males, the relationship between testosterone and performance is positive, while the relationship for cortisol was unclear.

If these results were compared with the results in Tables 4A and B, then the morning and pre 30 min test hormone concentrations in the females had no obvious response pattern with performance. However, the situation for pre-post exercise changes in cortisol and DHEA were much clearer and the best performances during the ergometer session were associated with the biggest increases in these hormones. The comparison with pre-post exercise changes suggests larger increases in testosterone are associated the best performances of individuals for the males.

Table 6. Effect of within-athlete changes in the change in salivary hormone concentration in the 30-min test on changes in performance power (%) in the 30-min test. Changes in performance shown are for 2 standard deviations of change in an athlete's hormone concentration.

	Effect of change in hormone concentration on change (%) in performance (mean; ±90% confidence limits)		
	Cortisol	Testosterone	DHEA
Females	1.3; ±1.0 positive	0.4; ±1.1 unclear	1.5; ±1.0 positive
Males	0.0; ±3.6 unclear	2.7; ±3.3 positive	

An aim of this study was to question the significance of DHEA as an equivalent measure to testosterone in females. The release of testosterone and DHEA in females in response to exercise and over-reaching is not well understood, and there are very few studies reported in the literature. The results from this study show that DHEA responds in some situations with similarities to cortisol release and in others more like testosterone. To discuss the potential mechanisms behind DHEA, cortisol, and testosterone release fully is beyond the scope of this report, but a useful study on the potential mechanisms is reported by Fearon et al. (1998). For the present study, the observed data show that morning concentrations of testosterone and DHEA tend to correlate. That is, if females have high levels of DHEA then levels of

testosterone tend to be high, not surprisingly, as DHEA is a precursor for testosterone. However, while testosterone concentrations decreased during the study, DHEA concentrations did not. Furthermore, while testosterone responses to exercise decreased during the study, DHEA responses did not. DHEA tended to behave more like cortisol in the pre-post exercise situation. We conclude that DHEA is not an equivalent measure to testosterone, although absolute concentrations probably reflect general androgen status. More research is required to gain a better understanding of DHEA in female athletes.

The analysis of overall trends in morning concentrations of hormone and trends in the response to exercise are similar to those reported in the literature regarding endurance training. There are no known studies that report the results we have presented in Tables 3-6, where we have tried to link hormonal responses to individuals and performance. We believe that the findings reported in this study, such as decreases in testosterone and decreased responsiveness to exercise, are in agreement with current theories about endocrine function and over-reaching. However, it is difficult to assess the performance changes that are associated with hormonal responses to exercise at the level of the individual. This aspect will be further assessed when the full study is complete.

SUMMARY

Results from the weekly 30-minute ergometer session indicated performance increases of approximately 1% for the females and 2% for the males. This suggests over- reaching in the females, and that the males were possibly a little less over-reached than the females. The test was performed with no tapering period prior and after 4.5 days of cumulative exercise load.

Performance in the incremental lactate test improved by approximately 8% and 9%, in females and males, respectively. This test was performed at the beginning of the study and at the end. The final test was performed with only a short taper period of 1.5 days.

For the females, mean morning concentrations of cortisol and DHEA remained unchanged during the study, whereas testosterone decreased by 11%. For the males, both cortisol and testosterone remained unchanged.

Pre-post exercise changes in testosterone proved to be the most sensitive hormonal indicator of the increased workload as the study progressed. The females showed decreases in both pre-exercise testosterone concentrations (10%) and post-exercise concentrations (28%), whereas males showed only a post-exercise testosterone decrease (13%).

Creatine kinase increased in relation to the work load as expected and tended to decrease after the rest day on Sundays. Lactate dehydrogenase had a different profile and appeared to follow an adaptive response to the increased workload. These enzymes will be reviewed in the context of other measures in a future report.

Improvements in performance were associated with higher morning concentrations of testosterone and DHEA in the females. The association of performance with testosterone was unclear in the males.

The results from this study show that DHEA responds in some situations with similarities to cortisol release and in others more like testosterone. We conclude that DHEA is not an equivalent measure to testosterone in females, although absolute concentrations probably reflect general androgen status.

For females, testosterone and DHEA concentrations predict performance changes for the 2 and 4 mmol lactate test. That is, the higher the resting value and the smaller the change in concentration over the 30-minute exercise test, the greater the improvement in performance.

REFERENCES

Barron JL, Noakes TD, Levy W, Smith C, Millar RP 1985. Hypothalamic dysfunction in overtrained athletes. Journal of Clinical Endocrinology and Metabolism 60:803-806.

Bird SP, Tarpenning KM, Marino FE 2006. Independent and combined effects of liquid carbohydrate/essential amino acid ingestion on hormonal and muscular adaptations following resistance training in untrained men. Eur J Appl Physiol. 97: 225-238.

Budgett R 1998. Fatigue and underperformance in athletes: the overtraining syndrome. British Journal of Sports Medicine 32:107-110.

Crewther B, Lowe T, Weatherby R 2007. Salivary hormones and anaerobic performance in healthy males. Accepted abstract for the Australian College of Sports Medicine and Science. Annual conference, October 2007.

Fearon U, Clarke D, Mckenna TJ, Cunningham SK 1998. Intra-adrenal factors are not involved in the differential control of cortisol and adrenal androgens in human adrenals. European Journal of Endocrinology 138:567-573.

Filaire E, Lac G 2000. Dehydroepiandrosterone (DHEA) rather than testosterone shows saliva androgen responses to exercise in elite female handball players. International Journal of Sports Medicine 21:17-20.

Granger DA, Schwartz EB, Booth A, Arentz M 1999. Salivary testosterone determination in studies of child health and development. Hormones and Behavior 35(1), 18-27.

Granger DA, Schwartz EB, Booth A, Curran M, Zakaria D 1999. Assessing dehydroepiandrosterone in saliva: a simple radioimmunoassay for use in studies of children, adolescents and adults. Psychoneuroendocrinology 24:567-579.

Halson SL & Jeukendrup AE 2004. Does Overtraining Exist? Sports Medicine 34(14):967-981.

Howell BF, McCune S, Schaffer R 1979. Lactate-to-pyruvate or pyruvate to lactate assay for lactate dehydrogenase: A re-examination. Clinical Chemistry 25(2): 269-272.

Hug M, Mullis PE, Vogt M, Ventura N, Hoppeler H 2003. Training modalities: overreaching and overtraining in athletes, including a study of the role of hormones. Best Practice and Research Clinical Endocrinology and Metabolism Vol 17(2) 191-209.

Kuipers H & Keizer HA 1988. Overtraining in elite athletes. Sports Medicine 6:79-92. Lejeune-Lenain C, Van Cauter E, Decir D, Beyloos M, Franckson JRM 1987. Control of circadian and episodic variations of adrenal androgens secretion in man. Journal of Endocrinology Investigation 10:267-276.

Lucia A, Díaz Hoyos J, Fernández Villa G, Bandrés F, Chicharro JL 2001. Hormone levels of world class cyclists during the tour of Spain stage race. British Journal of Sports Medicine 53: 424-430.

