Confounders of haemoglobin mass in athletes: implications for anti-doping

A thesis submitted in the fulfilment of the requirements for the degree of

Doctor of Philosophy

of the University of Canberra

Clare E. Humberstone

BSc. (Hons) Sports Science and Physiology (Ind), University of Leeds, 2006

May 2012

Abstract

The Athlete Biological Passport (ABP) is used for the indirect detection of blood doping in athletes. The calculations within the ABP (Adaptive model) monitor longitudinal changes in characteristics of an athlete's blood, specifically haemoglobin concentration ([Hb]) and percent reticulocytes (%Ret), and 'flag' unusual variations that may be characteristic of doping. The inclusion of total haemoglobin mass (Hb_{mass}) as a marker in the ABP may improve the Passport's sensitivity. However, concerns have been expressed about the suitability of Hb_{mass} measurements for this purpose. This thesis investigated two potential barriers to including Hb_{mass} in the ABP: the lack of a quality control system for Hb_{mass} measurement and the potential confusion between longitudinal Hb_{mass} profiles of doped and non-doped athletes.

The use of custom-made quality control solutions eliminated the majority of between-laboratory differences in Hb_{mass} measures. Analytical error associated with making successive Hb_{mass} measurements in three different laboratories was reduced from 2.4% to 1.7% when the quality controls were used. These findings demonstrated that using quality control solutions would ensure that Hb_{mass} results from different laboratories were equivalent if Hb_{mass} was included as a marker in the ABP.

The effects of various confounders on Hb_{mass} in non-doped athletes were quantified. These confounders were investigated specifically for their potential to increase the biological variability of Hb_{mass}. For detection of blood doping, these confounders represent a 'worst-case-scenario' for the variability of Hb_{mass} in *non-doped* athletes. Ultra-endurance triathlon racing (+3.2%), Classical altitude training (+3.8%) and Live High:Train Low altitude training (two estimates, +4.0% and +4.3%) each caused substantial mean increases in Hb_{mass}. Conversely, reduced training (-2.3%) and surgery (-2.7%) in injured / ill athletes caused

substantial mean decreases in Hb_{mass}. Acute Intermittent Hypoxic Exposure did not substantially affect Hb_{mass} (-0.3%). The effects of microdoses of recombinant human erythropoietin (rHuEPO) on Hb_{mass} were also examined to represent a 'worst-case-scenario' for the detection of *doped* athletes using Hb_{mass} in the ABP. Over a 12 week period, rHuEPO microdosing caused a mean increase of Hb_{mass} of 11.0%, with individual responses ranging from increases of 1.4% to 19.2%.

Finally, an investigation was carried out to determine which of six different Hb_{mass} Adaptive models might be suitable for inclusion in the ABP. The sensitivities and specificities of these models were compared in a sample of 159 non-doped and 18 doped athletes. In models that used Hb_{mass} as a single marker, the sensitivity and specificity of the model was heavily influenced by the estimate of Hb_{mass} biological within-subject (BioWS) variance included in the calculations. These models were each named after the first author of a publication in which the BioWS variance of Hb_{mass} in athletes was estimated. Due to their low specificities in non-doped athletes, neither of two $Hb^{m\ (Prommer)}$ models would be suitable for inclusion in the ABP. In contrast, based on specificities close to 100%, any of the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ or $Hb^{m\ (Eastwood)}$ models may be suitable for inclusion, although each model only offered ~20% sensitivity to rHuEPO doping. The novel ON^{hm+ret} model, which combined Hb_{mass} and %Ret into a single marker, would not be useful in the ABP due to its low specificity.

Overall, the inclusion of Hb_{mass} may improve the sensitivity of the ABP, particularly to microdose rHuEPO doping. The sensitivities of the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models should be examined in a larger sample of doped athletes. Unfortunately, these models may be susceptible to recording false-positive results in some extreme cases of Hb_{mass} perturbation in ill or injured athletes.

Acknowledgements

I initially had real doubts about whether a PhD was the right path for me, but with encouragement from the lovely Dan and our friend, Jo Vaile, I embarked on a journey that has enriched me with a level of skill, knowledge and a confidence that I never anticipated. It has been an experience that I will look back on with great fondness and I have many people to thank for getting me to this point.

To the lovely Dan. Thank you for being the super boyfriend and now the super husband. Your support for me is unwavering and I couldn't have achieved all that I have over the past few years without you. Everything is more fun when you're there. I love you. Thank you xxx

To Mum, Dad, Catherine and Mel. We have a wonderful family and the support you have all given me my whole life has got me to this point and beyond. Thank you for all your love, generosity, and teaching, and for having faith in me to make the right decisions. I'm sad that we now live so far apart, but the distance is only physical and I think we are all closer than ever.

To my good friend Jo Vaile. I feel so privileged to be your buddy. The encouragement, support, walks, enforced runs, and fun you contribute to my life are all equally important and without you I wouldn't have had the confidence to embark on the PhD in the first place. Thanks!

To my fantastic supervisors Prof. Chris Gore, Dr. Judith Anson and Dr. Philo Saunders. The guidance you have given me has been invaluable and I'm constantly amazed by the time you

dedicate to helping me. The opportunity to work with you has been a real highlight of the PhD experience and the skills I have learned from you will benefit me for a long time to come.

The advice of my Hb_{mass} PhD predecessors, Dr. Laura Garvican, Dr. Eileen Robertson and Dr. Annette Eastwood, has been invaluable and I want to thank you all for giving up your time for me to ask questions, compare notes and help with testing throughout my PhD.

To Dr Ken Sharpe, your willingness to advise me, help me investigate the data, and generally explain mathematical concepts to me was invaluable; thank you. To Dr. Mike Ashenden, thank you for being available to discuss my ideas and offer critical bits of advice. To Dr Pierre-Edouard Sottas, thank you for granting me access to the Athlete Biological Passport software and for offering your guidance and advice about its use.

To all the athletes who volunteered to take part in my studies, I can't thank you enough. You all participated with enthusiasm and great humour which made the hours of sitting at the OSM3 analysing blood worthwhile.

To everyone who helped me with the testing, thank you so much. I couldn't have done it without you: Nicola Bullock, Darrell Bonetti, Laura Garvican, Nicole Prommer, Melissa Arkinstall, Kiara Johnson, Jesse Featonby, Graeme Albon, Wei Chung, Lizzie Wraith, Peter Fowler, Annette Eastwood, Jo Vaile and Dean Higham.

To Maz, thanks for all your work behind the scenes, for the hugs and encouragement and for all times you said, "You know how I like to tell you....".

Finally, the attitude towards excellence demonstrated by all my colleagues at the Australian Institute of Sport is infectious and AIS Physiology will hold a cherished place in my heart for a long time to come.

Declaration

In this thesis I detail the findings from research carried out between January 2009 and May 2012. The research studies described in Chapters 3, 4, 5, 6, 7 and 8 were carried out in collaboration with my co-authors, the names of whom are listed at the start of each chapter. For each of these studies I took a lead role in the experimental design, subject recruitment, data collection and analysis, and I wrote the manuscripts.

Note, I got married in August 2011 and subsequently I changed my name from Clare E. Gough to Clare E. Humberstone. Consequently, I used my maiden name for the submission of publications from this thesis, whilst for this dissertation I have used my married name.

Publications and Presentations relevant to this thesis

Peer-reviewed journal articles

Gough C.E., Saunders P.U., Fowlie J., Savage B., Pyne D.B., Anson J.M., Wachsmuth N., Prommer N. and Gore C.J. Influence of altitude training modality on performance and total haemoglobin mass in elite swimmers. *European Journal of Applied Physiology*. 2012. [Epub ahead of print]. DOI: 10.1007/s00421-011-2291-7

Gough C.E., Sharpe K., Ashenden M.J., Anson J.M., Saunders P.U., Garvican L.A., Bonetti D.L., Gore C.J. and Prommer N. Quality control technique to reduce the variability of longitudinal measurement of hemoglobin mass. *Scandinavian Journal of Medicine and Science in Sports*. 2011 (6): e365-71. DOI: 10.1111/j.1600-0838.2011.01316.x.

Gough C.E., Eastwood A., Saunders P.U., Anson J.M. and Gore C.J. Spurious Hb mass increases following exercise. *International Journal of Sports Medicine*. 2012 (In press).

Gough C.E., Sharpe K., Garvican L.A., Anson J.M., Saunders P.U. and Gore C.J. The effects of injury and illness on haemoglobin mass. (Submitted to *International Journal of Sports Medicine* in March 2012; In Review)

Gough C.E., Saunders P.U., Bonetti D.L., Stephens S., Bullock N., Anson J.M. and Gore C.J. Comparison of Acute Intermittent Hypoxic Exposure and Live High:Train Low altitude. (Submitted to *International Journal of Sports Physiology and Performance* in January 2012; In Review)

Ashenden M., **Gough C.E.**, Garnham A., Gore C.J. and Sharpe K. Current markers of the Athlete Blood Passport do not flag microdose EPO doping. *European Journal of Applied Physiology*. 2011 111(9): 2307-14.

Peer reviewed conference proceedings

Gough C.E., Bullock, N., Bonetti D.L., Saunders P.U., Anson J., Stephens S. and Gore, C.J. Acute Intermittent Hypoxic exposure is inferior to live high/train low altitude for haemoglobin adaptation. *Proceedings of the 15th Annual Congress of the European College of Sport Science.* 2010, p 59.

Gough C.E., Eastwood A., Saunders P.U., Anson J. and Gore C.J. Ultra-endurance triathlon racing spuriously increases haemoglobin mass. *Proceedings of the 15th Annual Congress of the European College of Sport Science.* 2010, p 364.

Conference presentations

Gough C.E., Eastwood A., Saunders P.U., Anson J. and Gore C.J. Ultra-endurance triathlon racing spuriously increases haemoglobin mass (Oral presentation).
Congress of the European College of Sports Science, Turkey, 2010

Gough C.E., Bullock N., Bonetti D.L., Saunders P.U., Anson J., Stephens S. and Gore C.J. Acute Intermittent Hypoxic exposure is inferior to live high/train low altitude for haemoglobin adaptation (Poster presentation).
Congress of the European College of Sports Science, Turkey, 2010

Table of Contents

List of Tables

List of Figures

CHAPTER 7

CHAPTER 8

CHAPTER 9

List of Abbreviations

ABP	Athlete biological passport
AcIHE	Acute intermittent hypoxic exposure
BioWS variance	Biological within-subject variance
Classic	Classical altitude training
CI	Confidence interval
CL	Confidence limits
CO	Carbon monoxide
CV	Coefficient of variation
cCVH	Current cardiovascular training hours
cCVI	Current cardiovascular training intensity
cTTH	Current total training hours
CO_2	Carbon dioxide
EPO	Erythropoietin
ESAs	Erythropoiesis-stimulating agents
[Ferr]	Serum ferritin concentration
FINA	Federation Internationale de Natation
FIS	Federation Internationale de Ski
[Hap]	Serum haptoglobin concentration
Hb	Haemoglobin

[Hb]	Haemoglobin concentration
Hb_{mass}	Haemoglobin mass
Hct	Haematocrit
HIF-1	Hypoxia–inducible factor 1
HR	Heart rate
IPS	International point score
[Lac]	Blood lactate concentration
LHTL	Live High:Train Low altitude training
ln	Natural-log transfromed
Mb	Myoglobin
O_2	Oxygen
%Ret	Percent reticulocytes
rHuEPO	Recombinant human erythropoietin
RPE	Rating of perceived exertion
SD	Standard deviation
SpO_2	Peripheral oxygen saturation
[STfR]	Serum soluble transferrin receptor concentration
SWC	Smallest worthwhile change
TE	Typical error
TRIMP	Training impulse
TTE	Time to exhaustion

U23	Under-23
UCI	Union Cycliste Internationale
VO_2	Volume of oxygen consumed
VO_{2peak}	Peak oxygen consumption
vVO_{2peak}	Velocity at VO_{2peak}
WADA	World Anti-Doping Agency

Chapter 1

Introduction

Sporting success can lead to substantial riches, glory and fame. It is, therefore, not surprising that some athletes will seek to enhance their own sporting performance by unethical means. Blood doping, a general term for the illegal manipulation of an athlete's blood, may improve performance by 2-4% (Brien et al. 1989; Brien and Simon 1987). Consequently, and because blood doping is also contrary to the spirit of sport as well as potentially harmful to health, it is outlawed by the World Anti-Doping Agency (WADA).

The Athlete Biological Passport (ABP) is a longitudinal monitoring system used by WADA for the indirect detection of blood doping. The Adaptive model of the ABP compares an athlete's current blood test results to their own previous results, with any suspicious longitudinal changes in the athlete's blood profile being 'flagged' for review by a panel of experts. Currently, the key blood markers included in the Adaptive models of the ABP are haemoglobin concentration ([Hb]), percentage reticulocytes (%Ret; the proportion of red blood cells in a sample that are immature) and an integrated measure of these two markers, the OFF-hr Score. A number of researchers have suggested that the inclusion of total haemoglobin mass (Hb_{mass}) as an additional marker would improve the sensitivity (rate of correct identification of doped athletes) of the ABP (Morkeberg et al. 2011; Prommer et al. 2008). However, doubts have also been expressed about the suitability of Hb_{mass} measurement for anti-doping purposes (Eastwood et al. 2011b; Lundby and Robach 2010; Schumacher and Pottgiesser 2010).

First, there is no quality control system for Hb_{mass} available to ensure equivalency of measurements made in different laboratories (Schumacher and Pottgiesser 2010). Every year, athletes travel extensively for training and competition. Anti-doping authorities may request that an athlete submits a blood sample for analysis in the ABP at any time, with the sample being analysed at the nearest of 33 WADA-accredited laboratories (World Anti-Doping Agency 2010). Consequently, sequential ABP results for an individual athlete may originate from a number of different laboratories and it is crucial that the between-laboratory differences in measures are minimised. Although all WADA-accredited laboratories participate in quality control programs for their haematology analysers to ensure the equivalency of all [Hb] and %Ret results, currently no such quality control procedures exist for Hb_{mass}. This remains a major barrier to the inclusion of Hb_{mass} in the ABP.

Second, doping-induced changes in Hb_{mass} may not be distinguishable from the normal within-subject variation in Hb_{mass} (Eastwood et al. 2011b; Lundby and Robach 2010), which could compromise the specificity (rate of correct identification of non-doped athletes) of Hb_{mass} measurement for anti-doping purposes. Typically, the day-to-day fluctuations of Hb_{mass} in non-doped athletes (~2-4%) (Eastwood et al. 2011b; Prommer et al. 2008) are smaller than the changes in Hb_{mass} that result from blood doping (~6-20%) (Parisotto et al. 2000a; Pottgiesser et al. 2009b). However, some common factors in the lives of athletes, such as prolonged exercise, illness, injury, detraining and altitude exposure, have the potential to increase the fluctuations in Hb_{mass} in non-doped athletes. Additionally, microdoses of recombinant human erythropoietin (rHuEPO) by doped athletes may result in small but beneficial increases in Hb_{mass}. Collectively, these factors may influence Hb_{mass} in ways that blur the line between doped and non-doped athletes but the effects of these confounders have not, as yet, been examined in sufficient detail. In-depth investigations would allow anti-

doping authorities to determine the sensitivity and specificity of different Adaptive models based on Hb_{mass}, and would allow an informed decision to be made about whether to include a particular model in the ABP.

Aims

The aims of this thesis were:

(i) to develop a quality control system for Hb_{mass} measurement;

(ii) to quantify the potential confounding effects of prolonged exercise, illness, injury, different forms of hypoxia and microdoses of rHuEPO on Hb_{mass}; and,

(iii) to examine the sensitivities and specificities of different Adaptive models based on Hb_{mass} in the ABP.

The specific aims of each of the studies contained in this thesis were:

Chapter 3: Quality control technique to reduce variability of longitudinal measurement of haemoglobin mass

- to examine the variation in Hb_{mass} when consecutive measurements are made over a short period of time in different laboratories
- to investigate the efficacy of using custom-made quality control solutions to reduce the variability between Hb_{mass} measurements made in different laboratories

Chapter 4: Spurious Hb_{mass} increases following exercise

- to examine the immediate effect of ultra-endurance triathlon racing on Hb_{mass}

Chapter 5: The effects of injury and illness on haemoglobin mass

- to quantify the effects of reduced training, surgery and changes in body mass on Hb_{mass} in injured or ill athletes
- to model the dose-response effect of a decrease in training on Hb_{mass}

Chapter 6: Influence of altitude training modality on performance and total haemoglobin mass in elite swimmers

- to quantify and compare the effects of Classical (Classic) altitude training and Live High:Train Low (LHTL) altitude training on Hb_{mass} in elite swimmers
- to quantify and compare the effects of Classic altitude training and LHTL altitude training on swimming performance
- to examine the relationship between changes in Hb_{mass} and changes in swimming performance following Classic altitude training and LHTL altitude training

Chapter 7: Comparison of Acute Intermittent Hypoxic Exposure and Live High:Train Low altitude

- to compare the effects of LHTL altitude training and Acute Intermittent Hypoxic Exposure (AcIHE) on Hb_{mass}
- to compare the effects of LHTL altitude training and AcIHE on the physiological characteristics of running

Chapter 8: Does the inclusion of haemoglobin mass in the Athlete Biological Passport improve detection of microdose rHuEPO doping?

- to assess the sensitivity and specificity of six Adaptive models based on Hb_{mass} to microdose rHuEPO doping

Chapter 9: Discussion

- to assess the sensitivity and specificity of six Adaptive models based on Hb_{mass} in a large sample of athletes, with considerations made for the effects of the biological within-subject (BioWS) variance, testing frequency and the utility of combining Hb_{mass} and %Ret

Chapter 2

Literature Review

Introduction

At the top level, races can be won and lost by one thousandth of a second. In more general terms, it has been demonstrated that a worthwhile improvement (one that increases the likelihood of medalling) is ~0.4% (Hopkins 2005; Pyne et al. 2004). Thus, it is not surprising that athletes may resort to methods that could boost their performance. Blood doping is one favoured method that may improve performance by ~1-4% (Brien et al. 1989; Brien and Simon 1987). However, blood doping, the illegal manipulation of an athlete's blood for the purpose of enhancing athletic performance, is prohibited by the World Anti-Doping Agency (WADA) (World Anti-Doping Agency 2012). One important predicator of endurance performance is maximal oxygen uptake (VO_{2max}) (Jacobs et al. 2011; Saunders et al. 2010b). In well-trained athletes, oxygen supply to working muscles is the major limiting factor to VO_{2max} (di Prampero 2003; Wagner 2006). As a result, the common aim of all blood doping methods is to increase the oxygen-carrying capacity of the blood to artificially enhance VO_{2max} and, thus, performance (Gaudard et al. 2003).

Haemoglobin (Hb) is a protein contained within red blood cells that reversibly binds to oxygen in the capillaries of the lungs and carries oxygen to respiring cells around the body (Hsia 1998). The most common way for blood doping methods to increase the oxygen-carrying capacity of the blood is to increase the amount of Hb in the circulation (Gaudard et al. 2003). This may be achieved in a variety of ways (Figure 2.1). Consequently, the

longitudinal monitoring of Hb in the blood of athletes is an important indirect method of detecting blood doping.

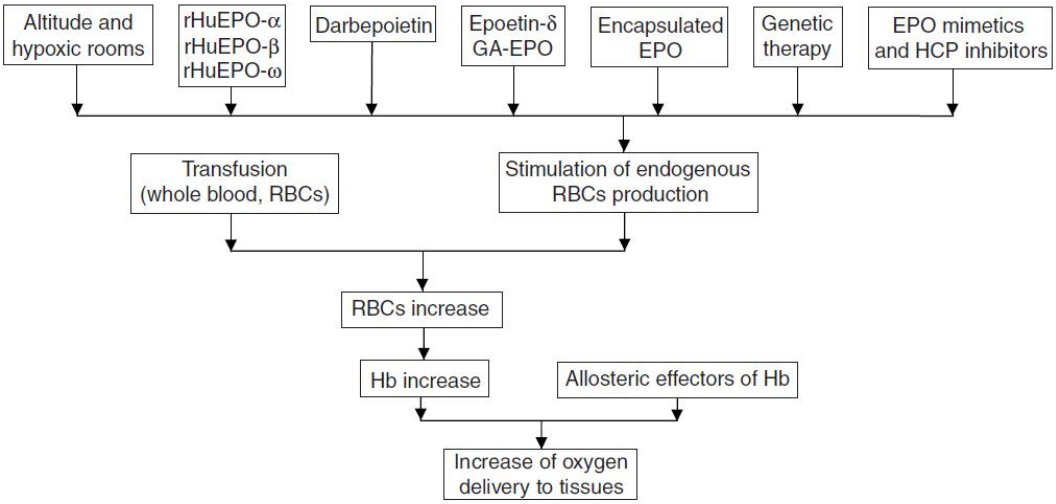

Figure 2.1: Methods to increase oxygen delivery to the tissues by direct action on haemoglobin. All methods listed here, except 'Altitude and hypoxic rooms', are blood doping and are prohibited by the World Anti-Doping Agency. EPO = erythropoietin; GA-EPO = gene-activated erythropoietin; HCP = haematopoietic cell phosphatase; RBCs = red blood cells; rHuEPO = recombinant human erythropoietin. Figure sourced from Gaudard et al. (2003).

Since the 1990s, international sport federations and anti-doping organisations have serially monitored Hb concentration ([Hb]) in athletes. In 1997, the Federation International de Ski (FIS; the international governing body for skiing) implemented a "no start" ruling for athletes whose [Hb] exceeded the cut-offs of 185 $g.L^{-1}$ for men and 165 $g.L^{-1}$ for women (Videman et al. 2000). Since 2000, models tracking a range of blood markers in athletes, including [Hb], have been used to detect blood doping. The first and second generation models compared an athlete's blood profile against those of a typical athletic population (Gore et al. 2003;

Parisotto et al. 2000a). Athletes whose results fell outside the normal range of the population were suspected of blood doping. In the third generation (3G) model, the emphasis shifted from comparing athletes against population norms to, instead, comparing the athlete's blood results against their own previous test results (Sharpe et al. 2006). The most recently developed model, the Adaptive model of the Athlete Biological Passport (ABP) (Sottas et al. 2010), has adopted a similar philosophy to the 3G model in that it compares athletes against individualised reference ranges. These ranges are generated primarily from an individual's test history but also include a contribution from population normative data.

The ABP is a longitudinal tracking system of the biological characteristics of athletes, currently endorsed by WADA for use by national anti-doping organisations and international sports federations to detect blood doping. The ABP was originally developed to allow detection of androgen doping, growth hormone doping and blood doping, although the procedures for the detection of androgen and growth hormone doping are still being refined and have not yet been implemented. The haematological module of the ABP is concerned with detection of blood doping and monitors changes in [Hb], percent reticulocytes (%Ret; the proportion of red blood cells in a sample that are immature) and an integrated measure of these two markers, the OFF-hr score. The calculations used within the Adaptive model progressively integrate an individual athlete's past test results with population values to create an individualised reference range against which subsequent test results are compared (Sottas et al. 2007).

Although the theoretical basis behind the adoption of the ABP has, in general, been lauded as a step-forward for blood doping detection (Lundby et al. 2012; Pottgiesser et al. 2011), concerns have been expressed about the susceptibility of [Hb] and %Ret to manipulation

(Sanchis-Gomar et al. 2010a; Schumacher and Pottgiesser 2010). The success of an anti-doping model is determined by its sensitivity (rate of correct identification of doped athletes) and specificity (rate of correct identification of non-doped athletes). If athletes are able to artificially influence [Hb] and %Ret in a way that hides the effects of doping, the sensitivity of the ABP could be substantially reduced. For example, the use of plasma volume expanders such as hydroxyethyl starch can mask a doping-induced increase in [Hb] (Sanchis-Gomar et al. 2010b). In addition, by using an intelligent (but illegal) combination of blood transfusions, recombinant human erythropoietin (rHuEPO) injections and altitude exposure, the fluctuations in %Ret that are characteristic of doping can be mostly neutralised (Sanchis-Gomar et al. 2010a). One possible solution is to supplement the existing ABP markers with the measurement of an additional marker, total haemoglobin mass (Hb_{mass}). By measuring the *amount* of Hb in the blood, rather than its *concentration*, the effect of plasma volume fluctuation is negated. Furthermore, by including Hb_{mass} in the ABP, the precise variable that doped athletes are aiming to increase would be directly monitored (Prommer et al. 2008). It has been argued by a number of researchers that the addition of Hb_{mass} to existing anti-doping markers would improve the detection sensitivity of the ABP (Morkeberg et al. 2011; Pottgiesser et al. 2007; Prommer et al. 2008; Sottas et al. 2010). Others have expressed doubts, however, about whether the perturbation in Hb_{mass} associated with doping would be distinguishable from the normal within-subject variation of Hb_{mass}, calling into question the specificity of Hb_{mass} in an anti-doping context (Eastwood et al. 2011b; Lundby and Robach 2010). Clearly, there are questions that remain unanswered about the suitability of Hb_{mass} as a marker in the ABP.