Mäestu J, Jürimäe J, Jürimäe T 2003. Hormonal reactions during heavy training stress and following tapering in highly trained male rowers. Horm. Metab Res 35:109-113.

Meeusen R, Piacentini MF, Busschaert B, BuyseL, De Schutter G, Stary-Gundersen J 2004.

Hormonal responses in athletes: the use of a two bout exercise protocol to detect subtle differences in (over)training status. European Journal of Applied Physiology 91:140-146.

Meeusen R, Duclos M, Gleeson M, Rietjens G, Steinacker J, Urhausen A 2006 Prevention, diagnosis and treatment of the Overtraining Syndrome - ECSS Position Statement 'Task Force' European Journal of Sports Science 6 (1): 1-14.

Morelius E, Nelson N, Theodorsson E 2004. Salivary cortisol and administration of concentrated oral glucose in newborn infants: improved detection limit and smaller sample volume without glucose interference. Scand J Clin Lab Invest. 64: 113-118.

Mujika I, Chatard J, Padilla S, Guezennec CY, Geyssant A 1996. Hormonal responses to training and its tapering off in competitive swimmers: relationships with performance. European Journal of Applied Physiology 74:361-366.

Pruessner JC, Wolf OT, Hellhammer DH, Buske-Kirschbaum A, von Auer K, Jobst S, Kaspers F, Kirschbaum C 1997. Free cortisol levels after awakening: A reliable biological marker for the assessment of adrenocortical activity. Life Sciences 61(26):2539-2549.

Schumann G, Aoki R, Ferrero CA, Ehlers G, Ferard G, Gella FJ, Jorgensen PJ, Kanno T, Kessner A, Klauke R, Kytzia HJ, Lessinger JM, Miller WG, Nagel R, Pauwels J,

Schimmel H, Siekmann L, Weidemann G, Yoshida K, Ceriotti F 2006. IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 degrees C. Clin Chem Lab Med 44(9):1146-55.

Smith LL 2004. Tissue trauma; The underlying cause of overtraining syndrome? Journal of Strength and Conditioning 18(1): 185-193.

Steinacker JM, Lormes W, Kellmann M, Liu Y, Reibnecker S, Opitz-Gress A, Baller B, Gunther K, Petersen KG, Kallus KW, Lehmann M, Altenburg D 2000. Training of junior rowers before world championships. Effects on performance, mood state and selected hormonal and metabolic responses. The Journal of Sports Medicine and Physical Fitness 40(4): 327-335.

Steinacker JM, Laske R, Hetzel WD, Lormes W, Liu Y, Stauch M 1993. Metabolic and hormonal reactions during training in junior oarsmen. International Journal of Sports Medicine 14:S24-28 Suppl. 1.

Uusitalo ALT, Huttunen P, Hanin Y, Uusitalo AJ, Rusko HK 1998. Hormonal responses to endurance training and over-training in Female Athletes. Clinical Journal of Sport Medicine 8:178-186.

Veroon C, Quist AM, Vermulst LJM, Erich WBM, de Vries WR, Thijssen JHH 1991. The behaviour of the plasma free testosterone/cortisol ratio during a season of elite rowing training. Internation Journal of Sports Medicine 12:257-263.

Williams Textbook of Endocrinology 2003. Saunders Pennsylvania, PR Larsen, HM Kronenberg, S Melmed, KS Polonsky Edts.

Wilmore JH, Costill DL 1999. Physiology of Sport and Exercise (2nd edt), Human Kinetics Champaign: IL, United States.

Wirtz PH, von Känel R, Emini L, Ruedisueli K, Groessbauer S, Maercker A. Ehlert U 2007.

Evidence for altered hypothalamus-pituitary-adrenal axis functioning in systemic hypertension: Blunted cortisol response to awakening and lower negative feedback sensitivity. Psychoneuroendocrinology 32:430-436.

Yuan Y, Kwong AWK, Kaptein WA, Fong C, Tse M, Glatz JFC, Chan C, Renneberg R 2003.

The response of fatty acid-binding protein and creatine kinase to acute and chronic exercise in junior rowers. Research Quarterly for Exercise and Sport 74(3)277-283.

ACKNOWLEDGEMENTS

APPENDIX C

Psychometric Recall Inventories for the Study in Chapter 3.

Name:_____   **Date**_____

PANAS SCALE

Below there is a list of twenty words that represent different moods people can experience. Indicate on the scale to what extent you have experienced these moods during the previous week.

	Not at all	A Little	Moderately	Quite a bit	Very much
Hostile	1	2	3	4	5
Distressed	1	2	3	4	5
Scared	1	2	3	4	5
Active	1	2	3	4	5
Determined	1	2	3	4	5
Irritable	1	2	3	4	5
Guilty	1	2	3	4	5
Upset	1	2	3	4	5
Nervous	1	2	3	4	5
Ashamed	1	2	3	4	5
Proud	1	2	3	4	5
Inspired	1	2	3	4	5
Attentive	1	2	3	4	5
Enthusiastic	1	2	3	4	5
Strong	1	2	3	4	5
Jittery	1	2	3	4	5
Interested	1	2	3	4	5
Excited	1	2	3	4	5
Alert	1	2	3	4	5
Afraid	1	2	3	4	5

	Sun	Mon	Tue	Wed	Thur	Fri	Sat
Time you got up?							
About what time did you fall asleep?							
How many times did you wake during the night?							
How many hours sleep do you think you had?							
What type of exercise (physical activity undertaken to improve or maintain fitness or strength)							
What was the total amount of time that you exercised for (write time in minute for each day)							
Fatigue levels in morning (immediately prior to breakfast)							
Fatigue at end of day (immediately prior to bed)							

Fatigue scale	Not at all	A Little	Moderately	Quite a bit	Very much
	1	2	3	4	5

APPENDIX D

Ethics Approval and Relevant Subject Information for Study in Chapter 4.

AUT UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

MEMORANDUM

Auckland University of Technology Ethics Committee (AUTEC)

To: Will Hopkins
From: **Madeline Banda** Executive Secretary, AUTEC
Date: 31 October 2007
Subject: Ethics Application Number 06/231 **Maximising 'supercompensatory' gains from overreaching in endurance athletes.**

Dear Will

I am pleased to advise that the Chair of the Auckland University of Technology Ethics Committee (AUTEC) has approved minor amendments to your ethics application, allowing a second stage that repeats the first stage without the hormonal measures and with a diary and use of heart rate variability measures already collected as part of the athletes' training routines. This delegated approval is made in accordance with section 5.3.2 of AUTEC's *Applying for Ethics Approval: Guidelines and Procedures* and is subject to endorsement at AUTEC's meeting on 12 November 2007.

I remind you that as part of the ethics approval process, you are required to submit to AUTEC the following:

- A brief annual progress report indicating compliance with the ethical approval given using form EA2, which is available online through *http://www.aut.ac.nz/about/ethics*, including when necessary a request for extension of the approval one month prior to its expiry on 7 March 2010;

- A brief report on the status of the project using form EA3, which is available online through *http://www.aut.ac.nz/about/ethics*. This report is to be submitted either when the approval expires on 7 March 2010 or on completion of the project, whichever comes sooner;

It is also a condition of approval that AUTEC is notified of any adverse events or if the research does not commence and that AUTEC approval is sought for any alteration to the research, including any alteration of or addition to the participant documents involved.