Indirect detection of blood doping and the Athlete Biological Passport

The practice of athletes using blood transfusions to boost their Hb_{mass} and improve endurance performance dates back at least 40 years (Eichner 2007). Homologous blood transfusion refers to the practice of transfusing blood from another person compatible for ABO and Rhesus D blood groups, whereas autologous blood transfusion involves the withdrawal and subsequent reinfusion of one's own blood (Giraud et al. 2010). It is well known that both homologous and autologous methods of doping were used by athletes in the 1970s and 1980s, and although no reliable detection tests existed at the time, both practices were banned by the International Olympic Committee after the 1984 Olympics (Lippi and Banfi 2006). However, transfusions carry significant safety risks (such as, transmission of disease between individuals and illness resulting from inappropriate blood storage) and present logistical difficulties (transporting and storing the blood) (Lippi et al. 2006). After the inception of genetic engineering in the late 1980s, the availability of rHuEPO on the black market provided athletes with an easier method of performance enhancement and, in the absence of a test for detecting its use, rHuEPO became the doping method of choice during the 1990s (Lippi et al. 2006).

Since there was no direct test capable of detecting rHuEPO use, sports authorities began to prevent the participation of athletes whose blood characteristics were suggestive of doping. The Union Cycliste Internationale (UCI; the international governing body of cycling) imposed a "no start" rule for athletes with a haematocrit (Hct) >50% (males) or >47% (females), similar to the FIS's upper limits on [Hb] in skiiers (Morkeberg et al. 2008; Videman et al. 2000). Although a direct test for detecting rHuEPO in urine was published and implemented by WADA in 2001 (Lasne 2001), indirect methods of detection continued to be

an important aspect of the fight against blood doping. After 2001, autologous transfusions once again became popular as a method to avoid detection of illicit activity (Lippi and Banfi 2006) and to this day, there is no direct test for autologous doping available. Since rHuEPO initially became available, a range of Erythropoiesis-Stimulating Agents (ESAs) have come onto the market, leaving anti-doping authorities constantly scrambling to modify their direct testing protocols to be specific and sensitive to the new drugs (Reichel 2011). The advantage of indirect methods of detection via modelling of expected physiological changes is that they facilitate detection of doping, even when the method of doping is otherwise undetectable.

In the early 2000s, more sophisticated models for indirect detection of doping were created, which defined population limits in much greater detail and took into account confounding factors such as age, ethnicity, sport, and altitude exposure (Gore et al. 2003; Parisotto et al. 2000a; Sharpe et al. 2002). Concurrently, the concept of using individualised rather than population-derived reference ranges, the *haematologic passport* concept, was being discussed as the future direction of anti-doping efforts (Ashenden 2002; Cazzola 2000; Malcovati et al. 2003). In 2006, Sharpe et al. (2006) published their 3G model for rHuEPO detection, so called because it built upon two previous population-based models. This refined model compared an athlete's blood test results to their own historical baseline rather than population-derived reference ranges. The results of the 3G model study supported the theory that a hematologic passport approach would be sensitive to doping. Unfortunately, a higher-than-expected incidence of false-positive results was recorded for the 3G model, leading the authors to conclude that further investigations were required before the approach could be implemented by anti-doping authorities (Sharpe et al. 2006). Subsequently, the Adaptive model was developed by scientists from a separate research team based at the Swiss Laboratory for Doping Analyses (Sottas et al. 2010). The Adaptive model differed from the

3G approach because, although primarily comparing athletes to individualised reference ranges, an athlete's past test results were also combined with population-derived estimates to create these ranges (Sottas et al. 2010). In December 2009, WADA officially endorsed the haematological module of the ABP to be used by international sports federations and national anti-doping organisations for detection of blood doping.

There are two methods of computation used by the Adaptive model for doping detection: first, each individual result is compared to the athlete's past test results, and, second, the variability of a complete sequence of results is compared to the variability in a control population.

For the analysis of individual test results, each time an athlete is requested to provide a blood sample their [Hb] and OFF-hr results are compared to the individualised range that is generated by the Adaptive model. Any results that fall outside this expected range are 'flagged' by the ABP software for the attention of anti-doping authorities. When the first result from an individual athlete is entered into the Adaptive model, the expected range is wholly determined by three population-derived values: the population mean, the between-subject variance and the within-subject variance. As more results are entered, the expected range becomes progressively less dependent on the population mean and between-subject variance, and more dependent on the within-subject variance in conjunction with the athlete's past test results (Sottas et al. 2010). The width of the expected range is determined by the magnitude of the within-subject variance of the blood marker and its associated probability-based confidence limits. The confidence limits act as a safety barrier for innocent athletes. WADA recommends the use of 99.9% confidence limits, which equate to a 1 in 1000 chance of a non-doped athlete's result falling outside the expected range (1 in 1000 chance of a false-positive) (World Anti-Doping Agency 2010). However, individual anti-doping organisations

may choose a lower limit at their own discretion (such as 99% limits, which equate to a 1 in 100 chance of a false-positive result). If a result falls outside of the expected range, the athlete's full blood profile ([Hb], %Ret, OFF-hr score and six other blood markers taken from the venous blood samples provided by the athlete) is reviewed by an expert panel.

Additionally, the variability of an athlete's entire sequence of results is compared against sequences obtained from athletes who are known to be either doped or non-doped (typically volunteers who have taken part in clinical studies). The variability in a sequence of blood results is higher in doped than non-doped athletes (Pottgiesser et al. 2012; Sottas et al. 2010). Like the procedure for individual results, an athlete's results are reviewed by an expert panel if the variability of their sequence exceeds the confidence limits of a non-doped population.

In direct methods of blood doping detection, such as the test for recombinant erythropoietin traces in blood or urine (Lasne 2001), an adverse finding leads directly to the athlete being sanctioned. However, a flagged result in the ABP is first reviewed by an expert panel. The expert panel reviews the results of all athletes whose individual results or sequence of results are flagged by the ABP software, and sanction the athlete unless they are satisfied that the changes could have resulted from innocent means (World Anti-Doping Agency 2010). For simplicity, throughout this thesis I have referred to an innocent athlete's result being flagged by the ABP as a 'false-positive'. In reality, the judgements of the expert panel are an important intermediary step between a result being flagged and the athlete being sanctioned.

Figure 2.2 is an ABP software screen shot showing the OFF-hr profile of a male subject who participated in a series of blood withdrawals and reinfusions during a recent autologous blood doping research study (Pottgiesser et al. 2011). In this case, 99% confidence limits were applied. The upper and lower lines on the graph represent the upper and lower limits of the

expected range of results that are progressively adjusted to incorporate the athlete's recent test history. The central line is the series of test results from the athlete. The phrase 'seq OFF-hr: prob= 99.96%' shows that the variability of the sequence of results is at the 99.96th percentile of the normal range. The fourth and sixth test results, and the sequence of results as a whole, would have been flagged by the ABP software for exceeding the 99% limits.

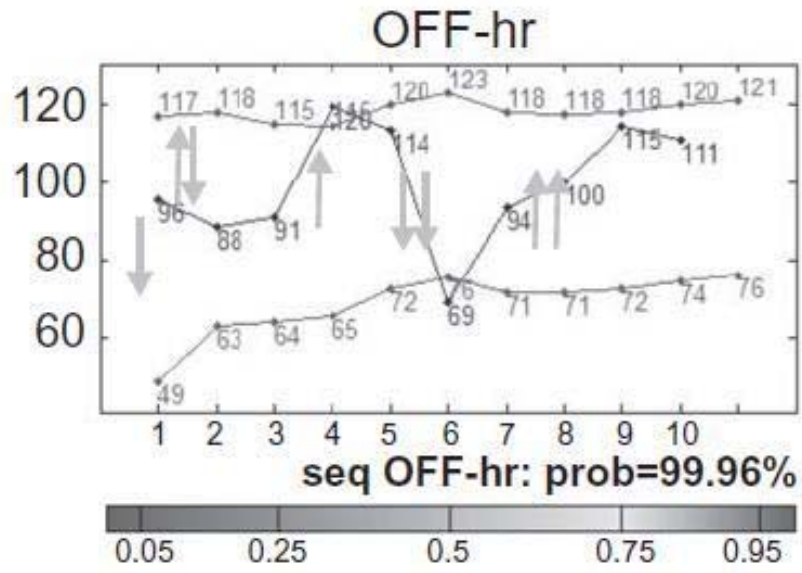

Figure 2.2: Screen shot from the ABP software of a male subject engaged in autologous blood doping. 99% limits of the expected range along with a series of ten test results and the variability of sequence of results are shown; individual results fell outside the expected range during 4th and 6th tests. "Arrow down" = withdrawal of 500 mL of whole blood; "Arrow up" = reinfusion of 280 mL of red blood cells. Figure adapted from Pottgiesser at al. (2011).

Practicalities of including Hb_{mass} in the ABP

Time efficiency

Total haemoglobin mass can be quantified indirectly using high purity carbon monoxide (CO) as a tracer substance for Hb, with the CO introduced via a re-breathing circuit. Carbon monoxide re-breathing methods for quantification of Hb_{mass} were first described in 1882 (Grehant 1882). Over the next ~100 years, continual improvements were made to make the method more accurate, reliable and time efficient (Burge and Skinner 1995; Myhre et al. 1968; Sjostrand 1948; Thomsen et al. 1991). In 2005, Schmidt and Prommer (2005) published a description of the 'Optimised CO re-breathing method', shortening the CO re-breathing period of previous methods from 15 minutes to two minutes and the total test time from 40 minutes to <15 minutes. The Optimised CO re-breathing method has repeatedly demonstrated good validity (Pottgiesser et al. 2007; Schmidt and Prommer 2005) and test-retest reliability (Eastwood et al. 2011b; Pottgiesser et al. 2009b; Schmidt and Prommer 2005). Therefore, in terms of time efficiency, Hb_{mass} measurement is now suitable for anti-doping purposes.

Quality control between laboratories

The sensitivity of the ABP is improved when tests are conducted as a mixture of 'In Competition' and 'Out of Competition' tests (World Anti-Doping Agency 2010). Given that modern elite competition schedules require athletes to travel worldwide, anti-doping agencies need to be able to conduct blood tests on athletes in a variety of different locations. This would require blood analyses to be conducted at the nearest of 33 WADA-accredited laboratories. As sequential ABP results for an individual athlete may originate from a number of different laboratories, the equivalency of all data must be ensured. For the existing ABP markers, this is achieved through the participation of all WADA-accredited laboratories in

designated internal and external quality control schemes for their haematology analysers (World Anti-Doping Agency 2009). One criticism that has been levelled at Hb_{mass} measurement as a potential ABP marker is the lack of quality control measures that currently exist between laboratories around the world (Sanchis-Gomar et al. 2010a; Schumacher and Pottgiesser 2010).

When Hb_{mass} is measured using the Optimised CO re-breathing method, the athlete's blood is measured for percent carboxyhaemoglobin (%HbCO) before and after the CO re-breathing period using a spectrophotometer (Schmidt and Prommer 2005). Current researchers use a variety of spectrophotometers. For example, numerous models from the manufacturer Radiometer (Copenhagen, Denmark) have been used: the OSM3 (Robertson et al. 2010b), ABL 520 (Schmidt and Prommer 2005), ABL700 (Lundby and Robach 2010) and ABL800-Flex (Steiner and Wehrlin 2010), as have two models manufactured by AVL Medizintechnik (Bad Homburg, Germany): the AVL Omni 5 (Ahlgrim et al. 2009) and AVL Omni 9 (Pottgiesser et al. 2007). In a recent study examining the feasibility of conducting longitudinal Hb_{mass} monitoring using two different spectrophotometers, the OSM3 (Radiometer, Copenhagen, Denmark) and the RapidLab 1245 (Siemens HealthCare Diagnostics, Sudbury, UK), Ulrich et al. (2011) reported significant differences between Hb_{mass} values calculated from these analysers. The authors concluded that one brand of spectrophotometer would need to be chosen for use in all laboratories if Hb_{mass} was to be included in the ABP, and that external quality control solutions would be required to ensure close agreement between analysers in different laboratories (Ulrich et al. 2011). No such external quality control solutions are currently available for %HbCO. The adoption of one specific model of spectrophotometer by all WADA-accredited laboratories would be relatively easy to implement, given that the Sysmex (Kobe, Japan) is the sole brand used by WADA

laboratories for haematological analyses. But it remains important to determine the extent to which external quality control solutions could reduce the variability of longitudinal Hb_{mass} measures conducted in different laboratories.

Sources of Hb_{mass} variation in non-doped athletes

In order for Hb_{mass} measurement to be sensitive to blood doping, the fluctuations of Hb_{mass} in non-doped athletes must be reliably distinguishable from the changes in Hb_{mass} in doped athletes. When athletes' results are entered into the ABP software, the Adaptive model defines a range in which subsequent results are expected to fall. The width of this expected range determines the limits of acceptable fluctuations of Hb_{mass} in non-doped athletes. It is, therefore, crucial that anti-doping authorities enforce the correct limits to maximise the sensitivity of the ABP to doping whilst maintaining a high level of specificity. After a few tests have been recorded for an athlete, the width of the expected range is primarily determined by the value of within-subject variance that has been entered into the Adaptive model. Consequently, the within-subject variance of Hb_{mass} in non-doped athletes must be well-defined and must be small enough that, once 99.9% confidence limits are applied, the results of doped athletes exceed the limits.

In the literature, the within-subject variance of Hb_{mass} has frequently been reported in the form of the within-subject standard deviation (SD). Mathematically, the SD is the square root of the variance. The within-subject SD of Hb_{mass} originates from both biological and analytical sources, which combine as follows:

$$Within\text{-}subject\ SD\ of\ Hb_{mass} = \sqrt{Analytical\ SD^2 + Biological\ SD^2}$$

(Fraser and Harris 1989; Gore et al. 2005)

Biological SD can be considered as the true change in Hb_{mass}, whilst analytical SD (otherwise known as error) originates from numerous aspects of the measurement method (such as, reliability of the spectrophotometer, operator's consistency of measuring the CO dose and purity of the CO) (Gore et al. 2005).

In the Adaptive model for the current haematological markers of the ABP ([Hb] and %Ret) a single value of within-subject variance is used, which incorporates both analytical and biological elements (Sottas et al. 2010). However, the Adaptive model for Hb_{mass} that has been included in the ABP software (for the purposes of exploration rather than implementation at this stage) requires separate estimates of analytical SD and biological variance of Hb_{mass} to be entered. The analytical component is the Typical Error (TE) for the method, expressed as a percentage. The TE is defined as the standard deviation of test results, after any shifts in the mean have been taken into account (Hopkins 2000). The biological component is the within-subject biological variance (BioWS variance), an absolute value, expressed in grams squared (g^2). The within-subject variance has been split in this way to better reflect the characteristics of a large sample of pilot data that has been examined by the scientists at the Swiss Laboratory for Doping Analyses, and to enable the TE value to be altered to accurately reflect the analytical SD recorded in individual laboratories (Personal communication, P.E. Sottas, 2012).

Currently, the analytical SD of Hb_{mass} is well-defined, but the effects of some key influences on the BioWS variance of Hb_{mass} in non-doped athletes remain unknown.

Analytical variability

The analytical SD of the Optimised CO re-breathing method has typically been estimated by researchers making two measurements of Hb_{mass} within a few days of each other in a group of

athletes. Over such a short time period, biological variation is likely negligible and therefore all variation is assumed to be analytical. In Schmidt and Prommer's (2005) original description of the Optimised CO re-breathing method, they calculated a TE of 1.7%. Although there have been a number of modifications made to the method with the aim of reducing analytical error (Alexander et al. 2011; Gore et al. 2006a; Prommer and Schmidt 2007), the TE of the method is typically reported to be between 1.5% and 2.0% in research conducted in Australia (Eastwood et al. 2011b; Garvican et al. 2011b), Switzerland (Steiner and Wehrlin 2010) and Germany (Pottgiesser et al. 2012), with occasional reports of studies demonstrating a smaller TE (Pottgiesser et al. 2009b; Prommer et al. 2008). With a TE <2%, the Optimised CO re-breathing method compares favourably against a number of other methods used for measuring blood volume (Gore et al. 2005), meaning that it would be an appropriate method to use in the ABP.

Normal BioWS variance in athletes

Four different researchers have quantified the within-subject variability of Hb_{mass} in an athlete population with the aim of differentiating non-doped from doped athletes (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008). Each of these analyses targeted an athletic population and included in their estimates normal influences that are part of the athlete lifestyle, such as training, off-season and competitive season differences, body mass changes and maturation.

Prommer et al. (2008) examined changes in Hb_{mass} in 24 semi-elite athletes and 6 leisure athletes over a period of 1 year, involving a total of 128 measures. The authors quantified the TE of their measures to be 1.4% and estimated a fixed 7.5g (0.8%) biological SD over the course of the year. This is equivalent to a BioWS variance of 56.25 g^2. However, using an

alternative method of calculation, the authors also calculated the within-subject variance of Hb_{mass} to be 408 g^2. By using this estimate along with the TE of 1.4%, the BioWS variance could, in fact, have been 244 g^2 (1.7%) (see Appendix 1 for details of calculations). It is not clear which of these estimates, 56.25 g^2 or 244 g^2, best represented the biological variation within the group. These authors noted a large difference between the mean Hb_{mass} of the athlete depending on their competitive level, but found no significant effect of competitive season or training on Hb_{mass}. Of note was the decision of these authors to exclude the data of any athletes who experienced a long-lasting injury, or altitude sojourns, meaning that the effect of these two factors, which can frequently occur in the lives of elite athletes, were not quantified.

Instead of excluding data from athletes associated with altitude exposure, Eastwood et al. (2011b) estimated this effect separately, as well as the effect of reduced training on Hb_{mass}. Without the influence of these extraneous factors, their estimates of within-subject SD were 3.4% for males and 4.0% for females in semi-elite and elite athletes. In this study, analytical SD was estimated as ~2.0%, meaning that biological SD was equal to 2.8% (830 g^2) in males and 3.5% (573 g^2) in females, much larger estimates than those of Prommer et al. (2008). The sample size of the Eastwood et al. (2011b) study was much larger than that of the other studies mentioned here; they tracked 130 athletes over the course of about one year, ~900 Hb_{mass} measures in total. It is also the only study to estimate the variability of Hb_{mass} in males and females separately. The authors suggested, however, that the gender difference may be due to sampling variation rather than a true difference between males and females.

Pottgiesser et al. (2012) reported an analytical SD of 1.7% and a BioWS variance of 550 g^2 (equivalent to a biological SD of 2.6%) using Hb_{mass} results from 10 male recreational athletes

over a 1 year period. The description by these authors of their subject group was relatively minimal, although they did explain that their only altitude exposure would have been occasional sojourns (< 7 days) for skiing.

Morkeberg et al. (2011) estimated the within-subject variance of Hb_{mass} to be 0.001 when expressed as a natural logarithm. At a higher level of precision, the estimate was, in fact, 0.000912 (Personal communication, J. Morkeberg, 2012). This is equivalent to a within-subject SD of 3.07%. The participants were 58 male elite athletes whose Hb_{mass} was measured ~3 times over a 24 month period. These authors did not divide this estimate into its analytical and biological components. However, they did report a TE of 2.0%, which means that biological SD would have been ~2.3% ($\sqrt{(3.07)^2 - (2.0)^2}$). In this population, where the mean Hb_{mass} was 1064 g, this equates to a BioWS variance of 611 g^2.

In summary, estimates of BioWS variance of Hb_{mass} have ranged from 56.25 g^2 to 830 g^2 (~0.8% to 3.5%) when measured in athletes. These estimates do not include the influence of a number of confounding variables (such as, the acute effect of exercise, illness, injury, detraining, and hypoxia) that may contribute additional BioWS variance in non-doped athletes.

Additional BioWS variance: Acute effect of exercise

The measurement of [Hb] is susceptible to variation in the case of recent exercise due to exercise-induced plasma volume concentration (Harrison 1985; Schumacher et al. 2010). Accordingly, the existing ABP guidelines stipulate that blood tests must not be conducted within 2 hours of physical exertion, in order to control for this effect (World Anti-Doping Agency 2010). One potential advantage of including Hb_{mass} in the ABP is that it is not

affected by changes in plasma volume, meaning that measurements could, theoretically, be made immediately after exercise (Sottas et al. 2010).

However, although Hb_{mass} is independent of plasma volume, the reliability of the CO re-breathing method is heavily influenced by the consistency of the rate of whole-body circulation (Garvican et al. 2010a). The effect of cycle stage-racing on daily measures of Hb_{mass} in elite Under-23 (U23) male cyclists has been examined by two different researchers (Garvican et al. 2010b; Schumacher et al. 2008c). Schumacher et al. (2008c) found increased variability of Hb_{mass} in the cyclists when measured daily following cycling during a 5-day stage race. In contrast, Garvican et al. (2010b) found no difference in the daily variability in Hb_{mass} between the cyclists and a matched control group over a 6-day period of cycle stage racing. The equivocal nature of these findings suggests that further investigation into the reliability of Hb_{mass} measures made shortly after racing is required.

Furthermore, it is important to extend the investigation to sports other than cycling as the ABP is applied to athletes from a variety of different sports. One factor that may influence Hb_{mass} specifically after running events is the destruction of red blood cells (intravascular haemolysis) commonly associated with footstrike (Miller et al. 1988; O'Toole et al. 1988; Telford et al. 2003). Following intravascular haemolysis, freely-circulating Hb binds to the protein haptoglobin in the blood until the haptoglobin stores are saturated. Haptoglobin-Hb complexes are then broken down in the liver and any remaining free Hb is excreted in the urine (McDonald 1984). The destruction of red blood cells in this manner, and resulting excretion of Hb may reduce Hb_{mass} for a period following long running events.

It is also possible that exercise-induced contraction of the spleen, which results in additional red cells being released into the circulation at times of physiological stress (Stewart and

McKenzie 2002), may cause a transient increase in Hb_{mass} shortly after exercise (Schmidt and Prommer 2010). However, an investigation into the effects of splenic contraction on Hb_{mass} in competitive apnea divers found no measurable difference in Hb_{mass} before and after diving, despite evidence of spleen contraction (Prommer et al. 2007a).

The acute effects of exercise on Hb_{mass}, other than cycling (Garvican et al. 2010b; Schumacher et al. 2008c), have not previously been examined. Given the reliance of Hb_{mass} measures on consistent rates of circulation and the potential influences of intravascular haemolysis and splenic contraction, the stability of Hb_{mass} following events involving running requires investigation. Research into this area will inform anti-doping authorities about the practicality of using Hb_{mass} shortly after racing or hard training.

Additional BioWS variance: Injury and illness

The only investigations to have quantified the effect of injury on Hb_{mass} are two case studies (Kjellberg et al. 1949; Schumacher et al. 2008b), but even with this limited evidence it is clear that injury has the potential to cause changes in Hb_{mass} that greatly exceed the normal variation of Hb_{mass} in non-doped athletes. Kjellberg (1949) reported a 15% decrease in Hb_{mass} in a well-trained female athlete following one and a half months of bed rest due to a badly healing leg fracture. Similarly, Schumacher et al. (2008b) reported a 19% decrease in Hb_{mass} in an elite female cyclist following four weeks of bed rest subsequent to a fracture of the patella. In the latter study, the authors estimated that 5% of the Hb_{mass} decrease was due to blood loss associated with the accident and surgery, with the remaining 14% presumed to be due to inactivity following the accident (Schumacher et al. 2008b).

The only illness to have been examined in detail with regards to its effect on Hb_{mass} is iron-deficiency, with the effect of iron supplementation on Hb_{mass} having been examined in three

separate studies. A remarkable 77% (~300 g) increase of Hb_{mass} was recently reported following 7 weeks of iron supplementation (intramuscular injection and oral supplementation) in a female endurance runner (Garvican et al. 2011a). This athlete had formally been diagnosed with anaemia and prior to treatment had a [Hb] of 88 $g.L^{-1}$. However, in athletes who have a low serum ferritin but are non-anaemic (normal [Hb]), both oral administration and intramuscular injection of iron have been shown to have no significant effect on Hb_{mass} (Ashenden et al. 1998; Friedmann et al. 2001). Therefore, it appears that iron supplementation may not disturb the normal variability of Hb_{mass} in non-doped athletes unless the athlete is clinically anaemic.

Considering the high incidence of injury in athletes (Gabbett and Ullah 2012) and the compelling, but limited, evidence of large Hb_{mass} changes in two injured athletes (Kjellberg et al. 1949; Schumacher et al. 2008b), it is important for the effect of injury on Hb_{mass} to be examined in more detail. Furthermore, the effect of illnesses, other than iron-deficiency, on Hb_{mass} clearly needs further investigation and quantification.

Additional BioWS variance: Training and detraining

The typical Hb_{mass} of trained endurance athletes is significantly higher than that of untrained individuals (Heinicke et al. 2001; Prommer et al. 2008), which has raised many questions about the influence of training on Hb_{mass}. Although this difference may imply that training has increased athletes' levels of Hb above those of untrained individuals, the description of some untrained individuals already possessing a high Hb_{mass} and a high VO_{2max} (Martino et al. 2002) suggests that an individual's Hb_{mass} has a genetically determined component (Schmidt and Prommer 2010).

A number of longitudinal studies have been conducted using untrained or moderately trained populations in an attempt to investigate the trainability of Hb_{mass}, but the results have been equivocal. In previously untrained adults, neither Eastwood et al. (2011a) nor Green et al. (1991) recorded a significant effect of a 6-8 week training program on Hb_{mass} or red cell mass, respectively. In contrast, Sawka et al. (1992) inferred an 8-10% increase in red cell mass following just three weeks of training in recreational athletes. This dramatic change over such a short period runs counter to the results of Eastwood et al. (2011a) and Green et al. (1991) and it has been suggested that this unlikely result could be a consequence of measuring red cell mass using the Evans Blue dye method, which has a TE of ~7% (Eastwood et al. 2011a; Gore et al. 2005; Schmidt and Prommer 2010). However, Schmidt and Prommer (2010) reported that following a much longer 9-month marathon training program, moderately trained runners increased Hb_{mass} by 6.4%. This suggests that a longer period of training may be necessary before a change is discernable. These mixed results make it difficult to come to a definitive conclusion about the influence of training on Hb_{mass} in untrained or moderately trained athletes.

There is, however, quite consistent evidence to suggest that in well-trained athletes, deviations from their normal high levels of training can influence Hb_{mass}; the extent to which Hb_{mass} decreases is related to the extent and duration of training reduction. In a year-long observation of the variation in athletes' Hb_{mass}, Eastwood et al. (2011b) calculated that when athletes self-reported that they had experienced a period of reduced training, Hb_{mass} was reduced by 2.8%. In this study the extent and duration of training reduction was not well described, but in another study by the same authors (Eastwood et al. 2012) they recorded a 3.1% decrease in Hb_{mass} when training was reduced by 87% during a one-month training break in nine ultra-endurance triathletes. In an eight-month observation of Hb_{mass} in elite

female cyclists, accompanied by a thorough description of training, Garvican et al. (2010c) found that a 10% change in training load over a 6-week period caused a 1% change in Hb_{mass} (a reduction in training led to a reduction in Hb_{mass} and vice versa). Using that relationship, the authors estimated that a 100% decrease in training load over a 6-week period would cause a 11.4% decrease in Hb_{mass}. This is a relatively close estimate to the 14% inactivity-related Hb_{mass} decrease evident in the injured female cyclist that was the subject of the case study by Schumacher et al. (2008b). In contrast, Prommer et al. (2008) also followed athletes over the course of a year but found no effect of increased or decreased training on Hb_{mass}. An explanation for the difference between this study and that of Garvican et al. (2010c) could be in the way training was analysed. Prommer et al. (2008) divided training into three categories and compared changes in Hb_{mass} according to volume alone (<10h, 10-20h, >20h per week). Garvican et al. (2010c) calculated training stress scores, which took into account both volume and intensity of training. They also conducted linear regression analyses using a continuous training scale rather than dividing the training into discreet categories. It is possible that this more complex analysis teased out the subtle relationship between small changes in training and Hb_{mass}. A description of the dose-response relationship between training and Hb_{mass} in elite athletes would be useful to inform the judgements of the ABP's expert panel in cases where an athlete asserts that an unusual Hb_{mass} result is a consequence of a training break or injury/illness-related detraining.

Additional BioWS variance: Hypoxia

When humans are exposed to an hypoxic environment for a prolonged period, the body responds by increasing the number of circulating red blood cells in order to ensure an adequate oxygen supply to respiring tissues (Berglund 1992). This response is mediated by activation of the hypoxia-inducible factor-1 (HIF-1) pathway and the resultant up-regulation

of the erythropoietin (EPO) gene. The end results of these physiological changes are increased serum EPO concentration and increased production of red cells in the bone marrow (Hopfl et al. 2003). Many researchers have investigated the effects of hypoxia on sports performance and the general consensus is that the practice contributes small but worthwhile performance benefits (for reviews, see Bonetti and Hopkins (2009) and Wilber (2007)). Most scientists also agree that given adequate hypoxic stimulus, and as long as the athlete is in an adaptive state (i.e. healthy, uninjured, and receiving good nutrition), Hb_{mass} is increased in athletes following hypoxic exposure (Rusko et al. 2004; Saunders et al. 2009a).

The classical form of altitude training (Classic) involves athletes living and training at a location greater than ~2000m above sea level. Live High:Train Low (LHTL) training, where athletes live at moderate altitude but train at a lower altitude, was first proposed by Levine and Stray-Gundersen (1997) as a way of circumventing the limitations on training intensity encountered by athletes during Classic altitude training. The hypoxic environment needed for LHTL altitude training can be either natural or simulated (e.g. altitude houses or altitude tents) (Wilber 2007). In recent years, various modern derivations of altitude training have been described that involve shorter periods of daily exposure to more severe forms of hypoxia, either at rest or whilst training (Millet et al. 2010). These methods have been developed with the aim of capitalising upon the physiological benefits associated with hypoxia whilst minimising the time commitment and financial costs.

For Classic altitude training, the minimum hypoxic 'dose' necessary for yielding an increase in Hb_{mass} has been described as >2000-2200 m for at least 3 weeks, and for LHTL altitude training it is >12 $h.day^{-1}$ for at least 3 weeks at an altitude between 2100 and 2500 m (Rusko et al. 2004). The collective findings of numerous studies examining this topic, although not

unanimous (Siebenmann et al. 2012), suggest that both Classic and LHTL altitude training camps lasting three or four weeks typically yield a 1.5-4% mean increase in athletes' Hb_{mass} (Bonetti and Hopkins 2009; Clark et al. 2009; Eastwood et al. 2011b; Garvican et al. 2012). Some studies have reported larger (5-9%) increases in group mean Hb_{mass} for both Classic (Friedmann et al. 2005; Heinicke et al. 2005) and LHTL (Robach et al. 2006; Wehrlin et al. 2006). In individual athletes, Hb_{mass} increases as large as 20% have been reported following Classic altitude training (Friedmann et al. 2005; Heinicke et al. 2005), although it has been suggested that these substantial changes may be a result of measurement error (Gore and Hahn 2005). However, it is common for at least one or two individual athletes in every altitude training study to experience an increase in Hb_{mass} >5% (Garvican et al. 2012; Garvican et al. 2011b; Pottgiesser et al. 2009a; Saunders et al. 2010a).

There have, as yet, been no studies conducted that have directly compared the changes in Hb_{mass} induced by Classic and LHTL altitude training. Although the hormonal response to both of these forms of altitude training has been reported to be similar (Koistinen et al. 2000), in a review of the literature Saunders et al. (2009a) estimated that 100 hours of hypoxic exposure would lead to a 3% Hb_{mass} increase in Classic but only a 1% increase in LHTL. This suggestion of a higher increase in Hb_{mass} following Classic altitude is supported by the findings of Levine and Stray-Gundersen (1997), who reported red cell mass increases of 10.5% and 5.3% after Classic and LHTL altitude training, respectively. However, these changes in red cell mass were measured using the Evans Blue dye method, which is less reliable than the CO re-breathing method for measuring small changes in blood volume (Gore et al. 2005). Consequently, it cannot be assumed that the same pattern would be observed if Hb_{mass} changes following Classic and LHTL altitude training were compared using this latter method.

Acute Intermittent Hypoxic Exposure (AcIHE) is a modern derivation of altitude training that has yielded substantial performance benefits in sub-elite athletes, but is relatively untested in elite athletes (Bonetti and Hopkins 2009). AcIHE involves ~three weeks of daily 60-90 minute intermittent exposures to hypoxia, typically simulating altitudes of between 3500 and 6000 m, for at least three weeks. Thus far, no research studies have measured changes in Hb_{mass} following AcIHE, although authors who have found increased [Hb] (Bonetti et al. 2006; Bonetti et al. 2009; Burtscher et al. 2010) have speculated that an increase in Hb_{mass} may have occurred. However, Gore et al. (2006b) demonstrated that no significant Hb_{mass} increase resulted from four weeks of 3 $h.day^{-1}$ continuous hypoxia equivalent to 5000m, which is likely to be a higher hypoxic 'dose' than the daily 60-90 minute intermittent exposures of AcIHE. Furthermore, since at least 90 minutes of continuous exposure to an altitude equivalent to 5500m is required for serum EPO to be increased (Knaupp et al. 1992; Rodriguez et al. 2000), it is unlikely that AcIHE constitutes a sufficient hypoxic dose to stimulate erythropoiesis.

Hypoxic exposure has long been acknowledged as a complication in the detection of blood doping using indirect blood models (Ashenden et al. 2004; Sharpe et al. 2002). When an athlete submits a blood sample for ABP analysis, the athlete declares any recent hypoxic exposure to the doping control officer (World Anti-Doping Agency 2010). These details of recent exposure are documented so that an adjustment can be made to the predicted [Hb] and %Ret value for that athlete within the Adaptive model (Swiss Laboratory for Doping Analyses 2009). The records may also be taken into account by the expert panel in the event of an athlete's result being flagged as abnormal. The experts would then be required to decide whether the fluctuation in an athlete's blood could be attributable to the hypoxia rather than doping. If Hb_{mass} is to be included in the ABP, it is vital that the extent of Hb_{mass} deviations

resulting from different forms of altitude training is defined. These deviations may be included as allowances in the Adaptive model for Hb_{mass}, similar to the way allowances are currently made for [Hb] and %Ret. Alternatively, the expert panel may simply use their knowledge of hypoxia-induced changes in Hb_{mass} to assist with their adjudication of flagged results.

Sources of Hb_{mass} variation in doped athletes

Indirect models of blood doping detection, such as the Adaptive model, play an important role in combating the illegal use of blood transfusions and ESAs (Lundby et al. 2012). The implementation of direct tests for rHuEPO and homologous doping in 2001 and 2004, respectively, stimulated resurgence in the popularity of autologous transfusions (Lippi and Banfi 2006). Documents retrieved from the 'Operation Puerto' doping scandal in 2006 revealed a complex schedule of blood withdrawal and reinfusion being followed by top cyclists (Lundby et al. 2012). These included reports of acute autologous doping, the practice of infusing blood shortly before a race and removing it immediately afterwards. Despite a great deal of research investment, to date, there remains no direct test for detecting autologous transfusions. Consequently, the ABP is seen as a valuable tool in the fight against autologous doping. It would be very beneficial if Hb_{mass} models were sensitive to this form of doping. Although a direct test now exists for rHuEPO, it has been suggested that athletes have learned to titrate their dosages of the hormone to reduce the window of detection from 3 days post-injection to as little as 12-18 hours (Ashenden et al. 2006; Nissen-Lie et al. 2004). The effects of rHuEPO on the athlete's blood profile are visible for a much longer period, making Hb_{mass} a potentially valuable tool in the detection of these smaller 'microdoses' of rHuEPO.

Blood transfusion

The Hb content of 1 unit (450 mL) of blood is 48-60 g (Morkeberg et al. 2011; Pottgiesser et al. 2009b), of 2 units is ~108 g (Pottgiesser et al. 2007) and of 3 units is ~135 g (Morkeberg et al. 2011). If the red blood cells are subsequently stored by freezing, the Hb content of blood is reduced from 48 g to 42 g in 1 unit and from 135 g to 86 g in 3 units (Morkeberg et al. 2011). The precision with which changes of Hb_{mass} can be measured using the Optimised CO re-breathing technique has been shown to be within 20 g of the actual change; increases of 70 g and 90 g were measured after reinfusion of 1 and 2 units, respectively (Pottgiesser et al. 2007). Assuming a Hb_{mass} of ~1000 g in an elite male endurance athlete (Heinicke et al. 2001), these 70 g and 90 g increases are equivalent to Hb_{mass} changes of 7% to 9%, although the change may be as little as 4.2% based on reinfusion of 1 unit of frozen blood containing ~42 g of Hb.

Using Hb_{mass} in the ABP would theoretically present two opportunities for detection of autologous doping: for a period of time after blood withdrawal and for a period after reinfusion. After donation of 1 unit of blood, there is a 36-day window when Hb_{mass} is reduced (Pottgiesser et al. 2008) until the body completely replenishes red cell volume through a period of accelerated erythropoiesis. However, it is only for the first 12-days of the 36-day period that this reduction in Hb_{mass} is distinguishable from the normal analytical variation of Hb_{mass} (i.e. >4.4%) (Prommer et al. 2007b). In a separate experiment, it was demonstrated that for 28 days following reinfusion of 1 unit of blood, Hb_{mass} was ~6% higher than baseline, before a marked drop in Hb_{mass} would render the transfusion again indistinguishable from the normal analytical variation of Hb_{mass} (Pottgiesser et al. 2009b). From the perspective of anti-doping authorities, both large decreases and large increases in Hb_{mass} would be suspicious and may be indicative of blood doping. However, the timing of

testing is crucial for successful detection and the window of opportunity may be shorter following blood withdrawal (~12 days) than following reinfusion (~28 days). In the case of acute autologous blood doping, timing is even more critical with tests needing to be conducted immediately after racing for there to be any hope that the doping may be detected.

rHuEPO doping

Changes in Hb_{mass} of a similar magnitude to autologous blood doping have been reported following rHuEPO doping. Group mean Hb_{mass} increases of 6.9% and 12.0% resulted from 25 days of rHuEPO administration (150 $IU.kg.wk^{-1}$) together with intramuscular or oral iron supplementation, respectively (Parisotto et al. 2000a). However, since the introduction of a direct test for the detection of rHuEPO in urine (Lasne 2001), it has been reported that athletes have begun using smaller 'microdoses' of rHuEPO in an effort to evade detection (Ashenden et al. 2006). Ashenden (2006) demonstrated that following 1 month of a microdosing regime (~75 $IU.kg.wk^{-1}$) in two subjects, [Hb] was consistently maintained at doped levels whilst the detection of the rHuEPO in the urine was no longer possible 24 hours after the injection.

There is one research study that has examined changes in Hb_{mass} resulting from rHuEPO microdosing (Lundby and Robach 2010). In this study, the dosage regimen consisted of a 'boosting' period of three weeks when three 5000 IU injections were given per week (~180 $IU.kg.wk^{-1}$, although dosage was not individualised according to body mass), followed by 11 weeks of microdosing consisting of one 5000 IU injection per week (~60 $IU.kg.wk^{-1}$, although again not individualised). The changes in Hb_{mass} from baseline varied substantially between subjects (Figure 2.3). In general, Hb_{mass} was increased by ~10% after the boosting period. Subsequently, the microdosing phase could be interpreted as being successful in three

subjects whose Hb_{mass} was maintained at ~7-10% above baseline throughout, and a failure in four subjects whose Hb_{mass} returned to just ~4% above baseline. One athlete had a vastly different response, their Hb_{mass} having increased by only 1.3% after the boosting stage but increasing to ~20% above baseline after the microdose period. These authors did not put forward any explanation as to why there may have been such a varied Hb_{mass} response to the doping regimen. One possibility was the researchers' decision to give a uniform dose of rHuEPO rather than individualised treatment according to body mass, which would have resulted in a relatively higher dose in some athletes than others. However, on close examination of the data, there appears to be no relationship between the Hb_{mass} response and body mass of the subjects. Therefore, the Hb_{mass} response to rHuEPO doping appears to be highly individual.

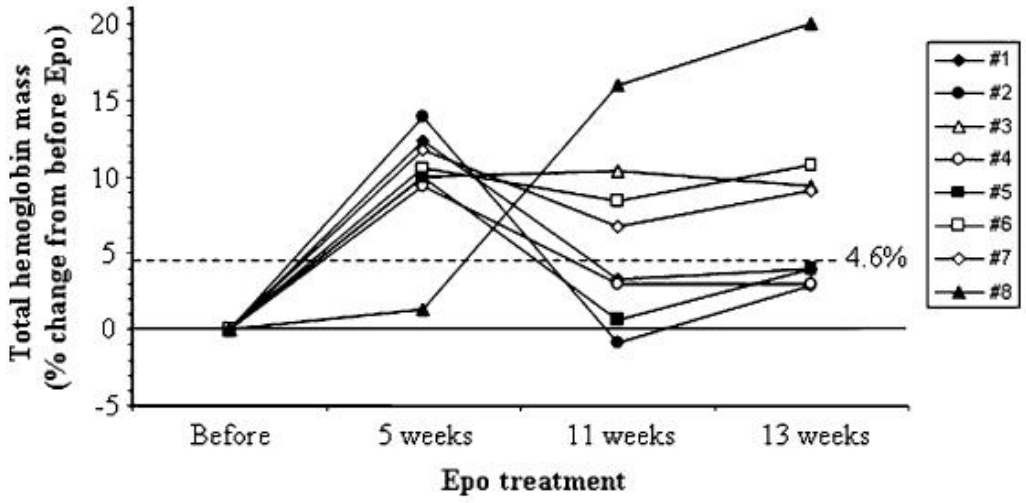

Figure 2.3: Percent change in Hb_{mass} after 5, 11 and 13 weeks of rHuEPO. From Lundby and Robach (2010).

Given the variable nature of individuals' responses to rHuEPO doping, it is likely that athletes would adopt an intelligent approach to doping. This may involve the modification of the dosages according to feedback from blood test results with the help of doctors or scientists. The study conducted by Lundby and Robach (2010) did quantify the effect of one schedule of rHuEPO microdosing on Hb_{mass}. However, the likelihood of athletes using that protocol is minimal due to the high disturbances in Hb_{mass} (and likely other markers) that occurred following the boosting period, with the associated increased risk of detection. It is more likely that an athlete would use a very low dose of rHuEPO, progressing to higher doses slowly if feedback from their blood values suggested they could do so without risk of detection. Such a protocol may result in small, but worthwhile, changes in Hb_{mass} that blur the boundaries between doping and non-doping. What is yet to be determined is whether including Hb_{mass} in the ABP would offer any additional sensitivity, above that of the current ABP markers, to realistic protocols of rHuEPO doping.

ABP models incorporating Hb_{mass}

Two Hb_{mass} models have already been incorporated into the ABP software by the scientists from the Swiss Laboratory for Doping Analyses, although these models are not currently in use by anti-doping organisations. One of these models uses Hb_{mass} as a single marker, and the other combines Hb_{mass} with %Ret. So far, there have been two research studies that have examined the utility of Hb_{mass} models for detection of autologous doping.

Models using Hb_{mass} as a single marker

Morkeberg et al. (2011) were the first researchers to examine the sensitivity and specificity of an Adaptive model based on Hb_{mass}. Twenty-nine recreationally active males were transfused with either one (n=8) or three (n=21) units of autologous blood and their Hb_{mass} and venous

blood markers were monitored for the four weeks following reinfusion. Overall, the sensitivity of the Hb_{mass} model was equivalent to the existing OFF-hr model (20%) but Hb_{mass} was only able to detect doping of larger volumes of blood. However, in the first five days after reinfusion, the Hb_{mass} model was far superior to the existing ABP models (40% versus 15%), highlighting the potential for Hb_{mass} to improve detection of acute autologous doping. The model yielded one false-positive result when the Hb_{mass} results of 60 German athletes were analysed, although the authors expressed doubts about the innocence of the athlete in question and considered that the result may, in fact, have been a 'true' positive.

Recently, Pottgiesser et al. (2012) also examined the sensitivity of Hb_{mass} to autologous doping. However, they designed their study with real-world anti-doping procedures in mind, blinding one senior investigator to the group allocation (doped or control) and asking them to play the role of 'doping control investigator'. The doping control investigator could order blood tests to be made at specific time points, up to a maximum of 10 times in each experimental subject, over a one year period. By comparing the results of two publications originating from this same study (Pottgiesser et al. 2012; Pottgiesser et al. 2011) it is clear that including Hb_{mass} as an additional marker in the ABP would improve the ABP's sensitivity to autologous doping. Using the existing ABP markers ([Hb] and OFF-hr score) or Hb_{mass} alone yielded identical sensitivity, with 6 out of 11 doped subjects exceeding the 99.9% confidence limits. However, there were two subjects who were flagged by Hb_{mass} but not by the existing markers, and vice versa. Consequently, using all of these markers in the ABP would have increased the sensitivity to 8 out of 11 subjects. Furthermore, no false-positive results from Hb_{mass} were reported in this study at WADA's recommended 99.9% level of specificity. This suggests that the specificity of Hb_{mass} is also high, making it suitable for inclusion in the ABP.

Models combining Hb_{mass} with %Ret

In the anti-doping setting, it has previously been demonstrated that multi-parameter indices (e.g. OFF-hr Score) can be more sensitive than single parameters alone, because the combination of unusual changes in multiple parameters can be a stronger indicator of blood doping (Parisotto et al. 2000a). Both Morkeberg et al. (2011) and Pottgiesser et al. (2012) demonstrated an additional advantage of combining Hb_{mass} and %Ret into a single model (termed the 'Hbmr' model by Morkeberg et al. (2011) and the 'OFF-mass' model by Pottgiesser et al. (2012)). The 'Hbmr' model was more sensitive to infusion of a large volume of blood than models using [Hb], OFF-hr or Hb_{mass} alone (Morkeberg et al. 2011) and the OFF-mass model allowed the detection of two additional doped subjects in the study of Pottgiesser et al. (2012). There were no false-positive results using either of the combined Hb_{mass} and %Ret models.

All together, these studies suggest that the inclusion of Hb_{mass} as an extra marker in the ABP could improve the sensitivity of the ABP to autologous doping. Additionally, combining Hb_{mass} with %Ret could further improve the ABP's sensitivity. However, further testing of these models is required before they could be implemented as part of the ABP.

There have been no studies conducted thus far that have assessed the sensitivity of Hb_{mass} models to rHuEPO doping. Lundby et al. (2010) dismissed the idea of Hb_{mass} being capable of detecting rHuEPO doping after they measured small doping-induced Hb_{mass} changes in half of their experimental subjects (Figure 2.3). However, they did not formally assess the sensitivity of the Adaptive model, instead drawing an arbitrary line at the level of a 4.6% increase of Hb_{mass} to represent detection success or failure. This interpretation is too simplistic and should

not be considered to represent the true sensitivity of Hb_{mass} to rHuEPO doping if it were to be included in the ABP.

Another aspect that has not, as yet, been comprehensively examined is the specificity of Hb_{mass} models in non-doped populations. In the studies of Morkeberg et al. (2011) and Pottgiesser at al. (2012) the specificity of the Hb_{mass}, Hbmr and OFF-mass models were close to 100%, with only one doubtful possible false-positive result reported by Morkeberg et al. (2011). However, all models that are considered for inclusion in the ABP would need to be examined for their specificity in a population that includes the possible confounding factors of acute exercise, injury, illness, detraining and hypoxic exposure. These factors represent the most likely causes of the Hb_{mass} results of non-doped athletes being flagged by the ABP, and are therefore an important test case to ensure that Hb_{mass} models do not compromise the specificity of the ABP as a whole.

Finally, the decision about what values of TE and BioWS variance should be entered into the Adaptive model for Hb_{mass} will play a key role in determining the model's sensitivity and specificity. Morkeberg et al. (2011) and Pottgiesser at al. (2012) each included their own estimates of TE and BioWS variance and achieved high specificities and sensitivities. It is, perhaps, not coincidental that their estimates of BioWS variance were very similar (2.5%-2.6%) and this value may represent a suitable estimate to be included universally. However, the estimates of BioWS variance of Prommer et al. (2008) were smaller (0.8%-1.7%) relative to those of Eastwood et al.(2011b) (2.8%-3.5%). It is yet to be determined which of these several estimates would provide an appropriate balance between sensitivity and specificity in a large and diverse athlete population if Hb_{mass} models were included in the ABP.