You are also reminded that, as applicant, you are responsible for ensuring that any research undertaken under this approval is carried out within the parameters approved for your application. Any change to the research outside the parameters of this approval must be submitted to AUTEC for approval before that change is implemented.

Please note that AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to make the arrangements necessary to obtain this.

To enable us to provide you with efficient service, we ask that you use the application number and study title in all written and verbal correspondence with us. Should you have any further enquiries regarding this matter, you are welcome to contact Charles Grinter, Ethics Coordinator, by email at charles.grinter@aut.ac.nz or by telephone on 921 9999 at extension 8860.

On behalf of the Committee and myself, I wish you success with your research and look forward to reading about it in your reports.

Yours sincerely

Madeline Banda

Participant Information Sheet

AUT
UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Date Information Sheet Produced: 25th October 2007

Project Title: Maximising 'supercompensatory' gains from overreaching in endurance athletes

An Invitation

You are invited to take part in the research project stated above. This project is part of Brett Smith's PhD and involves collaboration between Auckland University of Technology, Rowing New Zealand (RNZ), Sport and Recreation New Zealand and HortResearch. Your participation in this study is voluntary, and you may withdraw from the project at any time, without reason, and without any disadvantage.

What is the purpose of this research?

Elite endurance athletes are required to continually push their bodies to their physiological limit so as to ensure optimal conditioning and therefore run a high risk of overtraining. Overtraining can result in failure to fully recover from exercise leading to a reduction in conditioning and performance which in extreme cases can persist for weeks or months.

The aim of this study is to measure heart rate variability as an indicator of autonomic function and psychological stress to determine if these are useful as markers and predictors of overtraining.

How was I chosen for this invitation?

Effective measures of overtraining in elite athletes cannot be extrapolated very easily from non elite athletes who often don't consistently undertake the workloads necessary to provide the level of stress that can lead to overtraining. As an elite endurance athlete you are ideally suited to participate in this study and potentially benefit from the results.

What will happen in this research?

This project will be timetabled during a normal training month and will necessitate no changes to your training or testing regimes. All members of the New Zealand Rowing team currently wear heart rate monitors which are regularly downloaded so the only difference will be that after downloading this data it will be further analysed to assess patterns of heart rate variability. Participants will also be asked to complete charts that record their rate of perceived exertion, mood state and sleep patterns at the end of each day to help assess psychological stress so as to examine any relationship between them and the markers of overtraining being examined in this study.

The results of the previous measures will be compared to the comprehensive monitoring of training already in place in RNZ's elite program to help analyse whether any of the markers of training correlate with the various performance monitors. While one could argue that reduction in performance capacity or physiological conditioning is the ideal measure of overtraining there is some evidence that recovery from this state is very difficult and a predictive tool that could allow modification of training before this stage would be useful

What are the discomforts and risks?

The normal training and testing regime undertaken in the RNZ elite program involves a number of discomforts and some potential risks which RNZ seeks to minimise through a comprehensive safety plan and associated monitoring procedures.

How will these discomforts and risks be alleviated?

You are asked to inform the researchers prior to completing any of the measures if you are uncomfortable with the procedures and they will attempt to find an option you are comfortable

with. If during or after providing one of these samples you don't feel comfortable please tell the researchers and they will attempt to find a solution. Participants who are unhappy with the discomforts and risks can pull out of the study at any time without having to provide a reason and with no negative impacts.

What are the benefits?

The aim of this project is to develop non-invasive measures of overtraining that will help to maximise performance gains from training and reduce the risks of overtraining.

What compensation is available for injury or negligence?

Compensation is available through the Accident Compensation Corporation within its normal limitations.

How will my privacy be protected?

All participants will be coded with a number that will be used throughout the study. No participant names will be released in any publications arising from this study. The data from this study will be stored indefinitely in a secure location by Rowing New Zealand, HORT research and Brett Smith.

What are the costs of participating in this research?

The rate of perceived exertion, mood state and sleep pattern charts will be completed at the beginning and end of each day and will take approximately 2-3 minutes to complete. Extrapolating the time periods involved in other studies of this nature suggest the sampling may inconvenience you for 30-45 minutes per week.

What opportunity do I have to consider this invitation?

One week

How do I agree to participate in this research?

If you agree to participate in this research then please read and sign the attached consent form.

Will I receive feedback on the results of this research?

Yes. There will be a presentation to all participants at the end of the trial. Anonymity will be strictly adhered to.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Either, Brett Smith, brett@waikato.ac.nz, mob 021 627863 ph 07 8484500, or Will Hopkins will.hopkins@aut.ac.nz ph 09 921 9999

Concerns regarding the conduct of the research should be notified to the Executive Secretary, AUTEC, Madeline Banda, madeline.banda@aut.ac.nz , 921 9999 ext 8044.

Whom do I contact for further information about this research?

Either, Brett Smith, brett@waikato.ac.nz, 021 627863, or Will Hopkins will.hopkins@aut.ac.nz ph 09 921 9999

Consent Form

For use when laboratory or field testing is involved.

Project title: Maximising 'supercompensatory' gains from overreaching in endurance athletes

Project Supervisor: Professor Will Hopkins

Researchers: Brett Smith, Tim Lowe

○ I have read and understood the information provided about this research project in the Information Sheet dated 31 October 2007.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ I am not suffering from heart disease, high blood pressure, any respiratory condition (mild asthma excluded), any illness or injury that impairs my physical performance.

○ I agree to take part in this research.

○ I wish to receive a copy of the report from the research (tick one): Yes ○ No ○

Participant's
signature:...………………………………………………

Participant's name:...................................………………………………………………

Participants ethnicity :............................………………………………………………

Participant's Contact Details (if appropriate):

…………………………………………………………………………..

…………………………………………………………………………..

…………………………………………………………………………..

…………………………………………………………………………..

Date:

***Approved by the Auckland University of Technology Ethics Committee on 5th March 2007** AUTEC Reference number 06/231*

Note: The Participant should retain a copy of this form.

APPENDIX E

Psychometric Diary for the Study in Chapter 4.

All the daily recordings were identical so only one day was included in these appendices.

ROWING NEW ZEALAND

2007-2008

DAILY DIARY

MOOD STATE

QUALITY OF SLEEP

RATING OF FATIGUE

PARTICIPANT CODE_____

MONDAY _____(enter date)

Below there is a list of twenty words that represent different moods people can experience. Indicate on the scale to what extent you have experienced these moods during the previous day.

	Not at all	A Little	Moderately	Quite a bit	Very much
Hostile	1	2	3	4	5
Distressed	1	2	3	4	5
Scared	1	2	3	4	5
Active	1	2	3	4	5
Determined	1	2	3	4	5
Irritable	1	2	3	4	5
Guilty	1	2	3	4	5
Upset	1	2	3	4	5
Nervous	1	2	3	4	5
Ashamed	1	2	3	4	5
Proud	1	2	3	4	5
Inspired	1	2	3	4	5
Attentive	1	2	3	4	5
Enthusiastic	1	2	3	4	5
Strong	1	2	3	4	5
Jittery	1	2	3	4	5
Interested	1	2	3	4	5
Excited	1	2	3	4	5
Alert	1	2	3	4	5
Afraid	1	2	3	4	5

Time you got up this morning?	
About what time did you fall asleep last night?	
How many times did you wake during the night?	
How many hours sleep do you think you had?	
Fatigue levels this morning (immediately after waking)	
Fatigue levels last night (immediately prior to bed)	

	Not at all	**A Little**	**Moderately**	**Quite a bit**	**Very much**
Fatigue scale	1	2	3	4	5

RATINGS OF PERCIEVED FATIGUE (RPE) FROM TRAINING SESSIONS

Session details (type of training session)	Start time	Duration (minute)	RPE	Comments	RPE scale
					0 - Nothing at all
					0.5 Extremely light
					1 Very light
					2 Light
					3 Moderate
					4 Somewhat hard
					5 Hard
					6 –
					7 Very hard
					8 –
					9 Very very hard
					10 Extremely hard (almost max)
					* - Maximal

General Comments

APPENDIX F

Salivary Steroid Hormones as Markers of Performance and Overreaching in Elite Rowers

[1,3]T Brett Smith, [2,3]Tim E Lowe, [3]Will G Hopkins

[1]Sport and Leisure Studies, University of Waikato, Hamilton, NZ;

[2]Health and Food Group, HortResearch, Hamilton, NZ;

[3]Institute of Sport and Recreation Research, AUT University, Auckland, NZ

This is the abstract for the podium presentation I made at the New Zealand Sports Medicine and Science Conference in Dunedin 2008.

ABSTRACT

The use of cortisol and testosterone as indicators of overreaching and overtraining in athletes has attracted considerable interest, as changes in these hormones can reflect dysfunction in the role of the hypothalamic-pituitary axis in exercise. Resting concentrations of these hormones have been monitored in many training studies, but there has been little investigation of changes in the hormonal responses to exercise.

Twenty-three members of the New Zealand elite rowing team (13 women, 10 men) were monitored over one month of intensive aerobic training to examine the relationship of salivary cortisol and testosterone to changes in performance. Performance tests, conducted on a rowing ergometer, consisted of a weekly 30-min time trial for assessment of mean power and an incremental step test at the beginning and end of the study to determine power at the lactate threshold. Salivary samples were collected on six mornings each week and immediately before and after the time trials. Saliva samples were analysed for cortisol and testosterone using radioimmunoassay.

Over the course of the study there was a mean increase in power in the 30-min test in females (1.0%, 90% confidence limits ±1.2%) and males (1.9%, ±2.2%). Approximately 30% of the rowers failed to improve and were therefore candidates for over-reaching. Lactate-threshold power showed much larger improvements that were usual for this phase of training (females 8.0%, ±2.5%; males 8.7%, ±2.5%). A possible explanation for the differences between the two tests is that the 30 minute test was conducted near the end of the training week when the athletes were more likely to be over-reached while the lactate threshold test was conducted after a day and a half rest. Morning testosterone declined by 11% (±9%) in females during the study, whereas there was little change for the males. There was little change in morning cortisol and in the cortisol response to exercise as the study progressed, but testosterone each side of the time trial was a sensitive indicator of the increased workload: females showed clear decreases before and after the test (10%, ±10%; 28%, ±12%), and males showed a clear decrease after the test (13% ±17%).

In conclusion, a decrease in the testosterone response to exercise was a useful indicator for over-reaching associated with increased training load. However, there was no such evidence for cortisol. The best way to measure suppression of testosterone is to conduct a standardised exercise test at same time of the day, perhaps on a weekly basis. This approach would allow the monitoring of hormone baseline and responsiveness to exercise. The effectiveness of testosterone as a predictor of performance at an individual level remains to be determined.

APPENDIX G

Effects of Overload Training on Physiology, Psychology and Performance of Elite Rowers

Brett Smith[1,3], Will G Hopkins[3], Tim E Lowe[2,3]

[1]Sport and Leisure Studies, University of Waikato, Hamilton, NZ;

[2]School of Applied Science, Bay of Plenty Polytechnic, Tauranga, NZ;

[3]Institute of Sport and Recreation Research, AUT University, Auckland, NZ

This is the abstract for the podium presentation I made at the New Zealand Sports Medicine and Science Conference in Rotorua 2009. With this presentation I won the emerging researchers award.

ABSTRACT

Background: Establishing physiological and psychological predictors of performance change would help to optimise individual training in competitive athletes.

Aim: To monitor immune and other potential predictors during four weeks of intense overload training in 12 female and 10 male elite rowers preparing for a world championship.

Method: Change in performance was determined from mean power in a 30-min rowing ergometer test and from lactate-threshold power (evoking 4 mmol/L) in a discontinuous incremental rowing ergometer test. The 30-min test provided changes in performance on a weekly basis throughout the overload, while the incremental test was undertaken at the beginning of the study and following a one-week taper after the overload. Linear modelling was used to estimate the effect on performance of two standard deviations of difference or change in a predictor, and inferences about effects on performance were based on the location of 90% confidence limits in relation to magnitude thresholds of 1.0%, 3.0%, 5.3% and 8.3% for small, moderate, large and very large.

Results: There were small improvements in 30-min mean power over the period of overload (percent change in performance: females 1.0 ± 2.2, males 1.8 ± 4.6; mean ± SD), while lactate-threshold power showed large improvements (females 8.2 ± 4.3; males 8.4 ± 4.5). Many changes considered to predict over-training (e.g., worsening mood, decreasing morning testosterone, increased inflammatory response) actually had small to large positive associations with performance.

Conclusion: Markers of physiological and psychological stress during periods of overload training in highly trained athletes may be useful predictors of enhanced performance rather than over-training.

APPENDIX H

Are there Useful Physiological or Psychological Markers for Monitoring Overload Training in Elite Athletes?

Author Block: Tiaki B. Smith[1], Will G. Hopkins, FACSM[2], Timothy E. Lowe[3].
[1]University of Waikato, Hamilton, New Zealand. [2]AUT University, Auckland, New Zealand. [3]BOP Polytechnic, Tauranga, New Zealand.
Email: brett@waikato.ac.nz

This is the abstract for the podium presentation I made at the American College of Sports Medicine Annual Meeting in Baltimore 2010.