Summary

Although the inclusion of Hb_{mass} in the ABP may enhance the sensitivity of the ABP to blood doping, a review of the literature has highlighted the need for various investigations to be conducted before any final decision can be made about the suitability of Hb_{mass} for anti-doping purposes.

For the current markers of the ABP, the equivalency of results measured in different laboratories is ensured by the participation of all WADA-accredited laboratories in internal and external quality assurance schemes. Although various research groups from around the world have each demonstrated that they are capable of measuring Hb_{mass} in athletes with equivalent reliability, currently no such quality control procedures exist for Hb_{mass}. This is one limitation that would need to be overcome before Hb_{mass} could be included as a marker in the ABP.

The magnitude of normal day-to-day fluctuations in Hb_{mass} in non-doped athletes is markedly smaller than the Hb_{mass} changes that result from blood doping. However, there is evidence within the literature to suggest that some common influences in the lives of athletes may blur the line between doped and non-doped athletes. Specifically, increased fluctuations in Hb_{mass} may result from the acute effect of exercise, illness, injury, detraining and hypoxia. Conversely, the use of small microdoses of rHuEPO by doped athletes may induce small but beneficial increases in Hb_{mass}, which are indistinguishable from the fluctuations in non-doped athletes. The effects on Hb_{mass} of these confounding variables have not, as yet, been examined in sufficient detail.

Finally, initial investigations into the sensitivity and specificity of Hb_{mass} models included in the ABP have yielded very promising results. Their findings suggest a high specificity in non-doped athletes and additional sensitivity over the current markers of the ABP for the detection of autologous blood doping. However, these models have not been tested in a population of athletes where the line between doped and non-doped has been blurred by the influences of exercise, illness, injury, detraining, hypoxia and microdose rHuEPO. One important consideration to help optimise sensitivity and specificity of Hb_{mass} models in a situation where the difference between doped and non-doped athletes is small is an appropriate choice of the within-subject variance to be included in the model. In the literature, although estimates of the TE recorded for Hb_{mass} are relatively consistent, estimates of BioWS variance vary a great deal. By choosing to include a smaller estimate of BioWS variance, anti-doping authorities may record higher levels of sensitivity for the Hb_{mass} model but could sacrifice specificity. Conversely, choosing a larger estimate of BioWS variance may allow a higher specificity but could result in a low sensitivity. To date, no investigations have addressed the issue of which estimate of BioWS variance affords the optimum balance between sensitivity and specificity for Hb_{mass} models in the ABP.

Chapter 3

Quality control technique to reduce variability of longitudinal measurement of haemoglobin mass

Journal article accepted for publication in the

Scandinavian Journal of Medicine and Science in Sports 2011; 21: e365-371

Gough CE, Sharpe K, Ashenden MJ, Anson JM, Saunders PU,

Garvican LA, Bonetti DL, Gore CJ and Prommer N.

Presented here in the journal submission format

UNIVERSITY OF CANBERRA
AUSTRALIA'S CAPITAL UNIVERSITY

Chapter 3

This chapter has been removed due to copyright restrictions.

This chapter is available as:

Gough, C. E., Sharpe, K., Ashenden, M. J., Anson, J. M., Saunders, P. U., Garvican, L. A., Bonetti, D. L., Gore, C. J. & Prommer, N. (2011) Quality control technique to reduce the variability of longitudinal measurement of hemoglobin mass. *Scandinavian Journal of Medicine & Science in Sports*. 21(6): e365-e371.

Links to this chapter:

Print	http://webpac.canberra.edu.au/record=b1683904~S4
Online subscribed content (UC community)	http://ezproxy.canberra.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=67510655
Online general public	http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0838.2011.01316.x/abstract
DOI	10.1111/j.1600-0838.2011.01316.x

Abstract

The sensitivity of the athlete blood passport to detect blood doping may be improved by the inclusion of total hemoglobin mass (Hbmass), but the comparability of Hbmass from different laboratories is unknown. To optimize detection sensitivity, the analytical variability associated with Hbmass measurement must be minimized. The aim of this study was to investigate the efficacy of using quality controls to minimize the variation in Hbmass between laboratories. Three simulated laboratories were set up in one location. Nine participants completed three carbon monoxide (CO) re-breathing tests in each laboratory. One participant completed two CO re-breathing tests in each laboratory. Simultaneously, quality controls containing Low (1–3%) and High (8–11%) concentrations of percent carboxyhemoglobin (%HbCO) were measured to compare hemoximeters in each laboratory. Linear mixed modeling was used to estimate the within-subject variation in Hbmass, expressed as the coefficient of variation, and to estimate the effect of different laboratories. The analytic variation of Hbmass was 2.4% when tests were conducted in different laboratories, which reduced to 1.6% when the model accounted for between-laboratory differences. Adjustment of Hbmass values using quality controls achieved a comparable analytic variation of 1.7%. The majority of between-laboratory variation in Hbmass originated from the difference between hemoximeters, which could be eliminated using appropriate quality controls.

Chapter 4

Spurious Hb mass increases following exercise

Journal article accepted for publication in the

International Journal of Sports Medicine 2011; In Press

Gough CE, Eastwood A, Saunders P U, Anson J M and Gore CJ

Presented here in the journal submission format

UNIVERSITY OF CANBERRA
AUSTRALIA'S CAPITAL UNIVERSITY

Chapter 4

This chapter has been removed due to copyright restrictions.

This chapter is available as:

Gough, C. E., Eastwood, A., Saunders, P. U., Anson, J. M. & Gore, C. J. (2012) Spurious Hb mass increases following exercise. *International Journal of Sports Medicine*. 33(9): 691-695.

Links to this chapter:

Print	http://webpac.canberra.edu.au/record=b1683904~S4
Online subscribed content (UC community)	http://ezproxy.canberra.edu.au/login?url=https://www.thieme-connect.com/ejournals/abstract/10.1055/s-0031-1295441
Online general public	https://www.thieme-connect.com/ejournals/abstract/10.1055/s-0031-1295441
DOI	10.1055/s-0031-1295441

Abstract

Sensitivity of the Athlete Blood Passport for blood doping could be improved by including total haemoglobin mass (Hbmass), but this measure may be unreliable immediately following strenuous exercise. We examined the stability of Hbmass following ultra-endurance triathlon (3.8 km swim, 180 km bike, 42.2 km run). 26 male sub-elite triathletes, 18 Racers and 8 Controls, were tested for Hbmass using CO re-breathing, twice 1-5 days apart. Racers were measured before and 1-3 h after the triathlon. Controls did no vigorous exercise on either test day. Serum haptoglobin concentration and urine haemoglobin concentration were measured to assess intravascular haemolysis. There was a 3.2% ($p<0.01$) increase in Racers' Hbmass from pre-race (976 g±14.6%, mean ±% coefficient of variation) to post-race (1 007 g±13.8%), as opposed to a − 0.5% decrease in Controls (pre-race 900 g±13.9%, post-race 896 g±12.4%). Haptoglobin was − 67% ($p<0.01$) reduced in Racers (pre-race 0.48 g / L±150%, post-race 0.16 g / L±432%), compared to − 6% reduced in Controls (pre-race 1.08 g / L±37%, post-race 1.02 g / L±37%). Decreased serum haptoglobin concentration in Racers, which is suggestive of mild intravascular blood loss, was contrary to the apparent Hbmass increase post-race. Ultra-endurance triathlon racing may confound the accuracy of post-exercise Hbmass measures, possibly due to splenic contraction or an increased rate of CO diffusion to intramuscular myoglobin.

Chapter 5

The effects of injury and illness on haemoglobin mass

Journal article submitted to

International Journal of Sports Medicine; March 2012 (In Review)

Gough CE, Sharpe K, Garvican LA, Anson JM, Saunders PU and Gore CJ

Presented here in the journal submission format

UNIVERSITY OF CANBERRA
AUSTRALIA'S CAPITAL UNIVERSITY

Chapter 5

This chapter has been removed due to copyright restrictions.

This chapter is available as:

Gough, C. E., Sharpe, K., Garvican, L. A., Anson, J. M., Saunders, P. U. & Gore, C. J. (n.d.) The effects of injury and illness on haemoglobin mass. *International Journal of Sports Medicine*.

Links to this chapter:

Print	http://webpac.canberra.edu.au/record=b1683904~S4
Online subscribed content (UC community)	http://ezproxy.canberra.edu.au/login?url=https://www.thieme-connect.com/ejournals/abstract/10.1055/s-0033-1333692
Online general public	https://www.thieme-connect.com/ejournals/abstract/10.1055/s-0033-1333692
DOI	10.1055/s-0033-1333692

Abstract

This study sought to quantify the effects of reduced training, surgery and changes in body mass on haemoglobin mass (Hbmass) in athletes. Hbmass of 15 athletes (6 males, 9 females) was measured 9±6 (mean±SD) times over 162±198 days, during reduced training following injury or illness. Additionally, body mass (n=15 athletes) and episodes of altitude training (n=2), iron supplementation (n=5), or surgery (n=3) were documented. Training was recorded and compared with pre-injury levels. Analysis used linear mixed models for ln(Hbmass), with Sex, Altitude, Surgery, Iron, Training and log(Body Mass) as fixed effects, and Athlete as a fixed and random effect. Reduced training and surgery led to 2.3% (p=0.02) and 2.7% (p=0.04) decreases in Hbmass, respectively. Altitude and iron increased Hbmass by 2.4% (p=0.03) and 4.2% (p=0.05), respectively. The effect of changes in body mass on Hbmass was not statistically significant (p=0.435).The estimates for the effects of surgery and altitude on Hbmass should be confirmed by future research using a larger sample of athletes. These estimates could be used to inform the judgements of experts examining athlete biological passports, improving their interpretation of Hbmass perturbations, which athletes claim are related to injury, thereby protecting innocent athletes from unfair sanctioning.

Chapter 6

Influence of altitude training modality on performance and total haemoglobin mass in elite swimmers

Journal article accepted for publication in the

European Journal of Applied Physiology 2012; Epub ahead of print

Gough CE, Saunders PU, Fowlie J, Savage B, Pyne DB,

Anson JM, Wachsmuth N, Prommer N and Gore CJ

Presented here in the journal submission format

UNIVERSITY OF CANBERRA
AUSTRALIA'S CAPITAL UNIVERSITY

Chapter 6

This chapter has been removed due to copyright restrictions.

This chapter is available as:

Gough CE, Saunders PU, Fowlie J, Savage B, Pyne DB, Anson JM, Wachsmuth N, Prommer N and Gore CJ (2012) Influence of altitude training modality on performance and total haemoglobin mass in elite swimmers. *European Journal of Applied Physiology.* 112(9): 3275-3285.

Links to this chapter:

Print	http://webpac.canberra.edu.au/record=b1683904~S4
Online subscribed content (UC community)	http://ezproxy.canberra.edu.au/login?url=http://link.springer.com/article/10.1007/s00421-011-2291-7
Online general public	http://link.springer.com/article/10.1007/s00421-011-2291-7
DOI	10.1007/s00421-011-2291-7

Abstract

We compared changes in performance and total haemoglobin mass (tHb) of elite swimmers in the weeks following either Classic or Live High:Train Low (LHTL) altitude training. Twenty-six elite swimmers (15 male, 11 female, 21.4 ± 2.7 years; mean ± SD) were divided into two groups for 3 weeks of either Classic or LHTL altitude training. Swimming performances over 100 or 200 m were assessed before altitude, then 1, 7, 14 and 28 days after returning to sea-level. Total haemoglobin mass was measured twice before altitude, then 1 and 14 days after return to sea-level. Changes in swimming performance in the first week after Classic and LHTL were compared against those of Race Control (n = 11), a group of elite swimmers who did not complete altitude training. In addition, a season-long comparison of swimming performance between altitude and non-altitude groups was undertaken to compare the progression of performances over the course of a competitive season. Regardless of altitude training modality, swimming performances were substantially slower 1 day (Classic 1.4 ± 1.3% and LHTL 1.6 ± 1.6%; mean ± 90% confidence limits) and 7 days (0.9 ± 1.0% and 1.9 ± 1.1%) after altitude compared to Race Control. In both groups, performances 14 and 28 days after altitude were not different from pre-altitude. The season-long comparison indicated that no clear advantage was obtained by swimmers who completed altitude training. Both Classic and LHTL elicited ~4% increases in tHb. Although altitude training induced erythropoeisis, this physiological adaptation did not transfer directly into improved competitive performance in elite swimmers.

Chapter 7

Comparison of Acute Intermittent Hypoxic Exposure and Live High:Train Low altitude

Journal article submitted to

International Journal of Sports Physiology and Performance; January 2012 (In Review)

Gough CE, Saunders PU, Bonetti DL, Stephens S, Bullock N, Anson JM and Gore CJ

Presented here in the journal submission format

Abstract

Live High:Train Low (LHTL) altitude training is a popular ergogenic aid amongst athletes. However, high financial and logistical demands of LHTL make it inaccessible to many athletes. An alternative hypoxia protocol, Acute (60-90 min daily) Intermittent Hypoxic Exposure (AcIHE), has shown potential for improving athletic performance. This study is the first to measure the influence of AcIHE on athletes' Hb_{mass} to determine whether the brief hypoxic exposures during AcIHE are sufficient to stimulate additional erythropoiesis, a potential mechanism for improved performance. Changes in Hb_{mass}, peak oxygen consumption (VO_{2peak}), velocity at VO_{2peak} (vVO_{2peak}), running economy, maximal blood lactate concentration [Lac] and 3 mM lactate running speed were compared following 17 days of LHTL, AcIHE or Placebo treatment in 24 Australian National Team triathletes (7 female, 17 male). There was a clear 3.2 ± 4.8% (mean ± 90% confidence limits) increase in Hb_{mass} following LHTL compared with Placebo, whereas the corresponding change of -1.4 ±4.5% in AcIHE was unclear. There were no clear changes in running economy, VO_{2peak} and vVO_{2peak} following either method of hypoxia but an increase in 3mM [Lac] running speed and maximal [Lac] suggested a beneficial shift in the lactate-power profile following LHTL. The clear difference in Hb_{mass} response between LHTL and AcIHE indicates that any positive changes in athletic performance following AcIHE are unlikely to be due to haematological adaptation. However, the shared responses of decreased maximal [Lac] and HR between both LHTL and AcIHE may point to a common physiological adaptation for both methods of hypoxic exposure.

Introduction

Altitude training first became popular with athletes as part of their physical preparation for competition nearly fifty years ago, and over the intervening period many different altitude training protocols have evolved. In the past 15 years, numerous researchers have investigated the effects of Live High:Train Low (LHTL) altitude training, where athletes live at moderate altitude (2000-3000 m) but train near sea-level, on subsequent sports performance (Levine and Stray-Gundersen(1997) for instance). Providing athletes are exposed to an adequate 'dose' of altitude, LHTL can lead to worthwhile performance improvements (Bonetti and Hopkins 2009) and, therefore, is a popular ergogenic aid amongst elite athletes. The specific facilities required for LHTL altitude protocols can be logistically and financially inaccessible to many athletes, as either a location with rapid travel options between a low altitude training venue and a moderate altitude residential facility, or a special purpose 'altitude house' is required where the hypoxic environment can be simulated by reducing the oxygen content of the ambient air.

One alternative, Acute (60-90 min daily) Intermittent Hypoxic Exposures (AcIHE), was highlighted by a recent meta-analysis (Bonetti and Hopkins 2009) as one of the most beneficial forms of altitude training in sub-elite athletes. However, scientific opinion about the efficacy of AcIHE is divided; worthwhile improvements have been demonstrated by some researchers (Bonetti et al. 2006; Bonetti et al. 2009; Katayama et al. 2003; Wood 2006), whilst other researchers found unchanged or impaired performance (Hamlin et al. 2010; Julian et al. 2004). While AcIHE offers major practical advantages over LHTL, to date its utility in elite athletes has not been sufficiently explored. An investigation directly comparing these two forms of altitude training is, therefore, warranted.

Reflecting the relatively recent emergence of the AcIHE concept, very few studies have demonstrated physiological changes that could be responsible for the documented performance benefits of AcIHE. Increased haemoglobin mass (Hb_{mass}), which is recognised as a major contributor to performance improvements following LHTL (Robertson et al. 2010b), has not been examined following AcIHE. However, evidence of increased haemoglobin concentration ([Hb]) (Bonetti et al. 2006; Bonetti et al. 2009) and increased serum erythropoietin concentration (Rodriguez et al. 2000) following AcIHE are suggestive of a positive haematological adaptation. Considerations of non-haematological adaptations reveal a potentially common mechanism after altitude exposure; there is evidence of similar improvements in sub-maximal exercise efficiency in runners following AcIHE (Katayama et al. 2003) and LHTL (Saunders et al. 2009b).

The aim of this study was to compare directly the effects of two different forms of hypoxic exposure, LHTL and AcIHE, in elite athletes.

Methods

Study design

Twenty-four Australian National Team triathletes (7 female, 17 male) took part in this randomised placebo-controlled study. This study was approved by the Australian Institute of Sport Ethics Committee and all athletes provided their informed consent to participate. The athletes attended a 21-day running-focused training camp during the domestic competition season in Canberra, Australia (600 m) in which they were randomly assigned to one of three groups: LHTL, AcIHE or Placebo. The groups were evenly matched for peak oxygen consumption (VO_{2peak}), as measured during the incremental treadmill test at the start of the camp (Table 7.1). All athletes trained in the normal Canberra environment (minimum temp 16.5 ± 3.5ºC, maximum temp 31.6 ± 5.1ºC; mean ± SD), completing ~30 km swimming, ~400

km cycling, ~85 km running, and one strength session in the gym per week. The athletes recorded the duration (min), distance (km) and intensity (1-10 rating of perceived exertion scale) of all training completed one month prior to, and during the camp. The training impulse (TRIMP) for each session, which can be interpreted as the integrated training load (Banister et al. 1975), was calculated by multiplying duration and intensity of the session. Total body Hb_{mass}, serum ferritin ([Ferr]) and soluble transferrin receptor ([STfR]) concentrations, and various physiological variables associated with running were measured before and after the intervention (Figure 7.1).

Table 7.1: Physical characteristics of participants at baseline

	LHTL		AcIHE		Placebo	
	Mean (SD)		Mean (SD)		Mean (SD)	
Sex	m	f	m	f	m	f
n	5	2	5	2	6	3
Age (yr)	21.2 (1.6)	20.9 (3.5)	20.0 (2.6)	23.4 (6.4)	21.2 (1.6)	20.4 (3.9)
Height (m)	1.80 (0.08)	1.65 (0.04)	1.78 (0.04)	1.70 (0.01)	1.79 (0.04)	1.62 (0.04)
Mass (kg)	70.5 (5.9)	53.7 (1.2)	68.9 (3.3)	53.6 (1.2)	66.2 (6.1)	51.7 (0.8)
Hb_{mass} (g)	966 (91)	626 (6)	964 (90)	620 (23)	899 (109)	568 (43)
$VO_{2\ peak}$ ($mL.kg^{-1}.min^{-1}$)	74.8 (4.2)	63.3 (8.6)	72.4 (4.0)	60.6 (0.0)	70.5 (1.7)	64.2 (2.4)
$vVO_{2\ peak}$ ($km.h^{-1}$)	19.4 (0.6)	18.9 (0.2)	19.5 (0.7)	17.0 (1.1)	19.9 (0.6)	17.1 (0.9)

Total body haemoglobin mass: Hb_{mass}. The highest two consecutive 30 sec VO_2 values recorded during the pre-camp incremental treadmill test: VO_{2peak}. Running velocity at VO_{2peak} recorded during the pre-camp incremental treadmill test: vVO_{2peak}.

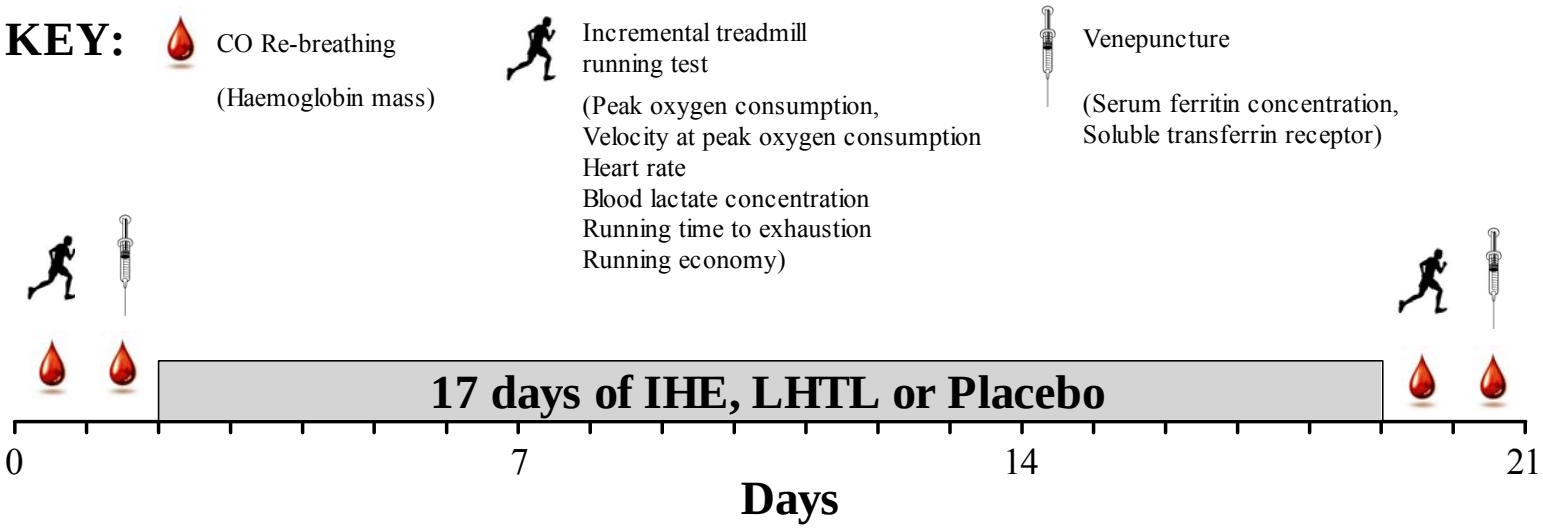

Figure 7.1: Outline of the study design illustrating the sequence of carbon monoxide (CO) re-breathing tests, incremental treadmill tests and venepuncture blood sampling.

Hypoxic exposure

The LHTL group (Table 7.1) spent 14.1 ± 0.1 $h.day^{-1}$ for 17 days in normobaric hypoxia equivalent to an altitude of 3000 m. For the AcIHE group, hypoxia was produced using a commercially available re-breathing device ($AltO_2lab$, Pharma Pacific, Phoenix, Arizona, USA), which uses spacers to increase respiratory deadspace and thereby reduce the partial pressure of oxygen in inspired air at the lungs. Expired air was passed through a soda-lime absorbent to reduce carbon dioxide (CO_2) in the $AltO_2lab$. Over a 17 day period, the AcIHE group completed one AcIHE session per day, during which they alternated breathing though the device for 6 min and normal room air for 4 min; this cycle was repeated six times, totalling 60 min. Peripheral oxygen saturation (SpO_2) was monitored continuously by experimenters via pulse oximetry (Avant 4000, Nonin Medical Inc., Plymouth, Minnesota, USA) and was progressively reduced from 90% on day 1, to 76% on days 14 to 17 by the addition of more spacers. The hypoxic stimulus was equivalent to that of 3500-6000 m altitudes (http://www.high-altitude-medicine.com/SaO2-table.html (Hackett and Roach 1995)). The Placebo group completed identical duration 'AcIHE' sessions daily for 17 days, but their re-breathing devices had been modified by removing the soda lime absorbent. This method of creating a Placebo AcIHE condition has been published (Wood 2006). In the current study, the placebo configuration of the re-breathing device had the effect of generating only mild hypoxia (SpO_2 96 ± 0.5%, mean ± SD) and mild hypercapnia during the 6 min intervals from days 1 to 17. The magnitude of hypercapnia during AcIHE and Placebo sessions was explored in a pilot study (n=4); end-expiratory CO_2 concentrations of 4.7 ± 0.3% and 5.7 ± 0.4% were recorded for AcIHE and Placebo, respectively, and the corresponding end-inspiratory CO_2 concentrations were 1.2 ± 0.3% and 2.9 ± 0.4%. By using telemetered pulse oximeters, athletes in both the AcIHE and Placebo groups were blinded to their SpO_2 throughout the intervention period.