Abstract:
There is a need for markers of an athlete's training status that would help determine when the training load is either insufficient or excessive. PURPOSE: To examine the relationship between changes in performance and changes in physiological and psychological markers during and following a period of overload training in 12 female and 10 male elite rowers preparing for a world championship. METHODS: Performance was assessed with a rowing ergometer as mean power in a 30-min time trial and as power at a blood-lactate concentration of 4-mmol.L^{-1} in an incremental test. The time-trial provided changes in performance during the 4 weeks of overload, while the incremental test was undertaken at the beginning of the study and following a one-week taper after the overload. The following markers were assayed throughout the overload: mood state, sleep quality, perception of fatigue, and concentrations of cortisol, testosterone, dehydroepiandrosterone, creatine kinase, lactate dehydrogenase, 14 cytokines and C reactive protein in saliva and/or plasma. Plots of change in performance versus the 4-week change in each marker were examined for evidence of an inverted- U relationship that would characterise under- and over-training. Simple linear modelling was also used to estimate the effect of changes in the marker on changes in performance. RESULTS: Evidence of an inverted- U was apparent only for performance in the incremental test versus some plasma cytokines, and the relationship arose only because of relative underperformance of one rower who later excelled in competitions. Many changes considered to predict over-training (e.g., worsening mood, decreasing morning testosterone, increasing inflammation) actually had small to large positive linear relationships with performance. CONCLUSION: The markers investigated in this study were not useful for identifying excessive training in elite rowers but could indicate the need for an increase in training load.

APPENDIX I

Variability and Predictability of Finals Times of Elite Rowers

Smith, TB,[1]. & Hopkins, WG.[2]
[1]: University of Waikato (Hamilton, New Zealand), [2]: AUT (Auckland, New Zealand)

This is the abstract for the poster I presented at the European College of Sport Sciences Conference in Liverpool 2011

Introduction
Little is known about the competitive performance characteristics of elite rowers. We report here analyses of performance times for finalists at world cups, world championships and Olympics from 1999 to 2009.

Methods
A linear mixed model of finals times for single and crewed boats provided estimates of within-boat variability and between-boat differences as coefficients of variation; the model included terms to account for differences or changes in performance between calendar years, venues and the various levels of finals (A, B, C…).

Results
Differences in the effects of environmental conditions, estimated as variability in mean race time between finals, were very large to extremely large (2.3-4.0%), with possibly greater variability for singles compared with quads and eights. Within-boat race-to-race variability was 0.6-1.5% (90% confidence limits $\times/\div 1.09$-1.21); singles were almost certainly more variable ($>1.10\times$) than quads and eights, males were likely more variable than females, and variability in A finals of singles was very likely less than that in other finals. Overall, the variability of performance was similar to that in comparable endurance sports performed against water or air resistance (kayaking, swimming and cycling). Smallest worthwhile enhancements in performance, given by $0.3\times$ within-boat variability, ranged from ~0.3-0.4% for the singles through to ~0.2% for the eights. Between-boat differences in a given final ranged from small to large (0.5-2.0%), with differences for females likely more than those for males. Predictability of performance, expressed as within-year intraclass correlation coefficients, was moderate to very high (0.33-0.83), with females very likely to be more predictable than males.

Discussion
Estimates of the smallest important performance enhancements and the effects of environment, size of boat, gender, and level of the final on the various aspects of variability and predictability will help inform investigations of factors affecting elite competitive rowing performance.

APPENDIX J

Ethics Approval and Relevant Subject Information for the Study in Chapter 5.

M E M O R A N D U M
Auckland University of Technology Ethics Committee (AUTEC)

To:	Will Hopkins
From:	**Madeline Banda** Executive Secretary, AUTEC
Date:	15 January 2008
Subject:	Ethics Application Number 06/231 **Maximising 'supercompensatory' gains from overreaching in endurance athletes.**

Dear Will

I am pleased to advise that on 14 January 2008, I approved a minor amendment to your ethics application allowing interviews with the coaches. This delegated approval is made in accordance with section 5.3.2 of AUTEC's *Applying for Ethics Approval: Guidelines and Procedures* and is subject to endorsement at AUTEC's meeting on 11 February 2008.

I remind you that as part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online through *http://www.aut.ac.nz/about/ethics*. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 7 March 2010;

- A brief report on the status of the project using form EA3, which is available online through *http://www.aut.ac.nz/about/ethics*. This report is to be submitted either when the approval expires on 7 March 2010 or on completion of the project, whichever comes sooner;

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are reminded that, as applicant, you are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

Please note that AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to make the arrangements necessary to obtain this.

When communicating with us about this application, we ask that you use the application number and study title to enable us to provide you with prompt service. Should you have any further enquiries regarding this matter, you are welcome to contact Charles Grinter, Ethics Coordinator, by email at charles.grinter@aut.ac.nz or by telephone on 921 9999 at extension 8860.

On behalf of the AUTEC and myself, I wish you success with your research and look forward to reading about it in your reports.

Yours sincerely

Madeline Banda
Executive Secretary

Auckland University of Technology Ethics Committee

Participant Information Sheet

Date Information Sheet Produced: 19th December 2007

Project Title: Maximising 'supercompensatory' gains from overreaching in endurance athletes

An Invitation

You are invited to take part in the research project stated above. This project is part of Brett Smith's PhD and involves collaboration between Auckland University of Technology, Rowing New Zealand (RNZ), Sport and Recreation New Zealand and HortResearch. Your participation in this study is voluntary, and you may withdraw from the project at any time, without reason, and without any disadvantage.

What is the purpose of this research?

Elite endurance athletes are required to continually push their bodies to their physiological limit so as to ensure optimal conditioning and therefore run a high risk of overtraining. Overtraining can result in failure to fully recover from exercise leading to a reduction in conditioning and performance which in extreme cases can persist for weeks or months.

The aim of this study is to examine the tools the coach utilises to manage their athletes training programs so as to avert possible performance mal-adaption.

How was I chosen for this invitation?

As coach of the elite athletes involved in this study you determine how and when to modify training in response to actual or predicted overtraining. The coach is a vital factor in the management of strategies to avoid performance mal-adaption and it is important that their role in this process is examined.

What will happen in this research?

This project will be timetabled during a normal training mesocycle and will necessitate no changes in your normal coaching regimes. You will be interviewed approximately once a week and asked to discuss whether you modified your training because of concerns related to overtraining. Furthermore, if modifications were made you will be asked to detail what prompted these changes, what the changes were, how they were implemented and the outcome.

What are the discomforts and risks?

The interview will take approximately 20-30 minutes once a week for the duration of the study.

How will these discomforts and risks be alleviated?

You are asked to inform the researchers prior to completing any of the measures if you are uncomfortable with the procedures and they will attempt to find an option you are comfortable with. If during or after the interview you don't feel comfortable please tell the researchers and they will attempt to find a solution. Participants who are unhappy with the discomforts and risks can pull out of the study at any time without having to provide a reason and with no negative impacts.

What are the benefits?

The aim of this project is to develop non-invasive measures of overtraining that will help to maximise performance gains from training and reduce the risks of overtraining. The development of effective strategies to mitigate the effects of overtraining often must be managed by the coach and it is therefore important that their role in this process is examined.

What compensation is available for injury or negligence?

Compensation is available through the Accident Compensation Corporation within its normal limitations.

How will my privacy be protected?

All participants will be coded with a number that will be used throughout the study. No participant names will be released in any publications arising from this study. The data from this study will be stored indefinitely in a secure location by Rowing New Zealand and Brett Smith.

What are the costs of participating in this research?