Blood parameters

Changes in Hb_{mass} from the start to the end of the training camp were assessed using the Optimised carbon monoxide (CO) re-breathing technique as published by our group previously (Gough et al. 2011). Duplicate measures of Hb_{mass} were made both pre- and post-intervention and averaged to a single value at each time point for analysis. The typical error of Hb_{mass} (with 90% confidence interval) was 2.4 (2.1 to 2.9)% from the pooled duplicate data of all three groups. Using the mean of the duplicate pairs reduced error by a factor of √2, when compared with singleton measures.

A venous blood sample (4mL) was collected from the athletes at the start and the end of the camp and analysed for [Ferr] and [STfR] using immunoturbidimetric assay on a Hitachii 911 automatic analyser (Boehringer Mannheim, Germany). Iron supplementation for all three groups (Ferrograd C, Ferrogradumet; Abbott Australia, Botany, Australia equivalent to 105 mg elemental iron per day) started two weeks prior to the first day of the camp and supplementation continued for the duration of the training camp.

Incremental treadmill test

Following a 5 min warm-up at 14 $km.h^{-1}$ (12 $km.h^{-1}$ for females), athletes ran at 14, 15, 16 and 17 $km.h^{-1}$ (12, 13, 14 and 15 $km.h^{-1}$ for females) for four mins at each speed, separated by 1-min rest periods. Heart rate (HR) was continuously recorded using short-range telemetry (Polar Vantage NV, Kempele, Finland). After each 4-min stage, a capillary blood sample was taken from the finger and measured for blood lactate concentration ([Lac]) using a portable analyser (Lactate Pro, Arkray, KDK Corporation, Kyoto, Japan). During pre-camp testing, the [Lac] of most athletes was >4 mM by the end of the fourth submaximal stage, but athletes for whom this was not the case (n=6) completed a fifth submaximal stage at 18 $km.h^{-1}$ (16 $km.h^{-1}$ for females). During post-camp testing, regardless of [Lac], athletes completed the same

number of submaximal stages as they had done pre-camp, which resulted in some athletes' [Lac] being < 4 mM post-camp. Consequently, in order to assess changes in the submaximal [Lac] profile from pre- to post- intervention, the running speed corresponding to 3 mM [Lac], rather than the traditional 4 mM, was calculated using an integrative technique of plotting speed versus [Lac] using an exponential fit. An in-house automated metabolic system, which has been described previously (Saunders et al. 2004), was used for measurement of oxygen consumption (VO_2) throughout the protocol. The VO_2 values for the final minute of each submaximal stage were pooled and averaged to give a measure of running economy. Upon completion of the final submaximal stage, participants rested for 5 min before completing an incremental run to maximal volitional fatigue, beginning at 16 $km.h^{-1}$ (14 $km.h^{-1}$ for females). The speed was increased by 1 $km.h^{-1}$ each minute until 20 $km.h^{-1}$ (18 $km.h^{-1}$ for females), then the gradient was increased by 0.5% per minute until volitional exhaustion. Every athlete was familiar with this test format from previous periodic testing. Time to exhaustion (TTE) during the maximal test and peak oxygen consumption ($VO_{2\ peak}$), taken as the highest two consecutive 30 sec VO_2 values, were recorded and the velocity at $VO_{2\ peak}$ ($vVO_{2\ peak}$) was calculated using an integrative technique of plotting speed versus VO_2 for the 4 submaximal stages and forming a regression equation that was solved for VO_{2peak} (Billat and Koralsztein 1996). For each athlete, the pre- and post-camp treadmill tests were completed at the same time of day, and the nutritional intake for 24 hours prior to the first test was recorded and athletes were asked to replicate that same diet prior to their second test.

Participation variations

Three athletes did not participate in the post-camp incremental treadmill tests due to injuries sustained in the latter stages of the training camp. One athlete completed the submaximal but not the maximal steps of the post-camp treadmill test due to an injury that limited top-speed running only, therefore submaximal running data only were included for this athlete. In

addition, the treadmill test results of a further two athletes were excluded because their data indicated a leak in the gas analysis system during one of their tests; the likely source was air leakage around the mouthpiece. All 24 athletes completed tests for blood parameters including Hb_{mass}. After exclusions, there were 6, 8 and 6 athletes in the LHTL, AcIHE and Placebo groups, respectively, for submaximal running variables; and the corresponding numbers for the maximal running variables were 5, 8 and 6 athletes.

Statistical analysis

Data were analysed using a contemporary analytical approach involving magnitude-based inferences (Hopkins et al. 2009), which enables small effects that are of practical importance in an elite athlete population to be detected. In order to reduce any effects of non-uniformity of error all measures were log-transformed before analysis. Preliminary analyses revealed large between-group differences in pre-camp training load (4922 ± 21.1%, 5246 ± 23.3% and 3618 ± 19.7% arbitrary TRIMP units for LHTL, AcIHE and Placebo, respectively) but only a small between-group difference in training load during the camp (5627 ± 36.8%, 6691 ± 28.4% and 6397 ± 6.4% arbitrary TRIMP units for LHTL, AcIHE and Placebo, respectively). To reduce the likelihood of training-induced changes impacting the findings of the study, the percent change in weekly training load from pre- to during-camp for each individual athlete (range 0.3% to 140%) was incorporated as a covariate in the analysis of the blood and running variables. The mean percent change in each of the blood and running variables from pre- to post-intervention was calculated and the differences in the response of each of the hypoxic groups were compared to changes in the Placebo group (± 90% confidence limits (CL)) using independent t-tests (Hopkins 2006). The magnitude of differences were assessed in relation to the smallest worthwhile change (SWC) which, for each variable, was calculated as one fifth of the between-subject standard deviation of athletes' baseline data. SWCs were calculated separately for male and female athletes, and the mean value taken as the final SWC because a

mixed-sex cohort led to large between-subject standard deviations (SD) in body mass-related variables such as Hb_{mass}, VO_{2peak} and running economy. Effects were termed positive, trivial or negative depending on the magnitude of the change relative to the SWC and the spreads of the 90% CL were used to ascertain the certainty with which the effects could be classified: 50-74% possibly; 75–95% likely; 95–99% very likely; and >99% almost certainly. The effect was deemed "unclear" if its confidence interval overlapped the SWC thresholds for both positive and negative change.

Results

Blood parameters

The 4.3% increase in Hb_{mass} following LHTL (pre 852 g ± 25%, post 881 g ± 26%; mean ± SD) was higher than the 1.1% increase in the Placebo group (pre 754 g ± 29%, post 768 g ± 29%; Table 7.2). The net difference between these two groups was a clear increase of 3.2 ± 4.8% (mean ± 90% CL) in LHTL compared with Placebo (Table 7.3). It was unclear whether the –0.3% decrease in Hb_{mass} following AcIHE (pre 827 g ± 27%, post 831 g ± 27%) was substantially different from the change in the Placebo group (Table 7.3).

Similarly, the change in [STfR] over the period of training was substantially higher in LHTL, but not in AcIHE, relative to Placebo. There were small and variable changes in [Ferr] in all groups resulting in group mean (± SD) post-camp values of 64 $ng.mL^{-1}$ ± 43% , 45 $ng.mL^{-1}$ ± 72% and 44 $ng.mL^{-1}$ ± 131% for LHTL, AcIHE and Placebo, respectively.

Running parameters

The changes to running economy were not different in LHTL or AcIHE compared to Placebo, but athletes in the LHTL group experienced a beneficial change in 3mM [Lac] running speed over that of Placebo (Table 7.3). Both LHTL and AcIHE demonstrated substantial decreases

in maximal [Lac] and HR, compared with Placebo. There were no clear changes in either VO_{2peak} or vVO_{2peak} in any group.

Table 7.2: Percent change in blood and running parameters after either Live High:Train Low (LHTL) altitude training, Intermittent Hypoxic Exposure (AcIHE), or Placebo.

	SWC (%)	LHTL		AcIHE		Placebo	
		Mean change (± %CV)	Qualitative Inference	Mean change (± %CV)	Qualitative Inference	Mean change (± %CV)	Qualitative Inference
Blood measures							
Haemoglobin mass	1.6	4.3 (± 2.5)	Likely higher	-0.3 (± 2.3)	Unclear	1.1 (± 4.1)	Unclear
Serum ferritin	8.2	4.3 (± 30.0)	Unclear	16.3 (± 28.3)	Possibly higher	-11.2 (± 31.8)	Possibly trivial
Soluble transferrin receptor	3.0	27.8 (± 10.1)	Very likely higher	15.7 (± 6.0)	Almost certainly higher	11.8 (± 13.2)	Likely higher
Submaximal running							
Running economy	1.9	-3.3 (± 1.3)	Likely lower	-0.7 (± 3.0)	Likely trivial	-2.4 (± 2.7)	Likely lower
3mM [Lac] running speed	0.8	4.2 (± 3.1)	Likely higher	2.5 (± 3.3)	Likely higher	-0.2 (± 2.2)	Unclear
Maximal running							
VO_{2peak}	1.5	0.1 (± 4.9)	Unclear	2.3 (± 5.3)	Unclear	-1.6 (± 4.6)	Unclear
vVO_{2peak}	0.9	0.5 (± 4.6)	Unclear	0.1 (± 2.3)	Unclear	-1.2 (± 5.6)	Unclear
TTE	3.0	4.0 (± 12.2)	Unclear	-0.8 (± 12.5)	Unclear	4.1 (± 3.6)	Possibly higher
Maximal heart rate	0.8	-5.3 (± 2.2)	Very likely lower	-2.5 (± 2.5)	Likely lower	0.4 (± 1.6)	Unclear
Maximal [Lac]	3.5	-25.6 (± 16.9)	Very likely lower	-12.8 (± 26.1)	Unclear	10.0 (± 11.5)	Unclear

Values are the group mean percent changes of the log-transformed data, measured from pre-camp to post-camp, ± standard deviation expressed as the percent coefficient of variation (% CV). Smallest Worthwhile Change in the variable, expressed as a percentage: SWC. Blood lactate concentration: [Lac]. Peak oxygen consumption: VO_{2peak}. Running velocity at peak oxygen consumption: vVO_{2peak}.

Table 7.3: Percent difference in the changes in blood and running parameters after either Live High:Train Low altitude training (LHTL) or Intermittent Hypoxic Exposure (AcIHE) compared with Placebo.

	LHTL		AcIHE	
	Difference from Placebo (± 90% CL)	Qualitative Inference	Difference from Placebo (± 90% CL)	Qualitative Inference
Blood measures				
Haemoglobin mass	3.2 (± 4.8)	Possibly higher	-1.4 (± 4.5)	Unclear
Serum ferritin	17.5 (± 46.6)	Unclear	31.0 (± 40.9)	Likely higher
Soluble transferrin receptor	14.3 (± 16.9)	Likely higher	3.4 (± 14.1)	Unclear
Submaximal running				
Running economy	-0.9 (± 4.0)	Unclear	1.8 (± 4.1)	Unclear
3mM [Lac] running speed	4.4 (± 4.5)	Likely higher	2.7 (± 3.7)	Unclear
Maximal running				
VO_{2peak}	1.7 (± 9.0)	Unclear	3.9 (± 7.1)	Unclear
vVO_{2peak}	1.7 (± 9.7)	Unclear	1.3 (± 8.3)	Unclear
TTE	-0.1 (± 20.0)	Unclear	-4.7 (± 12.2)	Unclear
Maximal HR	-5.7 (± 3.4)	Very likely lower	-3.0 (± 2.8)	Likely lower
Maximal [Lac]	-32.4 (± 27.2)	Very likely lower	-20.7 (± 26.8)	Likely lower

Values are the net difference between groups in the percent change in variables from pre-camp to post-camp with 90% confidence limits (± 90% CL). Blood lactate concentration: [Lac]. Peak oxygen consumption: VO_{2peak}. Running velocity at peak oxygen consumption: vVO_{2peak}.

Discussion

The clear increases in Hb_{mass} and [STfR] following LHTL demonstrate an erythropoietic response to this form of hypoxic exposure, which was absent following AcIHE. Increased [Ferr] in the AcIHE group confirms that iron availability was not a limiting factor for Hb_{mass} increases, and a more likely reason for the null haematological response is the difference in hypoxic "dose" between LHTL and AcIHE. It has been suggested previously that the minimum hypoxic "dose" needed to stimulate haematological adaptation is >12 $h.d^{-1}$ for at least 3 wk at an altitude or simulated altitude of 2100-2500 m (Rusko et al. 2004). The mean 3.2% increase in Hb_{mass} measured here after 17 d of LHTL, at 14 $h.d^{-1}$ and a simulated altitude of 3000 m, demonstrates that haematological adaptation can be achieved within fewer days if the severity of altitude and duration of exposure per day are increased. Although the hypoxia to which the athletes in the AcIHE group were exposed was more severe again (equivalent to 3500-6000 m), the findings of the present study confirm that this dose is insufficient to stimulate erythropoiesis since there was no increase in Hb_{mass}. This is the first time that changes in Hb_{mass} in response to AcIHE sessions lasting <3$h.d^{-1}$ have been examined. Our findings refute the suggestions of other researchers who, based on measured increases in Hb concentration and haematocrit coupled with decreased [Ferr] following 60-90 min AcIHE, concluded that Hb_{mass} may have increased (Bonetti et al. 2006; Bonetti et al. 2009). The current findings are unsurprising given that there was no increase in Hb_{mass} after 4 weeks of 3 $h.d^{-1}$ at 4000-5500 m (Gore et al. 2006b), where a larger cumulative dose of hypoxia is likely relative to the present study.

The absence of a haematological response, however, does not preclude the potential for a positive influence of AcIHE on athletic performance. Various non-haematological changes in athletes' physiology have been measured in response to hypoxia (Gore et al. 2007) and may

explain improved performance in the absence of increased Hb_{mass}. Improvements to the efficiency of oxygen usage during submaximal exercise is one such non-haematological change that has been demonstrated after LHTL (Gore et al. 2001; Saunders et al. 2009b) and AcIHE (Katayama et al. 2003). However, neither LHTL nor AcIHE resulted in worthwhile improvements in running economy in the current study. Running economy is one factor that, together with changes in VO_{2max} and lactate threshold, can account for 70% of the variance in endurance running performance (di Prampero et al. 1986).

Like running economy, there was no clear change in VO_{2peak} following both AcIHE and LHTL. This is somewhat surprising given the substantial increase in Hb_{mass} in the LHTL group that should theoretically transfer to a worthwhile improvement in maximal VO_2 of ~2% (Schmidt and Prommer 2010). Whilst there was an unclear 1.7% improvement in VO_{2peak} following LHTL compared with Placebo, the majority of this difference is due to a 1.6% decrease in VO_{2peak} in the Placebo group, not an increase in the LHTL group. One possible explanation for the incongruence between Hb_{mass} and VO_{2peak} in the LHTL group is that the decrease in maximal HR recorded after LHTL could have counteracted the positive effect of improved oxygen carrying capacity and resulted in no change to VO_{2peak}. A decreased maximal HR has previously been reported after LHTL at moderate altitude (Saunders et al. 2009b; Wehrlin et al. 2006) and similar changes after acclimatisation to severe and chronic altitude exposure have been attributed to changes in myocardial B-adrenergic and myocardial receptor density (Favret et al. 2001). Interestingly, since AcIHE also led to a similar decrease in maximal HR, these results together may be suggestive of a common mechanism of hypoxia adaptation for both LHTL and AcIHE.

A rightward shift in the lactate-power profile indicates that an athlete is able to run at a higher speed for the same or reduced lactate accumulation, and typically leads to improved running performance. The 4.4% increase in 3mM [Lac] running speed and decreased maximal [Lac]

following LHTL indicates a positive shift in the lactate profile, and although there were no clear changes in 3mM [Lac] running speed following AcIHE, there was decreased maximal [Lac] of a similar magnitude in both hypoxic exposure methods. Again, this similarity may be suggestive of a common hypoxia-induced adaptation. These changes in [Lac] are reminiscent of the "lactate paradox" (Hochachka et al. 2002): a well-known, although much debated, physiological adaptation that has been observed in subjects spending weeks in hypobaric hypoxia. The lactate paradox is characterised by rates of blood lactate accumulation during exercise lower than those recorded in normoxia, despite the limited-O_2 conditions. Although this phenomenon is most often discussed in terms of acclimatisation to much more severe hypoxia (>4500 m) and more prolonged periods (>4 wks) than those typically used for altitude training in athletes, rightward shifts in the lactate-power profile have been reported following LHTL (Nummela and Rusko 2000; Robertson et al. 2010b) and AcIHE (Bonetti et al. 2009; Wood 2006) protocols similar to those utilised in the current study. However, these changes do not appear to be consistent, with other researchers having reported no changes in the lactate profile following both LHTL (Gore et al. 2001; Robertson et al. 2010c) and AcIHE (Bonetti et al. 2006; Tadibi et al. 2007). In fact, the fickle nature of this adaptation has been demonstrated for both methods of hypoxia; the running speed corresponding to 4mM [Lac] was improved in one bout but not in a subsequent identical bout of LHTL (Robertson et al. 2010b), and despite using almost-identical AcIHE protocols, one research group recorded substantial changes in lactate profile (Wood 2006) whilst another found no such changes (Tadibi et al. 2007).

It is possible that, rather than hypoxia-induced adaptations, the decreases in maximal HR and [Lac] following LHTL and AcIHE were transient changes indicative of increased fatigue resulting from over-reaching (Meeusen et al. 2006) that have been observed a number of times following periods of intense training and could be reversed with a few days of sufficient

recovery (Faude et al. 2009). Hypoxia induces an additional physiological stress and can increase the occurrence of overtraining (Rusko et al. 2004). However, in this instance, if any groups were to suffer from undue fatigue, it is more likely that it would have been the Placebo group rather than the LHTL or AcIHE groups since the training load during the camp represented a much greater relative increase from their normal training. Unfortunately, due to the athletes' competition schedule it was not possible to delay the post-intervention treadmill tests until a few days after the end of the camp. This would have allowed a short period of recovery and may have reduced the possible influence of fatigue on the results.

It has been demonstrated previously that one additional parameter, vVO_{2max}, can alone predict up to 94% of the total variance in 16-km running performance (McLaughlin et al. 2010) and, as such, is a good indicator of endurance performance because it integrates both the maximal aerobic power and running economy (Billat and Koralsztein 1996). Again, there were no clear changes to this parameter relating to either method of hypoxic exposure in the current study. Of the four factors discussed here that have been shown to account for variance in running performance, only a positive change in 3mM [Lac] running speed in LHTL suggests any benefit for either hypoxic exposure method using the present protocol.

Limitations

The groups differed in the amount of training they had completed in the lead-up to the study, and consequently the training load of the camp would have served as a greater stimulus for some athletes than others. In order to neutralise the potential inequality of the training effect, the change in training load from pre-camp to during-camp was incorporated into the analyses as a covariate; however, we cannot discount the possibility that the results were affected by these training differences.

The number of participants for whom there are running data is less than those for blood parameters due to athlete injury drop-outs. Therefore, interpretation of these data is more difficult due to the effects being relatively small in magnitude with moderate variability between subjects.

Conclusions

The clear difference in Hb_{mass} response between LHTL and AcIHE indicates that any positive changes in athletic performance following AcIHE are unlikely to be due to haematological adaptation. There was little evidence for improved running performance after LHTL and AcIHE given the null responses of either hypoxic exposure method of the predicative parameters of running economy, VO_{2peak} and vVO_{2peak}. An increase in 3mM [Lac] running speed suggests a beneficial shift in the lactate-power profile following LHTL, and the shared responses of decreased maximal [Lac] and HR between both LHTL and AcIHE may point to a common physiological adaptation for both methods of hypoxic exposure.

Acknowledgements

The authors wish to thank all the triathletes and coaches who took part in this project. This project was jointly-funded by Triathlon Australia, the Australian Institute of Sport and the University of Canberra.

Chapter 8

Does the inclusion of haemoglobin mass in the Athlete Biological Passport improve detection of microdose rHuEPO doping?

This chapter is intended for submission to the

European Journal of Applied Physiology

Gough CE, Ashenden MJ, Sharpe K, Anson JM,

Saunders PU and Gore CJ

A separate manuscript originating from this study, for which I was not the lead author, was published in

European Journal of Applied Physiology 2011; 111 (9) 2307-14

Ashenden MJ, Gough CE, Garnham A, Gore CJ and Sharpe K

Abstract

Current markers of the Athlete Biological Passport (ABP) are unable to detect the use of microdoses of recombinant human erythropoietin (rHuEPO) in athletes (Ashenden et al. 2011). This study investigated whether the sensitivity of the ABP to microdose rHuEPO would be improved by the inclusion of total haemoglobin mass (Hb_{mass}). For 12 weeks, fifteen male recreational cyclists received two injections per week of either rHuEPO (n=10) or Placebo (n=5) and were measured for Hb_{mass} and %Ret fortnightly. Five models that used Hb_{mass} as their only variable ($Hb^{m\ (Prommer-small)}$, $Hb^{m\ (Prommer-large)}$, $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$) were examined for their sensitivity and specificity using the ABP software. These models differed only in the estimate of biological within-subject variance (BioWS variance) included in the calculations. Additionally, the sensitivity and specificity of one model combining Hb_{mass} and %Ret (ON^{hm+ret}) was assessed using calculations described in previous literature that simulated the computations of the ABP software. The sensitivity and specificity of all models were examined at both 99% and 99.9% specificity levels. Of the ten doped athletes, the $Hb^{m\ (Prommer-small)}$ model detected five, the $Hb^{m\ (Prommer-large)}$ detected three and the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models each detected two athletes at the 99.9% level. The ON^{hm+ret} model failed to detect any of the doped athletes at either the 99% or 99.9% levels. All models maintained 100% specificity at the 99.9% level, but the $Hb^{m\ (Prommer-small)}$ and $Hb^{m\ (Prommer-large)}$ models recorded three and one false-positives at the 99% level, respectively. The use of Hb_{mass} as a single marker in an ABP model resulted in improved sensitivity to microdose rHuEPO doping compared to existing ABP markers. The $Hb^{m\ (Prommer-small)}$ model was the most sensitive model, however, specificity remains an issue.

Introduction

Blood doping is a generic term for the illegal manipulation of athletes' blood for advantageous sporting outcomes. In 1989, a new form of blood doping emerged when recombinant human erythropoietin (rHuEPO) became available for medical use and quickly fell into the hands of unscrupulous athletes (Catlin et al. 2008). The injection of rHuEPO stimulates erythropoiesis, thereby increasing circulating haemoglobin levels and enhancing endurance sports performance due to the increased oxygen-carrying capacity of the blood (Parisotto et al. 2000a). The successful implementation of the isoelectric focusing technique (Lasne 2001) for direct detection of rHuEPO in urine has led athletes to use smaller 'microdoses' of rHuEPO, a technique which markedly improves their chances of evading detection (Ashenden et al. 2006). As well as the isoelectric focusing technique, the World Anti-Doping Agency (WADA) uses the haematological module of the Athlete Biological Passport (ABP) as an indirect method of blood doping detection. Using the ABP, WADA monitors longitudinal changes in [Hb], percent reticulocytes (%Ret) and an integrated measure of these two markers, the OFF-hr score, for signs of blood manipulation. An earlier publication from the present study demonstrated that a 12-week rHuEPO microdosing regimen was not detected by current markers of the ABP (Ashenden et al. 2011) despite the treatment inducing a substantial erythropoiesis. The red blood cell response was apparent from the mean 10% increase in the doped athletes' total haemoglobin mass (Hb_{mass}). The apparent insensitivity of existing anti-doping tests, both direct and indirect, to rHuEPO microdosing demonstrates the need for development of a new method for detection of this doping practice. A number of researchers have asserted that the inclusion of Hb_{mass} as an additional marker in the ABP could improve the sensitivity of the passport (Morkeberg et al. 2011; Pottgiesser et al. 2007; Prommer et al. 2008). Whilst it has been demonstrated that monitoring Hb_{mass}, alone and in combination with %Ret, enhanced the sensitivity of the ABP

to autologous blood doping (Morkeberg et al. 2011; Pottgiesser et al. 2012), these new blood markers have not yet been tested for their efficacy in detecting rHuEPO use.