The interview will take approximately 20-30 minutes per week.

What opportunity do I have to consider this invitation?

One week

How do I agree to participate in this research?

If you agree to participate in this research then please read and sign the attached consent form.

Will I receive feedback on the results of this research?

Yes. There will be a presentation to all participants at the end of the trial. Anonymity will be strictly adhered to.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Either, Brett Smith, brett@waikato.ac.nz, mob 021 627863 ph 07 8484500, or Will Hopkins will.hopkins@aut.ac.nz ph 09 921 9999

Concerns regarding the conduct of the research should be notified to the Executive Secretary, AUTEC, Madeline Banda, madeline.banda@aut.ac.nz , 921 9999 ext 8044.

Whom do I contact for further information about this research?

Either, Brett Smith, brett@waikato.ac.nz, 021 627863, or Will Hopkins will.hopkins@aut.ac.nz ph 09 921 9999

Consent Form

AUT UNIVERSITY
TE WĀNANGA ARONUI O TAMAKI MAKAU RAU

Project title: Maximising 'supercompensatory' gains from overreaching in endurance athletes

Project Supervisor: Professor Will Hopkins

Researcher: Brett Smith

○ I have read and understood the information provided about this research project in the Information Sheet dated 31 October 2007.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ I agree to take part in this research.

○ I wish to receive a copy of the report from the research (tick one): Yes ○ No ○

Participant's signature: ....………………………………………......

Participant's name: ....………………………………………….

Participants ethnicity : ....………………………………………….

Participant's Contact Details (if appropriate):

……………………………………………………………………………….

……………………………………………………………………………….

……………………………………………………………………………….

……………………………………………………………………………….

Date:

Note: The Participant should retain a copy of this form.

APPENDIX K

Examination of the Distance Recorded by the Nielsen Kellerman Impeller when Rowing over a 2000-m Rowing Course

Overview

From 2006 to 2010 we compared the distance recorded by the Nielsen Kellerman XL4 impeller (Nielsen Kellerman, Boothwyn, PA) with the 2000-m distance travelled in 61 races during various regattas. All the races were held on International Rowing Federation (FISA) sanctioned 2000-m rowing courses using FISA sanctioned timing systems. All Nielsen Kellerman (NK) units were positioned on the rowing boat hulls and calibrated prior to racing as per the directions in the installation and operation manual. Wind speed and direction were measured for each race. The NK units showed a negligible fixed error (0.1%), but there was a moderate random error of 1.2% even when wind direction and speed were taken into account. While this amount of error is only slightly larger than the 1% yardstick considered appropriate to accurately monitor training, the NK is not sufficiently accurate to quantify small but meaningful changes in 2000-m rowing boat speed. The NK may be useful for monitoring training speeds.

Introduction

A common method of measuring rowing boat speed is via impellers attached to the hull. The most popular device is the NK, which has two advantages over the GPS. The most important advantage is that it measures the speed of the boat relative to any water current, so impeller speed more accurately reflects the performance of the rowers. The other advantage of the NK is it potentially gives accurate readings over any distance, whereas GPS has large errors over short distances (see chapter 3 above). The impeller also has advantages over stopwatches, which require timing at each end of the course and accurate measurement of the distance. Despite the popularity of these impellers there are no published studies on their accuracy. We report here the standard error of estimate of the NK during regatta's over 2000-m adjusted for the effect of wind speed and direction.

Methods

Data was collected for 61 races from 10 different boats from the New Zealand elite rowing team. The impellers were placed on the rowing boat hulls as per the NK installation manual. The NK units were calibrated within a period of five days prior to racing, following the directions in the NK operation manual. All calibrations were conducted during periods of no wind on the regatta course where the race data was collected. The races were all held during regattas between 2006-2010 on various 2000-m FISA sanctioned International-rowing courses in Europe and New Zealand. The elapsed 2000-m times for each race were determined using a FISA sanctioned timing systems. During these regattas the rowing courses were fully buoyed and umpires ensured the competitors stayed within their allotted lanes throughout the entirety of their race.

The rowers turned the NK units on while sitting stationary in the starting pontoon immediately prior to the start of the race. The NK units were set to start recording as the first stroke of the race was taken. During the race the NK recorded distance, time and boat speed cumulatively for each successive stroke. The data was downloaded

from the NK units at the end of the race using the Nielsen Kellerman software and then uploaded to Microsoft Excel for further analysis.

For each 2000-m race, the official finish time recorded by the regatta timing system was matched to the cumulative time and associated cumulative distance recorded by the NK. This was done in excel by searching for the cumulative time recorded by the NK that matched the official finish time. As the NK unit only sampled at one-stroke intervals, the official finish time rarely matched any of the cumulative time measures recorded by the NK. We therefore interpolated the time and associated cumulative distance from the NK data that equated to the official finish time. In this manner we were able to calculate the distance recorded by the NK for the 2000-m regatta-course.

During each of the 61-recorded races the wind speed and direction were recorded at the finish line using a kestrel-3000 portable wind-meter. Above 1 m.sec^{-1} the wind speed was recorded to the nearest significant figure, while below 1 m.sec^{-1} it was expressed to the nearest multiple of 0.25 m.sec^{-1}. The direction of the wind was determined to the nearest 45 degree interval, these directions are tail, side tail, side, side head and head. We wished to express wind-speed and direction as a wind-speed value equivalent to either a direct headwind or tailwind, so as to enable these values to be entered into our analysis as a single covariate. To achieve this we used the spreadsheet developed by Kleshnev,[1] into which we entered the boat type, race time and average stroke rate for each race result. For each race we then entered the recorded wind-speed and direction, and the spreadsheet calculated how these conditions changed the boat speed. For each boat we then calculated the direct tailwind or headwind that caused the same change in boat speed. This direct wind-speed value was termed the derived wind-speed (-ve for headwind and +ve for tailwind) and used as the covariate in the analysis.

The distance recorded by the NK for each race was compared to the 2000-m race distance using the analysis of a post-only crossover trial spreadsheet, with adjustment for a covariate.[2] This spreadsheet provided the standard error of the estimate adjusted for the effect of the covariate. For this analysis the covariate was the calculated direct headwind or tailwind wind-speed value.

Results and Discussion

There was a moderate random error (1.2%) and the effect of the covariate (derived wind-speed) was trivial (0.1%). While the random error is only slightly larger than the 1% yardstick considered appropriate for monitor training (see Chapter 2), it is unlikely that this device is sufficiently sensitive to quantify small but meaningful changes in performance.

It is unlikely that the course distances and the finish times had more than trivial errors as FISA sets rigid standards and audits all the measurement devices. As most of this data was collected during International regattas, heavy security meant we were not able to independently audit the accuracy of these measures.

The NK units were calibrated in windless conditions, which led us to speculate that any wind during racing would have created currents and possibly effected steering leading to increased measurement error. It was therefore surprising to find that the derived wind speed had a trivial effect on this error. A possible limitation to our analysis on the effect of the wind, could have been the accuracy of the wind-speed and direction calculated by the portable wind-meter. It is also possible that the wind-speed measured at the finish line did not provide an accurate measure of the average wind conditions experienced by the boat as they raced along the 2000-m regatta course. If this experiment was to be repeated we would recommend a series of more advanced stationary weather stations placed just above water level along the length of the regatta course. We were also unable to confirm the accuracy of the calculations used to express wind-speed and direction as equivalent direct headwind or tailwind values.

Virtually all the boats in this experiment were coxless, therefore it is likely that the actual distances travelled by these boats during competition is longer than 2000m due to the difficulties in steering a perfect straight line while facing backwards. If only poor steering created the errors we would expect all the NK distances to be greater than 2000m, however 43% of the NK distances were less than 2000m, which can only be explained by measurement error.

Conclusion

The NK lacks the precision to accurately determine the effect of experimental or training interventions with elite rowers. It may however be a useful tool for monitoring training performance.

References

1.   Kleshnev V. Spreadsheet to calculate Training speeds at different stroke rates based on constant effective work per stroke. BioRow Ltd; 2006
2.   Hopkins WG. A spreadsheet for analysis of controlled trials with adjustments for a predictor. Sportscience 2006; 10: 46-50

APPENDIX L

Examination of the Distance Recorded by the GPSports Spi-Elite GPS when Rowing over a 2000-m Rowing Course

Overview

From 2006 to 2010 we compared the distance recorded by 10 SPI-Elite 1-Hz units (GPSports, Fyshwick, ACT) with the 2000-m distance travelled by 22 New Zealand rowing team boats during various regattas. All the races were held on International Rowing Federation (FISA) sanctioned 2000-m rowing courses using FISA sanctioned timing systems. All GPS units were positioned on the rowing boat to provide an unobstructed view of the sky and satellites. We determined a negligible standard error of the estimate (0.2%), which is sufficiently low to quantify small but meaningful changes in boat speed in 2000-m time trials. This accuracy requires perfect conditions (no wind or water currents) or constant wind and water currents during bouts of testing.

Introduction

While the most common tool for measuring boat speed is the Nielsen Kellerman impeller, GPS are becoming more popular as they are cheaper, far easier to install and don't require calibration. One weakness of these devices is that they measure boat speed relative to the land rather than the water, which will create inaccuracies in flowing water. Despite the increasing popularity of GPS in rowing, we weren't able to find any research that has examined the accuracy of a GPS during rowing. We report here the standard error of estimate of the Spi-Elite GPS during regatta's over 2000-m.

Methods

Data was collected for 22 races from 10 different Spi-Elite. The races were all held during regattas between 2006-2010 on various 2000-m FISA sanctioned International-rowing courses in Europe and New Zealand. The elapsed 2000-m times for each race were determined using a FISA sanctioned timing systems. During these regattas the rowing courses were fully buoyed and umpires ensured the competitors stayed within their allotted lanes throughout the entirety of their race.

Immediately prior to the crews going on the water for racing, the activated GPS were put in clear plastic waterproof holders that were positioned on the boat to provide an unobstructed view of the sky and satellites. Unfortunately, unlike the NK we were unable to set the Spi-Elite up to only recorded the race, instead it recorded everything while the boat was on the water. The data from the Spi-Elite was downloaded by Team AMS v2.0 software (GPSports, Australia), which presented the data as cumulative time, cumulative distance and velocity at one hertz from the time the unit started recording. To determine the distance recorded by the SPI elite during the 2000-m race we downloaded the data from the Team AMS software into Microsoft Excel.

The Spi-Elite start time and distance was determined as the point immediately prior to the first consistent increase above 0.1 m/sec on the velocity trace following the warm up period. This point was easy to determine, as all the boats were held stationary in the starting blocks for approximately 5 minutes prior to the race start. The Spi-Elite finish time for each race was calculated by adding the official race time to the Spi-Elite start time. As the Spi-Elite only sampled at 1-Hz the calculated finish time rarely matched any of the cumulative time measures recorded by the Spi-Elite. We therefore often had

to interpolate the Spi-Elite finish time and associated cumulative finish distance. The distance travelled by the Spi-Elite during the 2000-m race was this finish distance minus the start distance.

The distance recorded by the Spi-Elite for each race was compared to the 2000-m race distance. The standard error of the estimate was determined using Hopkins[1] (2000) analysis of validity by linear regression Excel spreadsheet.

Results and discussion

Calculations of the standard error of estimate for these results show both negligible fixed (0.1%) and random (0.2 %) errors. It is possible that this already negligible error was actually inflated due to poor steering causing the rowers to travel more than 2000m, slow reaction time at the start creating a difference between the official race time and the actual time taken to travel 2000m, and interpolation errors.

Unfortunately the wind conditions were not recorded when this validation was undertaken, but it is less likely that wind will have a major effect on the GPS, unless it is a wind that makes steering a straight course difficult. When we compare this error to the smallest worthwhile effect and race-to-race variability established in chapter 2, we can assert that the error of the Spi-Elite is acceptable for monitoring both racing and training for distances of 2000-m.

Conclusion

With ideal or constant environmental conditions the SPI-Elite has a measurement error sufficiently low to quantify small but meaningful changes boat speed in 2000-m time trials.

Reference

1.   Hopkins WG. Analysis of validity by linear regression (excel spreadsheet). A new view of statistics. Auckland: Sportsci.org: Internet Society for Sport Science; 2000

APPENDIX M

Change Scores and Technical Error of the Assays in Chapter 3

These data are discussed in Chapter 6.

Table 1. Standard deviations of the 4-wk change in steroid hormones and markers of muscle damage, and the technical error of the assays expressed as SD of change scores.

	FEMALE		MALE	
	4-wk change SD (%)	Technical error (%)	4-wk change SD (%)	Technical error (%)
STEROID HORMONES				
Testosterone: (pg/ml)	18	6	11	6
Cortisol: (ng/ml)	17	15	22	15
DHEA: (ng/ml)	25	14		
Test: within-wk chg	27	6	7.2	6
Cort: within-wk chg	31	15	43	15
DHEA: within-wk chg	23	14		
Test: Pre 30-min (pg/ml)	35	6	20	6
Cort: Pre 30-min (ng/ml)	69	15	86	15
DHEA: Pre 30-min (ng/ml)	55	14		
Test: post/pre 30-min	33	6	44	6
Cort: post/pre 30-min	63	15	82	15
DHEA: post/pre 30-min)	55	14		
MARKERS OF MUSCLE DAMAGE				
CK: (U/L)	53	5	46	5
LDH: (U/L)	25	3	14	3
CK: within-wk chg	36	5	17	5
LDH: within-wk chg	24	3	25	3

Change scores are linearized change over four tests (Weeks 1 to 4).

Technical error is the intra assay coefficient of variation for the marker multiplied by $\sqrt{2}$ and represents a standard deviation of change scores arising solely from the assay.

Test, Testosterone; Cort, Cortisol; within-wk chg, the linearized factor change from Monday to Friday determined from each individual's line of best fit for each day of the week; Pre 30-min, mean value for markers collected immediately before the 30-min test; Post/Pre 30-min, difference between the mean value for the marker collected immediately post the 30-min test and the mean value collected immediately pre the 30-min test.