Being found guilty of a blood doping offence has a damning effect on an athlete's career and reputation, and it is vital for anti-doping authorities to ensure adequate protection for innocent athletes against false-positive results. For each blood test result, the Adaptive model (calculations used within the ABP) defines a range in which the result is expected to fall, and results falling outside the specified range are 'flagged' for the attention of the anti-doping authorities. The upper and lower limits of the expected range are initially based upon population mean and variance values for the blood marker in question, but as more tests are conducted on an athlete, the range is progressively more reliant on that individual's past results (Sottas et al. 2010). Despite this adaptation, the estimate of within-subject variance entered into the Adaptive model for each blood marker remains unaltered by previous test results and is, therefore, crucial for determining the upper and lower limits of the expected range. In the case of Hb_{mass}, it has been demonstrated that a 'two-levels' error model that splits within-subject variance into analytical and biological components is most appropriate (Pottgiesser et al. 2012; Prommer et al. 2008). In the current version of the ABP software, an Adaptive model using Hb_{mass} as its sole marker is included, although it is for experimentation purposes rather than for implementation at this stage. This Adaptive model includes separate estimates of Typical Error (TE; which represents the analytical standard deviation, expressed as a percentage) and biological within-subject variance (BioWS variance) (Pottgiesser et al. 2012). Together, the TE and the BioWS variance comprise the within-subject variance of Hb_{mass}.

Four separate studies (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008) have generated estimations of the within-subject variance of Hb_{mass} in athletes (ranging between 1.6% and 4.0%). When these estimates are split into their analytical

and biological components, it is clear that these researchers are in closer agreement about the magnitude of the TE than about the magnitude of the BioWS variance. Estimates of the TE range from 1.4% to 2.0%, whereas estimates of BioWS variance range from 56 g^2 to 830 g^2 (equivalent to ~0.8% to 2.8%). The sensitivity (rate of correct identification of doped athletes) and specificity (rate of correct identification of non-doped athletes) of possible future Hb_{mass} models to be included in the ABP rely on the anti-doping authorities selecting the most appropriate values of TE and BioWS variance. Indeed, one research group (Lundby and Robach 2010) has asserted that Hb_{mass} should not be used in an anti-doping context because rHuEPO-induced changes could not be differentiated from within-subject variance in 50% of their subjects after 13 weeks of microdosing. However, these authors did not formally assess the sensitivity and specificity of an Adaptive model based on Hb_{mass}, which should be done before final conclusions can be made regarding its suitability for inclusion in the ABP.

The aim of this study was to assess the sensitivity and specificity of six Adaptive models based on Hb_{mass} to microdose rHuEPO doping. Five of the models used Hb_{mass} as a single marker, differing only in the BioWS variance values included in the calculations, whilst the sixth model combined Hb_{mass} and %Ret.

Methods

Participants

Fifteen healthy male recreational cyclists volunteered to participate in this double-blind study and after having the potential risks explained to them gave their informed consent. The participants, who were recruited from local sporting clubs, had a history of regular cycling training (at least 3 sessions per week, >2 hours per session) and were asked to continue with their normal training throughout the course of the study but were prohibited from competing

during, and for at least 6 weeks after, the study. The procedures were approved by the ethics committees of the Australian Institute of Sport and the University of Canberra.

Study design

The participants underwent medical screening to ensure they were free from injury, illness or high blood pressure. Subsequently, the participants completed an incremental cycle ergometer test for determination of maximal aerobic power (VO_{2max}), and were randomly divided into two groups matched for VO_{2max}. The characteristics of the rHuEPO group were: n=10, age 31.4 ± 7.0 yr, height 181 ± 7.2 cm, body mass 80.0 ± 9.1 kg, VO_{2max} 58.2 ± 5.2 ml.kg.min^{-1} (mean ± SD); the corresponding values for the Placebo group were: n=5, 35.2 ± 2.9 yr, 178 ± 6.3 cm, 74.5 ± 6.8 kg, 58.2 ± 3.7 ml.kg.min^{-1}. All members of the research team who were directly involved with testing were blinded to the groupings. Over a 12 week period, participants attended the laboratory twice weekly for intravenous injections of either rHuEPO (Neorecormon, Roche Diagnostics, Australia) or a saline solution. Eight of the 15 participants began the study two weeks later and therefore completed the study over a period of 10 weeks. Venous blood samples were obtained twice within one week before the start of the study for double baseline measures of [Hb] and %Ret, and then once per week during the study. Double baseline measures of Hb_{mass} were also made one week before the study and then single Hb_{mass} measures made fortnightly throughout the study period. All participants took daily oral iron supplements throughout the study, which provided 105 mg elemental iron per day (Ferro-Grad C, Abbott Australasia Pty Ltd, Australia), to ensure adequate iron stores for accelerated erythropoiesis (Berglund 1992).

Our aim was to recreate a rHuEPO microdosing regimen typical of that which may be used by athletes. Administration was cautious with dosages increased progressively in response to feedback about blood changes so as to maximise performance benefits whilst minimising the

chances of being caught. One member of the research team, who was not directly involved in testing, monitored weekly [Hb] and %Ret results along with fortnightly Hb_{mass} results to increase the rHuEPO dose over the study period with the aim of causing minimal fluctuation in %Ret. An initial 'titration' phase was completed to ensure that all participants received the same quantity of rHuEPO regardless of start date; those who started two weeks earlier were given 10 $IU.kg^{-1}$ body mass per week for the first 4 weeks and the others 20 $IU.kg^{-1}$ per week for their first two weeks. Dosages were subsequently prescribed on an individual basis according to the feedback from the weekly [Hb] and %Ret results; all subjects received 20 $IU.kg^{-1}$ for the next four weeks, then 30 $IU.kg^{-1}$ for the final 4 weeks, with the exception of 3 subjects in the rHuEPO group who were given 40 $IU.kg^{-1}$ for their final 3 injections.

Haemoglobin mass estimation

Total body haemoglobin mass was measured using the optimised carbon monoxide (CO) rebreathing technique (Gough et al. 2011). Briefly, subjects re-breathed a bolus of CO equivalent to 1.4 $mL.kg^{-1}$ of body mass through a glass spirometer (BloodTec, Germany) for two minutes. Percent carboxyhaemoglobin (%HbCO) in fingertip capillary blood was measured using an OSM3 hemoximeter (Radiometer, Copenhagen, Denmark) before and seven minutes after administration of the CO dose. Ten repeat measures of %HbCO were made for improved precision in Hb_{mass} estimation (Alexander et al. 2011). The same researcher completed all measures of Hb_{mass}. The typical error for Hb_{mass}, based on the baseline duplicate measures, was 2.8% (95% Confidence Limits (CL): 2.0 to 4.5%).

Venous blood analysis

Four millilitres of venous blood was drawn into K_3EDTA vacutainers by a qualified phlebotomist and refrigerated until analysis. The analysis was conducted within 24 hours of blood collection. A full blood count was conducted on one of two side-by-side Sysmex XE-

2100 instruments in a commercial pathology laboratory. Records from the external quality assurance scheme in which the laboratory participated confirmed that both instruments performed within the allowable limits of the programme (Royal College of Pathologists of Australia QAP programme).

Samples were homogenised by mixing on a roller mixer for at least 15 mins then measured in duplicate, in accordance with WADA guidelines (World Anti-Doping Agency 2010). Given a satisfactory difference between the two replicate measures (< 0.1g.dL^{-1} for [Hb] and <0.15 absolute difference for %Ret or <0.25 difference if the %Ret was >1.00%), only the first data point was included, or else the entire analyses were discarded and then repeated.

Calculations and Statistics

The percentage changes in Hb_{mass} and %Ret for each athlete were calculated at each timepoint compared with baseline values and results were displayed as the group mean ± standard deviation (SD). The intra-individual variability of Hb_{mass} over the 12-week period was calculated and expressed as the percent coefficient of variation (%CV).

Haemoglobin mass models

To examine the sensitivity and specificity of Hb_{mass} models to rHuEPO doping, six separate analyses were completed as part of this study. Five different Adaptive models that used Hb_{mass} as a single parameter (Hb^{m} models) were examined using the ABP software. Additionally, one novel model that combined Hb_{mass} and %Ret into a single parameter (ON^{hm+ret} model) was examined using calculations similar to those used in the Adaptive model, although this analysis was not conducted using the ABP software.

The Adaptive models for Hb_{mass} in the ABP software use four population-derived values to generate the expected ranges of test results for an individual athlete: the population mean,

between-subject variance, TE and BioWS variance (Pottgiesser et al. 2012). In this investigation the sensitivities and specificities of five Hb^m models with different values of BioWS variance were assessed: $Hb^{m\ (Prommer\ –\ small)}$ 56.25 g^2, $Hb^{m\ (Prommer\ –\ large)}$ 244 g^2, $Hb^{m\ (Pottgiesser)}$ 550 g^2, $Hb^{m\ (Morkeberg)}$ 611 g^2 and $Hb^{m\ (Eastwood)}$ 830 g^2. These estimates of BioWS variance were gathered from four research studies (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008). The Hb^m models are named after the first author of each study, with the exception of $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$ where two estimates of BioWS variance were derived from the same study (See Appendix 1 for a detailed description of how the estimates of BioWS variance were derived). Identical values for population mean (11.84 x body mass kg +149 g), between-subject variance (3994 g^2) and TE (1.7%) were used for all five models. The estimates of population mean and between-subject variance used in the Hb^m models were published by Prommer et al. (2008). The TE is the average analytical error from the four key studies (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008). The ABP software assessed each Hb_{mass} result individually (single result analysis), and results that fell outside the expected range were flagged. Additionally, the variability of the full sequence of results was compared against the expected variability of sequences of the same length (sequence analysis). A sequence of results with an abnormally high variance is indicative of doping (Pottgiesser et al. 2012; Sottas et al. 2010). These analyses presented two separate opportunities for doping detection, allowing identification of not only individual anomalous results but also abnormal variation in a sequence of results.

The flagged results that occurred on the first test for two athletes in the rHuEPO group were discounted. This is because their results fell below the lower limits of the expected range, simply reflecting the consequences of judging non-elite athletes against a population mean that is representative of the elite athlete population. Once these athletes' first test results were

entered, the Adaptive model of the ABP individualised the expected range for the next result, taking into account their first Hb_{mass} result (Sottas et al. 2010). So, although the first results for those two athletes were affected, all of their subsequent tests were unaffected and were included in the single results analyses. However, the abnormal first test results caused an artificially high sequence result for these two athletes. Consequently, the sequence analysis results of these two athletes were also discounted from the sensitivity results for all five Hb^{m} models.

The sensitivity and specificity of one model that combined Hb_{mass} and %Ret, the ON^{hm+ret} model, was also assessed. The ON^{hm+ret} model was generated using the raw data from the publication of Parisotto et al. (2000a) to combine Hb_{mass} and %Ret in a way that best differentiated between the 18 doped (rHuEPO injections) and 9 non-doped subjects who participated in that study (Equation (1)).

$$ON^{hm+ret} = \sqrt{\%\text{Ret}} + 2 \times \ln(\text{Hb}_{\text{mass}}) \qquad (1)$$

The same discriminant analysis modelling approach as previously described by Parisotto et al. (2001) was used to derive the ON^{hm+ret} model. A high ON^{hm+ret} score results from a combination of high Hb_{mass} and high %Ret, which is characteristic of rHuEPO use. The Hb_{mass} and %Ret results for the 15 athletes who participated in the present study were integrated at each time point, in accordance with equation (1), to create a series of ON^{hm+ret} scores for each individual athlete. Subsequently, each ON^{hm+ret} score was compared against the expected range generated for the individual athlete using calculations described in previous literature (Ashenden et al. 2011; Morkeberg et al. 2011; Sottas et al. 2010). These calculations are based on Bayesian network statistics and simulate the Adaptive model included in the ABP software. We verified the accuracy of these calculations using examples of [Hb] results from 3 subjects, comparing a series of upper and lower limits from our calculations to those

generated by the Adaptive model in the ABP software, and obtained identical figures. The following values were used in the calculations for the ON^{hm+ret} model: population mean of 14.70, between-subject variance of 0.235, within-subject variance of 0.0256. These estimates were derived from 184 observations on 34 elite and semi-elite athletes who served as control subjects in four research studies conducted by our group in recent years, including Garvican et al. (2012) Robertson et al. (2010b) and Saunders et al. (2010a). The ON^{hm+ret} model allowed each individual result to be compared against the individualised reference ranges, but it was not possible for a sequence analysis to be completed. The sensitivity and specificity results of the ON^{hm+ret} model are, therefore, based only on single result analyses.

Although %Ret and Hb_{mass} were measured weekly and fortnightly, respectively, in the present study, only data collected 4 weeks apart were analysed using the Hb^{m} and ON^{hm+ret} models so that a realistic anti-doping testing schedule of 8-12 tests per year was replicated (Zorzoli and Rossi 2010). Too frequent testing can simply cause the expected range to *follow* the path of the test results and reduce sensitivity to suspicious deviations (Ashenden et al. 2011). Accordingly, we included only the first baseline test results, then the results from wk 4, wk 8 and wk 12. Due to the later start of some athletes, there are results from wk 12 for only 7 of the 15 subjects.

For didactic purposes, in this study the sensitivity and specificity of each of the Hb^{m} and ON^{hm+ret} models were assessed at both 99% and 99.9% specificity levels. These levels refer to the probability of a non-doped athlete's result falling outside the expected range (a false-positive result); 99% limits correspond to a 1 in 100 chance of a false-positive occurring, and 99.9% limits correspond to a 1 in 1000 chance of a false-positive occurring. Whilst the WADA guidelines recommend that only results exceeding the 99.9% level in the ABP should be investigated further, they also advocate that individual national anti-doping organisations may choose a lower probability score for their own investigations (e.g. 99% limits),

particularly to identify athletes to be targeted subsequently for more frequent testing (World Anti-Doping Agency 2010).

Results

At the end of the 12-week period of treatment, Hb_{mass} was increased by 11.0 ± 6.0% (mean ± SD) from a baseline of 967 ± 135 g in the rHuEPO group, and by 2.8 ± 1.6% from a baseline of 930 ± 102 g in the Placebo group (Figure 8.1). The changes in %Ret over the course of the 12 week treatment period were very similar between the rHuEPO and Placebo groups: an increase from baseline (0.63 ± 0.21 and 0.62 ± 0.14 in rHuEPO and Placebo groups, respectively) by ~20% at week 4, with a decrease back to near-baseline levels by week 12 (Figure 8.1). Over the 12-week period, the within-subject variability of Hb_{mass} was higher in the rHuEPO group (5.2% CV) than in the Placebo group (3.2% CV).

The $Hb^{m\ (Prommer\ –\ small)}$ model recorded the highest sensitivity of all the models examined (Table 8.1), correctly flagging the results of seven out of ten athletes from the rHuEPO group at the 99% level, five of whom were also flagged at the 99.9% level. The sensitivity of the $Hb^{m\ (Prommer\ –\ large)}$ model was lower, flagging the results of three athletes from the rHuEPO group at the 99.9% level. The sensitivities of the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models were similar, each detecting two doped athletes at the 99.9% level. The ON^{hm+ret} model was not sensitive to microdose rHuEPO doping, failing to flag the results of any doped athlete.

In general, the sequence analysis flagged fewer athletes than the analysis of each single result. However, one rHuEPO athlete was flagged by the sequence analysis of the $Hb^{m\ (Prommer\ -\ small)}$ model at the 99.9% level without any single result being flagged in that athlete at the same level.

All models maintained 100% specificity at the 99.9% level and the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Eastwood)}$ and ON^{hm+ret} models also recorded 100% specificity at the 99% level. Both the $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$ models recorded false-positive results at the 99% level.

The Hb_{mass} results of two exemplar athletes are shown in Figure 8.2 for illustration of the difference in the expected ranges between the two Hb^{m} models that used the smallest and largest estimates of BioWS variance, $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Eastwood)}$ models, respectively.

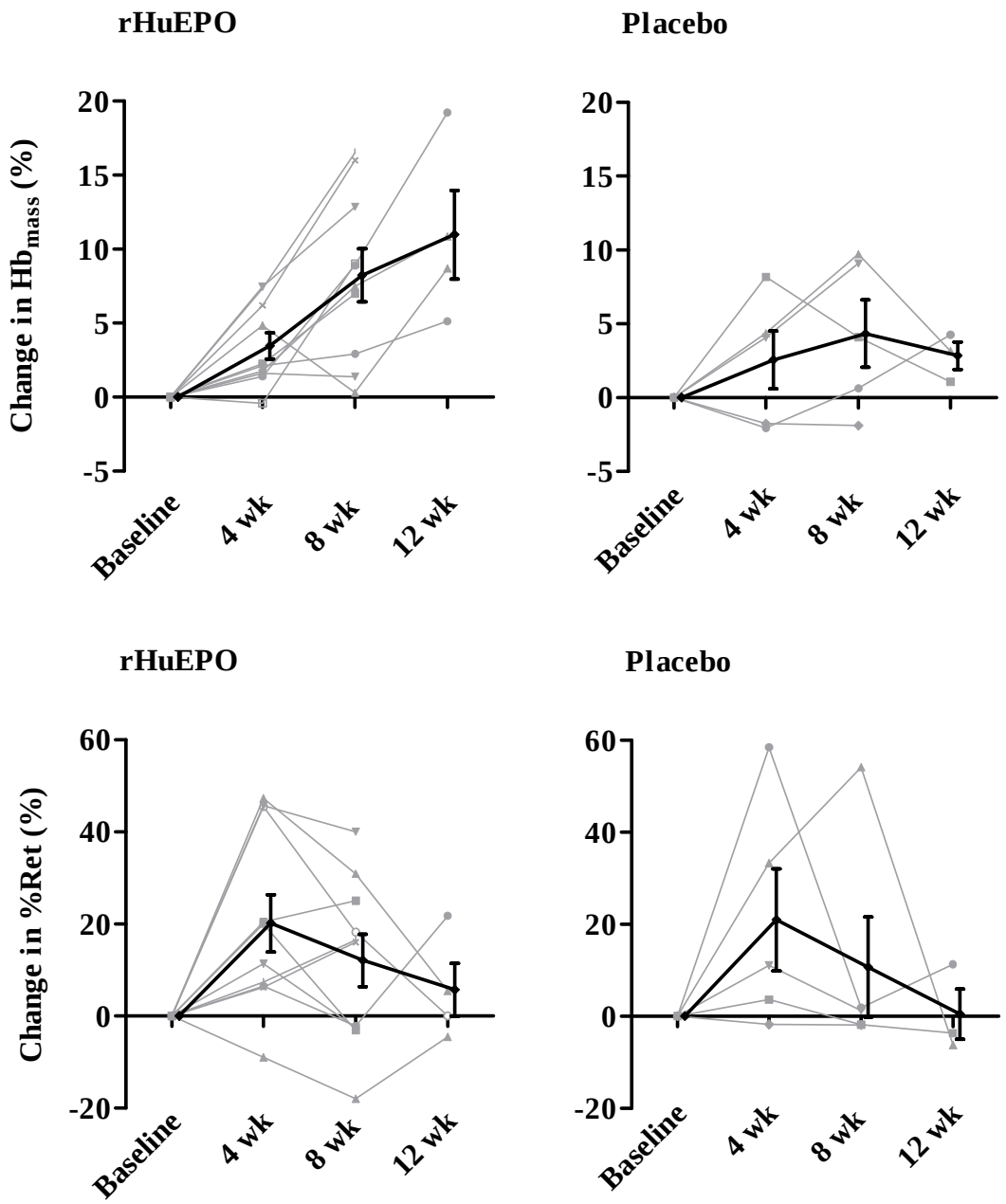

Figure 8.1: Percent changes in haemoglobin mass (Hb_{mass}) and percent reticulocytes (%Ret) over 12 weeks of a microdose recombinant human erythropoietin (rHuEPO) or Placebo regimen. Grey data points show individual athlete changes in rHuEPO (n=10) and Placebo (n=5) conditions, with black lines and error bars signifying group mean results ± SD, slightly offset for clarity.

Table 8.1: Number of athletes flagged by $Hb^{m\ (Prommer\ -\ small)}$, $Hb^{m\ (Prommer\ -\ large)}$, $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Eastwood)}$ and ON^{hm+ret} models during 12 weeks of a recombinant human erythropoietin (rHuEPO) or Placebo microdosing regimen.

Model	Analysis	No. of rHuEPO athletes flagged (n=10) #		No. of Placebo athletes flagged (n=5)		Sensitivity (%)		Specificity (%)	
		99%	99.9%	99%	99.9%	99%	99.9%	99%	99.9%
$Hb^{m\ (Prommer\ –\ small)}$	Single result	7	4	3	0	70	40	60	100
	Sequence	7	3	3	0	70	30	60	100
	Total	**7**	**5**	**3**	**0**	**70**	**50**	**60**	**100**
$Hb^{m\ (Prommer\ –\ large)}$	Single result	4	3	1	0	40	30	80	100
	Sequence	2	2	0	0	20	20	100	100
	Total	**4**	**3**	**1**	**0**	**40**	**30**	**80**	**100**
$Hb^{m\ (Pottgiesser)}$	Single result	2	2	0	0	20	20	100	100
	Sequence	2	2	0	0	20	20	100	100
	Total	**2**	**2**	**0**	**0**	**20**	**20**	**100**	**100**
$Hb^{m\ (Morkeberg)}$	Single result	2	2	0	0	20	20	100	100
	Sequence	2	2	0	0	20	20	100	100
	Total	**2**	**2**	**0**	**0**	**20**	**20**	**100**	**100**
$Hb^{m\ (Eastwood)}$	Single result	2	2	0	0	20	20	100	100
	Sequence	2	1	0	0	20	10	100	100
	Total	**2**	**2**	**0**	**0**	**20**	**20**	**100**	**100**
ON^{hm+ret}	**Single result**	**0**	**0**	**0**	**0**	**0**	**0**	**100**	**100**

n=8 for sequence analysis in rHuEPO group.

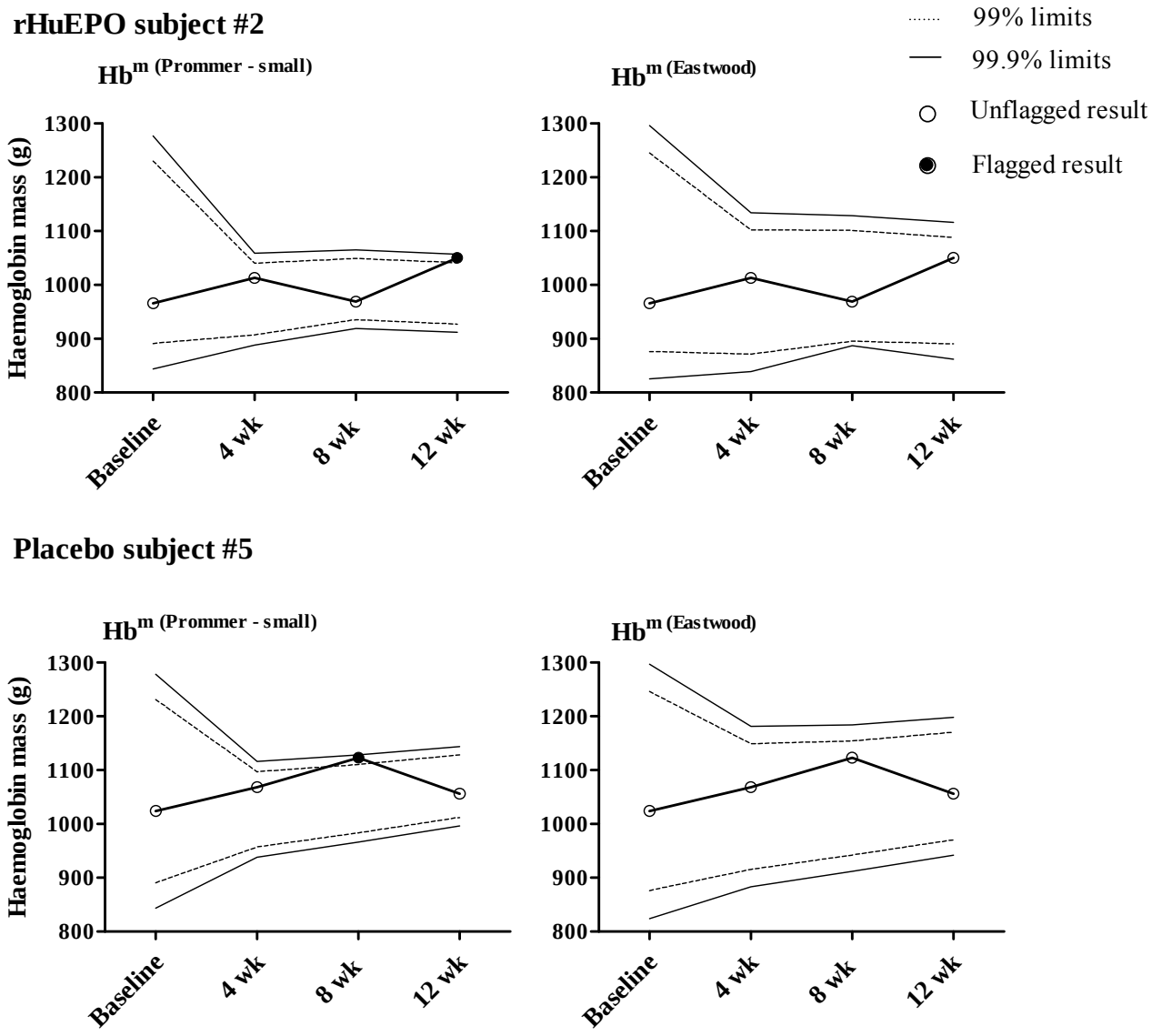

Figure 8.2: Exemplar graphs of two athletes, rHuEPO subject #2 and Placebo subject #5, illustrating the differences between two models, $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Eastwood)}$, in the expected ranges and the incidence of flagged results. Open circles indicate unflagged test results, whilst closed circles indicate results flagged at the 99% level.