Table 2. Standard deviations of the 4-wk change in the cytokines and CRP, and the technical error of the assays expressed as SD of change scores.

	FEMALE		MALE	
	4-wk change SD (%)	Technical error (%)	4-wk change SD (%)	Technical error (%)
CYTOKINES AND CRP				
CRP: plasma (ng/ml)	70	12	265	12
CRP: saliva (ng/ml)	297	12	187	12
IFNalpha: plasma (pg/ml)	146	12	377	12
IFNalpha: saliva (pg/ml)	161	12	197	12
IFNgamma: plasma (pg/ml)	81	12	127	12
IFNgamma: saliva (pg/ml)	242	12	184	12
IL1beta: plasma (pg/ml)	99	12	194	12
IL1beta: saliva (pg/ml)	72	12	66	12
IL2: plasma (pg/ml)	1511	12	706	12
IL2: saliva (pg/ml)	823	12	471	12
IL4: plasma (pg/ml)	1400	12	628	12
IL4: saliva (pg/ml)	743	12	494	12
IL5: plasma (pg/ml)	138	12	334	12
IL5: saliva (pg/ml)	182	12	116	12
IL6: plasma (pg/ml)	23	12	23	12
IL6: saliva (pg/ml)	15	12	19	12
IL8: plasma (pg/ml)	86	12	52	12
IL8: saliva (pg/ml)	78	12	63	12
IL10: plasma (pg/ml)	268	12	468	12
IL10: saliva (pg/ml)	341	12	167	12
IL12p70: plasma (pg/ml)	481	12	604	12
IL12p70: saliva (pg/ml)	251	12	151	12
IL18: plasma (pg/ml)	67	12	102	12
IL18: saliva (pg/ml)	123	12	120	12
MCP1: plasma (pg/ml)	53	12	38	12
MCP1: saliva (pg/ml)	68	12	101	12
TNFalpha: plasma (pg/ml)	82	12	151	12
TNFalpha: saliva (pg/ml)	55	12	54	12
TNFbeta: plasma (pg/ml)	459	12	805	12
TNFbeta: saliva (pg/ml)	607	12	2148	12

Change scores are linearized change over four tests (Weeks 1 to 4).

Technical error is the intra assay coefficient of variation for the marker multiplied by $\sqrt{2}$ and represents a standard deviation of change scores arising solely from the assay.

APPENDIX N:

International Rowing Federation Magazine Article

Feel terrible, row fast –
The science behind the Kiwis

New Zealand's rowing programme is going through a boom. In 1978, when the World Rowing Championships were last held in the country, one medal was obtained. In 2010 the medal count was 10. As in any success, there are always a number of factors behind it. One factor is the attention to detail and within this the use of sports science.

One of New Zealand's premier sports science experts who has played a key role in the success of the New Zealand rowing team over the past decade, Brett Smith, believes that the key has been coaches 'buying in' to the use of sports science. Smith is a sports physiologist consulting for Rowing New Zealand and also a lecturer at the University of Waikato.

New Zealand Rowing's head coach, Dick Tonks, has adopted the methods, Smith asserts. "Dick is the architect of all of the recent successes (in New Zealand). I have been very fortunate as he's been extremely supportive and allowed me to constantly experiment. So other coaches here have applied it as well."

Smith believes it is common for coaches to fail to take advantage of what sports science has to offer. Part of that is the fault of the sports scientists, says Smith. As a profession they have not always proved their worth in an applied sense. "Coaches

don't want scientific terms and theories. They want to know what is going to make the boat go faster, not sports science jargon."

"Most sports science is interested in the various theoretical mechanisms. If an athlete has gotten better on a test, the scientist will assume that the boat performance will also go up. So, if for example the VO2 max or lactate threshold improves, that's interesting, but for the New Zealand rowers we're more interested if performance improves. The bottom line is boat speed."

Smith notes that sports scientists also tend to look at signs and symptoms of overtraining and use these to advise the coaches when to lighten the training load. Smith has a different approach.

Eric Murray of New Zealand is congratulated by teammate Rebecca Scown after winning the Men's Pair Final.

Mahe Drysdale (NZL), 2009 World Champion

©Igor Miejer

©Hagen Hopkins/Getty Images

∨ Feel terrible, row fast – The science behind the Kiwis

∨ Most top coaches know their athletes will feel bad during hard training and also know that hard training is essential for success. "If you measure a large increase in some type of stress (eg. decreased immune function) most scientists will say this is bad and recommend that the coach 'reduce training'. But greater stress can also equal greater adaptation, so reducing training for a stressed athlete may possibly rob them of a gold medal. What we found after two years of comprehensive research examining a wide variety of stress markers was that all rowers with abnormally high increases in psycho-physiological measures of stress got medals."

Smith says that often research is done using high school or recreational athletes so if stress is too high then they will fall over. "But elite athletes respond differently. Science is all about a normal distribution curve, but with elite athletes you don't always have the bell curve shape because you're dealing with the abnormal people, the outliers (on the normal distribution curve). So gold medallists are often abnormal."

"Some people say athletes will choose the best training for themselves, but that is not always correct," says Smith. "The brain is often hardwired to reduce stress and the punishment of regular intensive rowing training and racing is arguably extremely stressful." So feedback from the brain may not be what is best for going fast. "The

normal brain's role is perhaps to inhibit intensive, potentially damaging training."

Smith commented that he found most of the athletes who had the worst mood changes and greatest sleep disturbances actually ended up performing best at the World Rowing Championships. "While a sports psychologist might look at these psychological measures and say, 'Hey, they're being over trained', I would say they're only 'overtraining' when their boat speed falls off."

But this does not mean ignoring the reaction of athletes. "Most top coaches know that athletes have to feel really bad and accept it's going to happen. The good coaches watch the boat (speed) and body language of the rowers and compare what they see with what their rowers tell them."

With Rowing New Zealand, Smith's role has been to give coaches feedback on every single training session. "The bottom line is boat speed. So, for example, in August we want to be doing 'x' percentage of the World Best Time for 'y' (kilo)metres."

Smith describes one of the changes they made after observing athlete behaviour. Smith and the coaches were looking at recovery as they tended to find crews getting slower between the heats and finals of the international regattas. Smith followed the lead taken by cycling where athletes employ a systematic post-race warm-down.

"The best way to recover is to keep the blood flowing and the body metabolising waste products to get back to normal homeostasis. If done properly, in some cases it could actually improve the athletes' performance for their next race."

So they worked on ways to get the athletes recovered as quickly as possible. "We used compression, cooling, active recovery and nutrition." This means that at all big regattas New Zealand has a warm-up and cool-down area which includes a mixture of rowing machines, bikes, compression gear, appropriate nutrition and cooling equipment.

This detail, Smith believes, is part of the current success being experienced by New Zealand rowers.

■ R.S.B.

Brett Smith
©Waikato Print

Joseph Sullivan (L) and Nathan Cohen of New Zealand hold their gold medals after the Men's Double Sculls Final at the 2010 World Rowing Championships.

© Sandra Mu/Getty Images