Discussion

The use of Hb_{mass} alone as a variable in the Adaptive model of the ABP resulted in sensitivities of between 20% and 70% to rHuEPO microdosing, with the estimate of BioWS variance incorporated in the model heavily influencing the rate of detection. At the specificity

level currently recommended by WADA (99.9%), the $Hb^{m\ (Pottgisser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models each detected two doped athletes, the $Hb^{m\ (Prommer\ -\ large)}$ model detected three, and the $Hb^{m\ (Prommer\ –\ small)}$ detected five doped athletes. In contrast, a previous publication from this study (Ashenden et al. 2011) described the failure of the existing markers of the ABP ([Hb] and OFF-Hr score) to detect doping in any of the ten subjects. The incorporation of % Ret with Hb_{mass} did not lead to improvements in rates of detection, with the ON^{hm+ret} model also failing to flag the results of any of the doped athletes.

If Hb_{mass} was to be incorporated into the ABP, serious consideration would need to be given to deciding which model should be used, in order to protect innocent athletes from false accusations whilst optimising sensitivity. There is an inverse relationship between the sensitivity and specificity of doping detection models. In this investigation, the differences in sensitivity and specificity between the Hb^{m} models were due solely to the different BioWS variance values used in each model. The $Hb^{m\ (Prommer\ –\ small)}$ and $Hb^{m\ (Prommer\ –\ large)}$ models, which used the two smallest estimates of BioWS variance, recorded the highest sensitivities but also the lowest specificities of all the models examined. The WADA currently uses the 99.9% limits in the ABP to flag abnormal blood results and at this level, all the models examined in the present study maintained 100% specificity. On this basis, the best Hb^{m} model to use in the ABP would be the $Hb^{m\ (Prommer\ –\ small)}$ model, due to its superior sensitivity. However, the 99% level false-positive results recorded by the $Hb^{m\ (Prommer\ –\ small)}$ and $Hb^{m\ (Prommer\ –\ large)}$ models suggest that the BioWS variance values used in these models may be too small. In order to get a true indication of the suitability of these models for inclusion in the ABP, a more thorough examination of the models' specificities would need to be conducted in a larger group of non-doped athletes. If future investigations reveal that the specificities of the $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$ models are too low, the 100% specificity records of

the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models suggest that any of those models could be used in the ABP instead, albeit with lower sensitivity.

The sensitivity and specificity results of the models should not be the only consideration when the decision is made about which estimate of BioWS variance to include in the Hb^m model. It is important that the BioWS variance accurately reflects the true biological variance of Hb_{mass} in the elite athlete population. The five estimates of BioWS variance of Hb_{mass}, 56.25 g^2 (Prommer et al. 2008), 244 g^2 (Prommer et al. 2008), 550 g^2 (Pottgiesser et al. 2012), 611 g^2 (Morkeberg et al. 2011) and 830 g^2 (Eastwood et al. 2011b) were each obtained using a different population of athletes. One key difference between the estimates was the number of observations upon which they were based: 128 for Prommer et al. (2008), ~90 for Pottgiesser at al. (2012), 186 for Morkeberg et al. (2011) and ~900 for Eastwood et al. (2011b). Morkeberg et al. (2011) noted the need for studies with a greater number of participants (> 186 observations) to more precisely determine the BioWS variance of Hb_{mass}. The higher sensitivity of the $Hb^{m\ (Prommer\ -\ small)}$ model makes it an appealing option when compared to the other Hb^m models. However, the estimate of BioWS variance included in the $Hb^{m\ (Eastwood)}$ model was based on the greatest number of Hb_{mass} observations and is, therefore, more likely to be a more accurate estimate of the true BioWS variance. There is a substantial difference between the estimates of BioWS variance included within the $Hb^{m\ (Prommer\ –\ small)}$ and $Hb^{m\ (Eastwood)}$ models (56.25 g^2 versus 830 g^2). Whilst the smaller estimate may afford greater sensitivity to doping, it seems unlikely that the $Hb^{m\ (Prommer\ –\ small)}$ model could record a high specificity in the population upon which Eastwood et al. (2011b) based their estimate of BioWS variance. When the precursor to the ABP, the third generation model for detection of blood manipulation, was first described (Sharpe et al. 2006), the authors faced a similar dilemma. They decided that the largest estimate of within-subject variance for [Hb] and OFF-hr would be incorporated into that model to make it less likely that non-doped athletes would

exceed the thresholds. Therefore, erring on the side of caution, $Hb^{m\ (Eastwood)}$ may be the most appropriate model to use because it includes an estimate of BioWS variance that was based on a large number of observations.

Multi-parameter models (those that integrate two or more blood markers) have previously been described as the most sensitive indirect models for blood doping detection (Parisotto et al. 2000b; Sottas et al. 2006). However, the combination of %Ret and Hb_{mass} as used in the ON^{hm+ret} model did not contribute any additional sensitivity to detecting microdose rHuEPO doping compared to current markers of the ABP, and performed worse than when Hb_{mass} alone was used. In response to a moderate dose of rHuEPO, %Ret is typically increased by ~100% for up to 6 weeks (Connes et al. 2004; Lundby et al. 2007; Parisotto et al. 2000a) as the hormone stimulates a greater number of young red blood cells to enter the athlete's circulation. In contrast, the group mean increases in %Ret in the present study for both the rHuEPO and Placebo groups were similar: +20% after 4 weeks of treatment before returning to baseline at 12 weeks. This reduced perturbation of %Ret during rHuEPO treatment is testament to skilful manipulation of the dosage and highlights the advantage of microdosing for dishonest athletes. It is possible that a different combination of Hb_{mass} and %Ret may still hold promise for rHuEPO detection. However, the muted change in %Ret, and the fact that after 12 weeks %Ret levels were back to baseline in the rHuEPO group, suggests that these models may not be as useful as previously proposed, whatever combination of variables is used.

The 20-70% sensitivities of the Hb^{m} models reported here are an improvement upon the 0% sensitivity of the current markers of the ABP that were reported previously (Ashenden et al. 2011). However, it is yet to be determined whether these rates are high enough for the models to be considered worthwhile. This is a complex issue since there is no absolute level of sensitivity that an anti-doping model must achieve before it can be implemented. In general,

the authors of studies who have reported sensitivities of various anti-doping methods exceeding 70% tend to write positively about the method (Parisotto et al. 2000b; Sharpe et al. 2006) whilst those with sensitivities <50% are generally critical of the method tested (Borno et al. 2010; Lundby and Robach 2010). Consequently, it is likely that even the 50% sensitivity of the $Hb^{m\ (Prommer\ -\ small)}$ model may not be considered worthwhile despite it being an improvement upon the 0% sensitivity of the existing ABP in this same group (Ashenden et al. 2011).

The use of the 99% limits to flag changes in blood parameters to be followed-up by target testing has been suggested previously (Sharpe et al. 2006; World Anti-Doping Agency 2010). Target testing could take the form of extra Hb_{mass} tests for use in the ABP, however, if test frequency is too high the expected range may *follow* the changes in the test result and sensitivity may be decreased (Ashenden et al. 2011). Alternatively, flagged Hb_{mass} results at the 99% level could be followed up by direct testing for rHuEPO in urine, which may result in prosecution of athletes who would otherwise go unpunished. Athletes may still evade detection if they have used microdoses of rHuEPO that reduce the window for detection in urine (Ashenden et al. 2006). In theory, any follow-up direct testing on innocent athletes should result in their vindication.

In the present study, microdose rHuEPO doping led to a mean Hb_{mass} increase of 8.2% in 8 weeks, which progressed to 11.0% in athletes who completed 12 weeks of the regimen. There was a notable inter-individual response, with Hb_{mass} increases from baseline ranging from 1.4 to 19.2% within the rHuEPO group over the course of treatment. Lundby and Robach (2010) reported a similarly variable Hb_{mass} response to low dosage rHuEPO treatment with individual responses ranging from 3% to 20% over 13 weeks of treatment. Eastwood et al. (2011b) correctly predicted that using their estimates of within-subject variance, the sensitivity of an Adaptive model using Hb_{mass} would be limited to changes of ~20%; in this study, the only two

athletes flagged by the $Hb^{m\ (Eastwood)}$ model at the 99.9% level had increased 19.8% and 18.4% from baseline and athletes whose Hb_{mass} had increased by up to 16.5% went undetected. The $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Morkeberg)}$ models experienced the same limits to their sensitivity. A 16.5% increase in Hb_{mass} would, theoretically, lead to a ~12% increase in VO_{2max} (assuming a Hb_{mass} of 950 g and VO_{2max} of 4.5 L) (Schmidt and Prommer 2010), and whilst differences in VO_{2max} are not a good predictor of endurance performance in homogenous groups (Faria et al. 2005), such a large change within an individual athlete would cause a substantial improvement in endurance performance (Brien and Simon 1987; Ekblom and Berglund 1991). For such large changes to go undetected would be extremely disappointing and yet may be the necessary reality if future investigations reveal that the specificities of the $Hb^{m\ (Prommer - small)}$ and $Hb^{m\ (Prommer - large)}$ models are not adequate. It may, therefore, be unrealistic for Hb_{mass} to be used as a single variable in the ABP (Eastwood et al. 2011b) but in combination with other markers of altered erythropoeisis the sensitivity may be improved. In the present study, the combination of Hb_{mass} and %Ret in the ON^{hm+ret} model failed to offer any additional sensitivity to microdose rHuEPO doping, mainly due to the negligible %Ret response to this form of doping. However, a model that combined Hb_{mass} and %Ret in a different way flagged 7 of 11 doped subjects in a recent study investigating the sensitivity of Hb_{mass} models to autologous blood doping (Pottgiesser et al. 2012). Previous multi-parameter models that have been created for use in indirect blood monitoring models have combined five (Parisotto et al. 2000a) or seven (Sottas et al. 2006) markers (such as haematocrit, endogenous EPO concentration, mean cell volume) to find the most sensitive combination. In order to find the best model in which to include Hb_{mass}, it may be necessary to investigate the optimal combination with some of these other blood markers, rather than just %Ret.

The Hb_{mass} in the Placebo group also showed notable individual changes (~8-10%) from baseline, larger than those to which our research group is accustomed to seeing in control

subjects. The coefficient of variation in the Placebo group of 3.2%, is larger than one previous estimate of 2.1% CV over 100 days (Eastwood et al. 2008) but is not unreasonably large given that within-subject variance in Hb_{mass} increases over time, from ~2% when measures are made a few days apart to ~4% when they are made several months apart (Eastwood et al. 2011b). One possible source of additional variation is unreliable measurement techniques. All Hb_{mass} measures in the present study were carried out by a single experienced investigator who usually records a TE for the CO re-breathing technique of <2% for measures taken a few days apart (Gough et al. 2011; Gough et al. 2012). The larger TE of the present study, 2.8%, was due to the inclusion of one athlete whose baseline Hb_{mass} readings differed by 120 g. Without including the baseline results of this individual athlete, the TE of Hb_{mass} in this study would have been 1.7% (1.2 to 2.5%; 90% CL), which is in line with values recorded in other studies (Pottgiesser et al. 2012) and was the TE used in all the Hb^{m} models in this investigation. Therefore, we are confident in the reliability of the measurement techniques used in the present study. Although it is clear that one of the baseline tests that differed by 120 g contained an unusual deviation, the experimenter was not able to identify any obvious reason for the discrepancy (e.g. there were no CO leaks recorded around the mouth or nose). In an anti-doping setting where double tests are unlikely, without a reason to discredit the results of a test, all Hb_{mass} values would be included for analysis in the ABP. Therefore, we considered that despite the anomalous result, it was necessary to replicate real-world practice and include all measurements.

An alternative explanation for the atypical individual variation in the Placebo group could be a true biological increase in Hb_{mass} as a result of training. Although participants were asked to maintain their normal training routine throughout the study, it is possible that the motivation involved with participation in a research study resulted in an increased training effort. A 10% increase in training load has previously been associated with a 1% increase in Hb_{mass}

(Garvican et al. 2010c), and the Placebo group mean Hb_{mass} increase of 2.8%, along with some large individual increases in %Ret at 4 weeks could suggest that training did indeed stimulate an unexpected erythropoietic response within the Placebo group (Ashenden et al. 1999).

Limitations

Time restrictions prevented us from continuing to monitor the subjects' blood in the weeks following the cessation of treatment. Continuing to do so enhances detection rates in the current ABP because the OFF-hr model is sensitive to the combination of increased [Hb] and decreased %Ret (Pottgiesser et al. 2011). Following cessation of microdose rHuEPO doping in the present study it is possible that %Ret would have reduced below baseline values as a homeostatic mechanism to restore the athletes' Hb_{mass} to baseline values. In this case, a new model combining Hb_{mass} and %Ret in a different way may have detected doping in the rHuEPO athletes. Two studies have demonstrated that using OFF models, combining high Hb_{mass} with low %Ret, contribute additional sensitivity to autologous doping compared to Hb_{mass} models alone (Morkeberg et al. 2011; Pottgiesser et al. 2012).

It is optimal for ABP models to compare the variability of an athlete's blood results to normative values in two different ways: each single value alone and the entire results sequence as a whole (Pottgiesser et al. 2011; Sottas et al. 2010). A sequence analysis of [Hb] and OFF-hr results caught one doping subject who otherwise would have gone undetected in a recent study examining the sensitivity of these ABP variables to autologous blood transfusion (Pottgiesser et al. 2012). In the present study, the sequence analysis of the $Hb^{m\ (Prommer - small)}$ model also detected one doped athlete who would not have been detected using the single results analysis alone. We were not able to analyse the variability of the sequence of ON^{hm+ret}

results and it remains unknown whether this method would have led to an increased rate of detection for the model integrating Hb_{mass} and %Ret.

Conclusion

The use of Hb_{mass} in an Adaptive model of the ABP resulted in improved sensitivity to microdose rHuEPO doping compared to existing ABP variables, but the combination of Hb_{mass} and %Ret did not provide any additional benefit. For a given specificity, sensitivity was heavily influenced by the estimate of BioWS variance of Hb_{mass} included in the Adaptive model. The $Hb^{m\ (Prommer\ –\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$ models correctly identified 50% and 30% of the doped athletes, respectively, but their false-positive results at the 99% level suggested that the BioWS variance values included in these models may be too small to be used reliably in a wider population. The larger estimates of BioWS variance included in the $Hb^{m\ (Pottgisser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models may be more appropriate but resulted in sensitivities of just 20%, failing to detect Hb_{mass} increases as large as 16.5%. It would be beneficial for future investigations to consider whether the combination of Hb_{mass} with a wider selection of blood markers may enhance the sensitivity of the ABP above that of Hb_{mass} alone.

Acknowledgements

We would like to express our sincerest gratitude to the athletes who volunteered to participate in this research study. We also thank Pierre-Edouard Sottas for allowing us to use the ABP software and for the guidance he offered. Our thanks also go to Melissa Arkinstall and Kiara Johnson for the expert organisational skills they contributed to the project.

Chapter 9

Specificity of haemoglobin mass in the Athlete Biological Passport

Introduction

The Athlete Biological Passport (ABP) tracks changes in the biological characteristics of individual athletes over time and 'flags' unusual variation in these markers, that are suggestive of an athlete using illegal performance-enhancing methods. The haematological module of the ABP is concerned with detection of blood doping and monitors changes in an athlete's haemoglobin concentration ([Hb]), percent reticulocytes (%Ret - the proportion of red blood cells in a sample that are immature) and an integrated measure of these two markers, the OFF-hr Score (World Anti-Doping Agency 2010). In recent years, a number of researchers have suggested that modification of the ABP to include a new marker, total haemoglobin mass (Hb_{mass}), would increase the detection rate of the ABP (Garvican et al. 2010b; Prommer et al. 2008). An abnormal increase in Hb_{mass} can be indicative of an athlete having infused additional blood into their circulation or having used erythropoiesis-stimulating agents (ESAs), such as recombinant human erythropoietin (rHuEPO). An abnormal decrease in Hb_{mass} can be indicative of an athlete having had blood removed for the purpose of reinfusion at a later date.

The main argument for the inclusion of Hb_{mass} in the ABP is that it would be an improvement upon the existing markers ([Hb] and OFF-hr score) that are susceptible to failure in cases of plasma volume manipulation and acute autologous blood doping (Morkeberg et al. 2011; Sanchis-Gomar et al. 2010b). However, the capacity for Hb_{mass} to successfully differentiate between non-doped and doped athletes has been called into question (Eastwood et al. 2011b;

Lundby and Robach 2010). Although Hb_{mass} is relatively stable in athletes compared with the large changes that can be induced by blood doping, larger-than-normal changes in non-doped athletes or smaller-than-normal changes in doped athletes have the potential to reduce the sensitivity (rate of correct detection of doped) and specificity (rate of correct detection of non-doped) of an ABP model based on Hb_{mass}. In Chapter 8, I demonstrated that the sensitivities of Hb_{mass} models to microdose rHuEPO doping ranged from 20% to 50%, but that the specificities of these models, in particular, need to be tested in a larger population of non-doped athletes. The current thesis has quantified variations in Hb_{mass} resulting from various forms of hypoxic exposure, illness, injury, inter-laboratory testing, ultra-endurance triathlon racing and microdoses of recombinant human erythropoietin (rHuEPO). Collectively, these data constitute an important test-case for sensitivity and specificity of Hb_{mass} models in the ABP, specifically focussing on scenarios that have the potential to blur the line between doped and non-doped athletes.

In this chapter of the thesis, I integrate the data collected during my candidature with additional data from four studies published by other researchers (Lundby and Robach 2010; Robertson et al. 2010b; Robertson et al. 2010c; Saunders et al. 2010a) in order to examine the feasibility of including Hb_{mass} in the ABP with a wider perspective than my research alone would permit. The analysis includes consideration of three additional factors that may influence the sensitivity and specificity of Hb_{mass} models used in the ABP: the choice of biological within-subject (BioWS) variance of Hb_{mass} included in the model, the frequency of Hb_{mass} testing and the utility of combining Hb_{mass} with %Ret in a model.

For each blood marker, the Adaptive model of the ABP uses an individual's past test results in conjunction with knowledge of the normal variation of the marker to define a probability range in which the athlete's next test result is expected to fall. For the current haematological markers of the ABP ([Hb], %Ret and OFF-hr score), when the first result from an individual

athlete is entered into the Adaptive model, the expected range is wholly determined by three values derived from an elite athlete population: the population mean, the between-subject variance and the within-subject variance. As more results are entered, the expected range becomes progressively less dependent on the population mean and between-subject variance, and more dependent on the within-subject variance (Sottas et al. 2010). For Hb_{mass}, a 'two-levels' error model that breaks down within-subject variance into analytical and biological sub-components is more appropriate than a single value of within-subject variance (Pottgiesser et al. 2012; Prommer et al. 2008). As a result, separate estimates of Typical Error (TE; the analytical standard deviation (SD), expressed as a percentage) and BioWS variance are entered into the Adaptive model for Hb_{mass}. Studies examining the within-subject variance of Hb_{mass} have demonstrated closer agreement between their estimates of TE (1.4% - 2.0%) than between their estimates of BioWS variance (56 g^2 – 830 g^2, equivalent to ~0.8% - 2.8%). The choice of the BioWS variance value to be used in the Adaptive model for Hb_{mass} heavily influences the sensitivity and specificity of the model, and must be considered in debates about the inclusion of Hb_{mass} in the ABP. Five different estimates of the BioWS variance of Hb_{mass} have been published in recent years (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008) and the sensitivities and specificities of Adaptive models based on each estimate are compared here: the first factor included in the analyses within this chapter.

In 2010 and 2011, elite cyclists were tested for the ABP on average five times per year by their international governing body, the Union Cycliste Internationale (Rossi et al. 2012). As well as the individual sports federations, national anti-doping organisations and the World Anti-Doping Agency (WADA) also schedule tests on elite athletes, so the frequency with which athletes are tested varies. Whilst it is important to assess anti-doping models using a realistic testing schedule, it is also worthwhile investigating whether the testing frequency

affects the sensitivity and specificity of Adaptive models. It has been reported that if athletes are tested too frequently the sensitivity of the model will be reduced because the expected range will simply *follow* the increases or decreases in an athlete's results (Ashenden et al. 2011). On the other hand, it seems intuitive that more frequent testing would increase the likelihood of a doped athlete being flagged because of the higher chance that a test would be conducted close to the time when the athlete has engaged in doping. Consequently, the influence of test frequency on sensitivity and specificity of Hb_{mass} models is the second factor included in these analyses.

The final area of analysis is an investigation into the utility of combining Hb_{mass} with another blood marker, %Ret, as an alternative to models based on Hb_{mass} alone. In the anti-doping setting, it has previously been demonstrated that multi-parameter models (e.g. OFF-hr Score) can be more sensitive than single parameters alone, because the combination of small changes in multiple parameters can be a strong indicator of blood doping (Parisotto et al. 2000a). Using %Ret in multi-parameter analyses is particularly beneficial for detection of blood doping because %Ret is sensitive to changes in the rate of erythropoeisis but unaffected by perturbations of plasma volume. An increase in %Ret can indicate recent injection of rHuEPO or blood withdrawal for the purpose of storage prior to autologous blood doping. Conversely, a decrease in %Ret can indicate recent cessation of rHuEPO usage or reinfusion of autologous blood. Two research studies (Morkeberg et al. 2011; Pottgiesser et al. 2012) have demonstrated experimentally that a combined Hb_{mass} and %Ret model was more sensitive to autologous blood doping than a model based on Hb_{mass} alone and both recorded a specificity of 100% (i.e. no false positive results in non-doped control subjects). In contrast, the results of Chapter 8 of this thesis suggested that a combined Hb_{mass} and %Ret model was not sensitive to microdose rHuEPO doping. In this chapter, I further explore the sensitivity of a combined

Hb_{mass} and %Ret model to rHuEPO doping and assess the specificity of this model in a population of non-doped athletes with a higher-than-normal variation in Hb_{mass}.

Methods

Study design

The sensitivity and specificity of six different Adaptive models were examined using longitudinal Hb_{mass} results of 159 athletes (100 male, 59 female; on average ~6 Hb_{mass} results per athlete) sourced from ten research studies. Five of the Adaptive models used Hb_{mass} as a single marker and one model used a combination of Hb_{mass} and %Ret. Two separate analyses of the data were conducted: the first included all available results for each individual athlete and the second included only test results that were separated by at least 21 days.

Sources of the data

The longitudinal Hb_{mass} results of individual athletes were sourced from ten research studies, six of which comprise the earlier chapters of this thesis and the other four having been published by various researchers in the past 3 years (Lundby and Robach 2010; Robertson et al. 2010b; Robertson et al. 2010c; Saunders et al. 2010a). A brief description of each study is outlined in Table 9.1, but individual publications should be consulted for detailed descriptions of the respective protocols. Data from nine of the ten studies were included in the analyses for specificity of the Adaptive models as they included non-doped subjects. Only data from two of the ten studies allowed calculations for model sensitivity since they were the only ones that included doped subjects. In four of the ten studies, measures of %Ret were made on the same day as Hb_{mass}, and these data were used in a model combining these two markers, the ON^{hm+ret} model. Hb_{mass} was measured using the optimised CO re-breathing technique (Prommer and Schmidt 2007) in all studies except one (Lundby and Robach 2010) where Hb_{mass} was measured using an earlier version of the CO re-breathing technique (Burge and Skinner

1995). In three of the studies where %Ret was measured, the blood analyser was the Advia 120 Hematology Analyzer (Bayer Diagnostics, Tarrytown, NY, USA), and in the fourth study (Chapter 8 of this thesis) a Sysmex XE-2100 analyser (Sysmex Corporation, Kobe, Japan) was used.

Adaptive models

The sensitivity and specificity of five Adaptive models using Hb_{mass} as a single marker (Hb^{m} models) were examined using the ABP software. The only difference between each of the Hb^{m} models was the BioWS variance value included in their calculations. The BioWS variances included in the Hb^{m} models were sourced from four published studies that each estimated the within-subject variance of Hb_{mass} in athletes (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008). The Hb^{m} models were named after the first author of each study, with the exception of the $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$ models where two estimates of BioWS variance were derived from the same study (Prommer et al. 2008). Refer to Appendix 1 for a detailed description of how each of the estimates of BioWS variance was derived.

Of the four studies from which estimates of BioWS variance were derived, the study by Eastwood et al. (2011b) was the only one to describe separate estimates for males and females. Two other studies used only male subjects (Morkeberg et al. 2011; Pottgiesser et al. 2012), and the final study published single values of BioWS variance and within-subject variance to characterise the variation in both their male and female subjects (Prommer et al. 2008). Consequently, for the purposes of this investigation, the $Hb^{m\ (Eastwood)}$ model used separate estimates of BioWS variance for males (830 g^2) and females (573 g^2), whilst the other Hb^{m} models used single estimates for both sexes: 56.25 g^2 for $Hb^{m\ (Prommer\ -\ small)}$, 244 g^2 for $Hb^{m\ (Prommer\ -\ large)}$, 550 g^2 for $Hb^{m\ (Pottgiesser)}$ and 611 g^2 for $Hb^{m\ (Morkeberg)}$. To enable a fair

comparison, the population mean, between-subject variance and TE values were kept consistent between all five Hb^m models. The population mean (11.84 x body mass kg +149 g for males, or 11.84 x body mass kg -174 g for females) and between-subject variance (3994 g^2) used in the Hb^m models were previously described by Prommer et al. (2008), and the TE (1.7%) is the average analytical error from the four key studies (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008). For all of the Hb^m models, each individual Hb_{mass} result as well as the variability of the full sequence of results were examined. Individual results that fell outside the expected range were flagged, as were sequences with an abnormally high variability (Sottas et al. 2010).

The results from the four studies that included measures of both Hb_{mass} and %Ret were analysed using a sixth, novel Adaptive model, the ON^{hm+ret} model, which combined the two markers according to the following calculation:

$$ON^{hm+ret} = \sqrt{\%Ret} + 2 \times \ln(Hb_{mass})$$

For a description of how the ON^{hm+ret} model was created, see Chapter 8 of this thesis. In brief. this combination of markers is not currently included in the ABP and, therefore, the ABP software could not be used for the ON^{hm+ret} model. Instead, the analyses for the ON^{hm+ret} model were conducted using calculations described in previous literature (Ashenden et al. 2011; Morkeberg et al. 2011; Sottas et al. 2010). These calculations are based on Bayesian network statistics and simulate the Adaptive model included in the ABP software. I verified the accuracy of these calculations using examples of [Hb] results from 3 subjects, comparing a series of upper and lower limits from my calculations to those generated by the Adaptive model in the ABP software, and obtained identical figures. The population mean and between-subject variance of 14.64 and 0.2314, respectively, were used for all subjects, whilst different values of within-subject variance were used for males (0.0169) and females

(0.0256). Unlike the Hb^m models, the within-subject variance values used in the ON^{hm+ret} model were not split into analytical and biological sub-components, because the calculations described in previous literature did not include these components separately. The calculations used for the ON^{hm+ret} model allowed each individual result to be compared against the individualised reference ranges, but it was not possible for a sequence analysis to be completed. Therefore, the sensitivity and specificity results of the ON^{hm+ret} model are based only on the analysis of individual results.

The completion of two separate analyses (All and Fewer) allowed the effect of testing frequency on the sensitivities and specificities of each model to be examined. In the 'All Tests' analysis, all available results from the ten studies were entered into each of the models. This included a total of 988 observations in 159 athletes. In contrast, the 'Fewer Tests' analysis included the first result recorded for each athlete, followed chronologically by any subsequent results that were separated by at least 21 days. This included 382 observations in 121 athletes (subjects from the Inter-laboratory and Racing studies were not included in the Fewer tests analysis because all measurements in these two studies were made within 21 days). The 'Fewer tests' analysis is a more realistic test schedule and did not include any double measures (made a few days apart) that were typically included at baseline in a number of the research studies.

Sensitivity and specificity calculations

In order to protect the rights of innocent athletes, the WADA has chosen a moderately high specificity level for the markers in the ABP: a 1 in 1000 chance of a non-doped athlete's result being flagged (99.9% specificity). For the purposes of gathering information on athletes who may be subsequently targeted with more frequent blood testing and/or testing of their urine for ESAs (Lasne et al. 2009), individual sports federations or national anti-doping

organisations may choose to reduce the specificity level to a 1 in 100 chance (99%) (World Anti-Doping Agency 2010). The sensitivities and specificities of each of the six models were assessed at both the 99% and 99.9% levels. Sensitivity was calculated as the number of athletes who recorded at least one flagged result, expressed as a percentage of all the athletes analysed. Specificity was calculated as the number of non-doped athletes that were not flagged, expressed as a percentage of all the athletes analysed. As in Chapter 8, any flagged results that occurred on the first test for any athlete were discounted because in all instances the non-elite athlete's result fell below the lower limit of the expected range for elite athletes. This simply reflects the consequences of judging non-elite athletes against a population mean that is representative of the elite athlete population. Once these athletes' first test results were entered, the Adaptive model of the ABP individualised the expected range for the next result, taking into account their first Hb_{mass} result (Sottas et al. 2010). Therefore, although the first results of these athletes were affected, all subsequent tests were unaffected and could be included in the analysis as they were judged against an individualised mean Hb_{mass}. However, the abnormal first test results caused an artificially high sequence result for these athletes. Consequently, the sequence analysis result for any athlete whose first test result fell outside the limits was excluded from the sensitivity results for all five Hb^{m} models.

Table 9.1: Description of studies included in the analyses of six Adaptive models that include Hb_{mass}.

Name	Reference	Includes %Ret	Description
Studies included in specificity calculations			
Inter-laboratory	Chapter 3 (Gough et al. 2011)	N	Effect of varied test location on Hb_{mass} in 10 recreational athletes, results entered before and after a calculated adjustment was used as a quality control treatment for the data.
Racing	Chapter 4 (Gough et al. In Press)	N	The acute effect of ultra endurance triathlon racing on Hb_{mass} in 18 well-trained triathletes, including 8 control subjects who rested between two measurements of Hb_{mass}.
Injury/Illness	Chapter 5 (Submitted to Int. J. Sports Med.)	N	Longitudinal monitoring of 15 athletes spanning a period of injury or illness, including the influences of Live High:Train Low (LHTL) altitude, iron supplementation, reduced training, surgery and changes in body mass.
LHTL and Classic	Chapter 6 (Gough et al. 2012)	N	Three weeks of LHTL or classical (Classic) altitude training in 27 elite swimmers.
AcIHE and LHTL	Chapter 7 (Submitted to Int. J. Sports Physiol. Perform.)	Y	Three weeks of Acute Intermittent Hypoxic Exposure (AcIHE), LHTL altitude or a Placebo condition in 24 semi-elite and elite triathletes.
LHTL and Placebo	(Saunders et al. 2010a)	Y	Three weeks of LHTL, Placebo LHTL or Control in 17 elite race walkers of national and international standard
LHTL+TH and TH alone	(Robertson et al. 2010c)	Y	Three weeks of LHTL with daily training in hypoxia (TH), compared to sea-level residence with TH in 17 well-trained runners.
LHTL repeat exposure	(Robertson et al. 2010b)	Y	Two three-week periods of LHTL altitude separated by six weeks of sea-level training (or, in control group, sea-level training throughout) in 16 well-trained runners.
Microdose rHuEPO (Placebo)	Chapter 8	Y	Twelve weeks of microdoses of Placebo injection in 5 recreational cyclists.
Studies included in sensitivity calculations			
Microdose rHuEPO (rHuEPO)	Chapter 8	Y	Twelve weeks of microdoses of rHuEPO injection in 10 recreational cyclists.
Boost/Microdose rHuEPO	(Lundby and Robach 2010)	N	Five weeks of boosting rHuEPO dose followed by eight weeks of microdosing in 8 male subjects.

Results

The $Hb^{m\ (Prommer\ -\ small)}$ model was the most sensitive to rHuEPO doping (Table 9.2), but also recorded the lowest specificity, resulting in false-positive results for 47 athletes at the 99% level and 17 athletes at the 99.9% level when all test results were included. The sensitivity of the $Hb^{m\ (Prommer\ –\ large)}$ model was lower (28% versus 72%) and the specificity was higher (97% versus 89%) than the $Hb^{m\ (Prommer\ -\ small)}$ model. The sensitivities and specificities of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models were similar, although the $Hb^{m\ (Morkeberg)}$ model was the only one that maintained 100% specificity at the 99.9% level.

The individual Hb_{mass} results of one non-doped athlete who was flagged at the 99.9% level by the $Hb^{m\ (Prommer\ –\ large)}$, $Hb^{m\ (Prommer\ –\ small)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models are shown in Figure 9.1. This athlete represents the single false-positive result at the 99.9% level for the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models.

The effect of reducing the frequency of testing was to reduce the sensitivities of all the Hb^{m} models to rHuEPO doping. In contrast, a reduction in testing frequency increased the specificity of the models, with one exception; the 'Fewer tests' analysis resulted in one additional false-positive result at the 99% level for the $Hb^{m\ (Pottgiesser)}$ model compared to the 'All tests' analysis.

The ON^{hm+ret} model, which analysed data from fewer studies than the Hb^{m} models, recorded false-positives for 9 athletes at the 99% level and 3 athletes at the 99.9% level in the 'All tests' analysis (Table 9.2). The ON^{hm+ret} model did not flag any doped athletes from the Microdose rHuEPO study, recording the lowest sensitivity of all the models examined.

Table 9.2: The sensitivities and specificities of six Adaptive models at the 99% and 99.9% levels, incorporating separate analyses for the inclusion of all test results and for the inclusion of only results separated by at least 21 days.

(a) Non-doped athletes

Study	No. of athletes	$Hb^{m\ (Prommer\ -\ small)}$ (99% / 99.9%)		$Hb^{m\ (Prommer\ –\ large)}$ (99% / 99.9%)		$Hb^{m\ (Pottgiesser)}$ (99% / 99.9%)		$Hb^{m\ (Morkeberg)}$ (99% / 99.9%)		$Hb^{m\ (Eastwood)}$ (99% / 99.9%)		ON^{hm+ret} (99% / 99.9%)	
		All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests
Injury/Illness	15	6 / 3	5 / 2	2 / 1	2 / 1	1 / 1	1 / 1	1 / 0	1 / 0	1 / 1	1 / 0	N/A[2]	N/A[2]
LHTL and Classic	27	8 / 2	6 / 1	1 / 0	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	N/A[2]	N/A[2]
Inter-laboratory	20[*]	1 / 1	N/A[1]	0 / 0	N/A[1]	0 / 0	N/A[1]	0 / 0	N/A[1]	0 / 0	N/A[1]	N/A[2]	N/A[2]
Racing	18	1 / 0	N/A[1]	0 / 0	N/A[1]	0 / 0	N/A[1]	0 / 0	N/A[1]	0 / 0	N/A[1]	N/A[2]	N/A[2]
AcIHE and LHTL	24	7 / 0	1 / 1	1 / 0	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
LHTL and Placebo	17	7 / 1	4 / 2	1 / 0	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	2 / 0	0 / 0
LHTL +TH and TH alone	17	5 / 2	2 / 1	3 / 1	1 / 0	1 / 0	1 / 0	1 / 0	0 / 0	0 / 0	0 / 0	1 / 1	1 / 0
LHTL repeat exposure	16	10 / 6	6 / 3	5 / 2	2 / 1	0 / 0	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	6 / 2	4 / 1
Microdose rHuEPO (Placebo group)	5	2 / 2	3 / 0	2 / 0	1 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
Totals	**159**	**47 / 17**	**27 / 10**	**15 / 4**	**9 / 2**	**2 / 1**	**3 / 1**	**2 / 0**	**1 / 0**	**1 / 1**	**1 / 0**	**9 / 3**	**5 / 1**
Specificity (%)		**70 / 89**	**83 / 94**	**91 / 97**	**94 / 99**	**99 / 99**	**98 / 99**	**99 / 100**	**99 / 100**	**99 / 99**	**99 / 100**	**94 / 98**	**97 / 99**

(b) Doped athletes

Study	No. of athletes	$Hb^{m\ (Prommer\ -\ small)}$ (99% / 99.9%)		$Hb^{m\ (Prommer\ –\ large)}$ (99% / 99.9%)		$Hb^{m\ (Pottgiesser)}$ (99% / 99.9%)		$Hb^{m\ (Morkeberg)}$ (99% / 99.9%)		$Hb^{m\ (Eastwood)}$ (99% / 99.9%)		ON^{hm+ret} (99% / 99.9%)	
		All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests	All tests	Fewer tests
Microdose rHuEPO (rHuEPO group)	10	8 / 7	7 / 5	7 / 3	4 / 3	4 / 2	2 / 2	4 / 2	2 / 2	3 / 2	2 / 2	0 / 0	0 / 0
Boost/Microdose rHuEPO	8	8 / 6	8 / 6	6 / 2	6 / 2	2 / 1	2 / 1	2 / 1	2 / 1	1 / 1	1 / 1	N/A[2]	N/A[2]
Totals	**18**	**16 / 13**	**15 / 11**	**13 / 5**	**10 / 5**	**6 / 3**	**4 / 3**	**6 / 3**	**4 / 3**	**4 / 3**	**3 / 3**	**0 / 0**	**0 / 0**
Sensitivity (%)		**89 / 72**	**83 / 61**	**72 / 28**	**56 / 28**	**33 / 17**	**22 / 17**	**33 / 17**	**22 / 17**	**22 / 17**	**17 / 17**	**0 / 0**	**0 / 0**

[*] Data analysed for the Inter-laboratory study contained data from 10 subjects that was analysed by all the models twice: before and after a calculated adjustment was applied to the data to make results from different laboratories equivalent. **N/A[1]:** No data were available for the 'Fewer tests' analysis for the Inter-laboratory and Racing studies because all measurements were made within 21 days. **N/A[2]:** Only data from studies in which %Ret measurements were made on the same day as Hb_{mass} measurements could be analysed using the ON^{hm+ret} model.

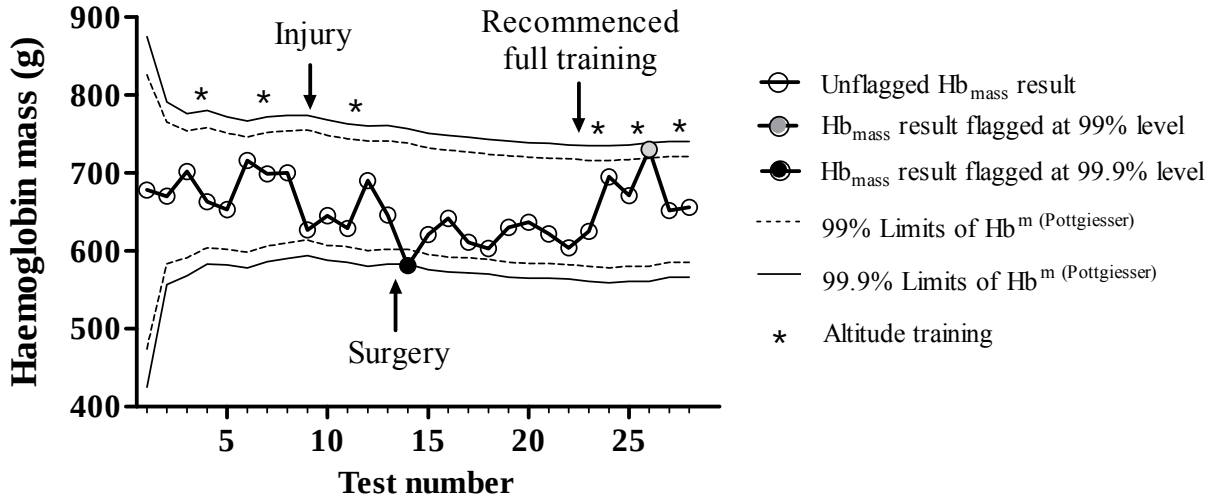

Figure 9.1: An exemplar graph depicting the Hb_{mass} results of one athlete from the injury/illness study along with 99% and 99.9% limits from the $Hb^{m\ (Pottgiesser)}$ model. X-axis represents sequential tests on an ordinal scale over a 25 month period, where the testing frequency was not consistent. (i.e. some consecutive tests separated by days, others separated by months). Grey circles represent results flagged by the $Hb^{m\ (Pottgiesser)}$ model at the 99% level and black circles represent results flagged by the $Hb^{m\ (Pottgiesser)}$ model at the 99.9% level. The lowest Hb_{mass} value, which was measured one month post-surgery, was flagged not only by the $Hb^{m\ (Pottgiesser)}$ model but also by the $Hb^{m\ (Prommer - small)}$, $Hb^{m\ (Prommer - large)}$, and $Hb^{m\ (Eastwood)}$ models at the 99.9% level. No results were flagged at the 99.9% level by the $Hb^{m\ (Morkeberg)}$ model.

Discussion

The results of the present study confirm that the sensitivity and specificity of an Adaptive model in the ABP using Hb_{mass} as a single marker would be heavily influenced by the choice of BioWS variance included in the model. The $Hb^{m\ (Prommer - small)}$ model recorded the highest sensitivity to doping of all the models tested, but also recorded the highest number of false-positive results among non-doped athletes. Conversely, the $Hb^{m\ (Morkeberg)}$ model did not result in a single false-positive at the 99.9% level but sensitivity was much reduced compared to the $Hb^{m\ (Prommer - small)}$ model. The other three Hb^{m} models recorded intermediate values for sensitivity and specificity. The ON^{hm+ret} model was apparently not useful, recording three

false-positives at the 99.9% level in the 'All tests' analysis and zero sensitivity to rHuEPO doping.

Specificity

In order to protect innocent athletes from being falsely accused of doping offences, it is vital that anti-doping authorities ensure that any models for potential inclusion in the ABP have a specificity close to 100%. The WADA has chosen to accept the risk of one false-positive result in one thousand cases. In this sample of 988 tests on 159 non-doped athletes, the 1 in 1000 standard would be equivalent to a maximum of one false-positive result at the 99.9% level for any of the models examined here. However, the $Hb^{m\ (Prommer\ -\ small)}$ model yielded 17 false-positives at the 99.9% level when all tests were used and 10 false-positives when only tests separated by 21 days were included. Changes in Hb_{mass} associated with LHTL altitude, Classic altitude, inter-laboratory measurement error, illness, injury and hypoxic training were all flagged incorrectly by the $Hb^{m\ (Prommer\ -\ small)}$ model. The $Hb^{m\ (Prommer\ -\ large)}$ model recorded four false-positive results at the 99.9% level when all test results were included and two false-positives when fewer tests were included. The ON^{hm+ret} model yielded three and one false-positives in the 'All Tests' and 'Fewer Tests' analyses, respectively, all resulting from Hb_{mass} tests conducted during LHTL altitude training. These levels of false-positives likely make the $Hb^{m\ (Prommer\ -\ small)}$, $Hb^{m\ (Prommer\ -\ large)}$, and ON^{hm+ret} models unsuitable for use in the ABP.

The model with the highest specificity was the $Hb^{m\ (Morkeberg)}$ model, which did not record any false-positives at the 99.9% level. The specificity results of the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models were also in line with WADA's desired target of 1 in 1000, each recording only one false-positive result at the 99.9% level, regardless of whether all tests or fewer tests were analysed. Although some researchers have previously expressed concern for the potential confusion between altitude and doping-induced increases in Hb_{mass} (Eastwood et al. 2011b;

Prommer et al. 2008), no false-positives at the 99.9% level arose from any form of hypoxic exposure in the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ or $Hb^{m\ (Eastwood)}$ models. However, the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models both mistakenly flagged one non-doped athlete whose Hb_{mass} was reduced after surgery and injury-related inactivity. This is a major flaw in the integrity of the Hb^{m} models because injury is a common occurrence in elite athletes and exemptions to prosecution given to injured athletes in the event of them being flagged by the ABP would be open to abuse. On the other hand, WADA routinely records athlete whereabouts information via the Anti-Doping Administration & Management System that could readily be used to identify events such as surgery (Zorzoli and Rossi 2010). Nevertheless, the fact that only one of the fifteen athletes in the injury/illness study flagged the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models at the 99.9% level illustrates the difficulty that anti-doping authorities would have in differentiating between injuries that might or might not be expected to cause large deviations in Hb_{mass}.

One additional non-doped athlete was flagged at the 99% level in the 'All tests' analysis by both the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Morkeberg)}$ models, due to a low Hb_{mass} result after one week of hypoxic training. This results is curious since altitude usually elevates Hb_{mass}, unless sick or injured. In addition, one other non-doped athlete was flagged at the 99% level of the $Hb^{m\ (Pottgiesser)}$ model in the 'Fewer tests' analysis only due to a high variability of their results sequence. National anti-doping authorities and individual sporting federations may choose to target these two athletes for further testing if they use the 99% specificity limits as an intelligence-gathering tool (Zorzoli and Rossi 2012). Although subsequent testing and expert review of the athletes' whole blood profiles may result in their vindication, it is undoubtedly a risk that many athletes would rather not have to take.

Sensitivity

The $Hb^{m\ (Prommer\ -\ small)}$ model was most sensitive to rHuEPO doping, flagging 72% of the doped athletes at the 99.9% level when all tests were used in the analysis. The corresponding sensitivity of the $Hb^{m\ (Prommer\ -\ large)}$ model was 28%, and of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models was 17%. Unlike the '1 in 1000' criterion for specificity, there is no set value for sensitivity that is required for an anti-doping test to be considered suitable for implementation. However, as stated previously, the authors of studies who have reported sensitivities of various anti-doping methods exceeding 70% generally write positively about the method (Parisotto et al. 2000b; Sharpe et al. 2006) and those with sensitivities <50% are generally critical of the method (Borno et al. 2010; Lundby and Robach 2010). Therefore, it is unlikely that any of the models that possess adequate specificity, the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ or $Hb^{m\ (Eastwood)}$ models, would be considered useful for inclusion in the ABP based on their sensitivities. But, it is important to acknowledge that the assessment of sensitivity for the models in this final chapter were based on only two studies and limited to rHuEPO doping only. Elsewhere, Hb_{mass} models have twice demonstrated superior detection sensitivity over existing ABP markers: Morkeberg et al. (2011) found that Hb_{mass} models had the highest sensitivity in the first 48 hours after autologous blood transfusions, and, in a year-long simulation of autologous doping and detection practices, Pottgiesser et al. (2012) described higher detection rates at the 99.9% level for a combination of Hb_{mass} models (8 of 11 doped athletes) than existing ABP models (6 of 11 doped athletes). Together with our findings of three rHuEPO-doped athletes being flagged by the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models who would otherwise have gone undetected (Ashenden et al. 2011), these studies demonstrate that the sensitivity of the ABP would be enhanced by the inclusion of Hb_{mass}. Intelligent testing of Hb_{mass} in close proximity to events when autologous transfusion or donation may have occurred could yield even greater sensitivities than those indicated

above. However, the introduction of Hb_{mass} into the ABP would necessitate worldwide implementation of a new testing procedure, CO re-breathing, which could be challenging since CO is a noxious gas, although not in the quantities used by researchers (Gough et al. 2011; Schmidt and Prommer 2005).

The ON^{hm+ret} model was not sensitive to microdose rHuEPO doping at either the 99% or 99.9% levels in any of the 10 subjects in the microdose rHuEPO study. In that study, despite up to 20% changes in Hb_{mass}, the %Ret deviated very little from baseline (See Chapter 8). By the final week of doping, despite this being the period of highest dosages of rHuEPO, the mean %Ret in the doping athletes had returned to baseline. By combining Hb_{mass} with %Ret, it was hoped that the ON^{hm+ret} model would be more sensitive to small haematological changes that Hb_{mass} alone. However, the muted %Ret response to microdoses of rHuEPO negated any potential benefits of the combined model. Given that the combination of high Hb_{mass} and high %Ret used in the ON^{hm+ret} model would not be suitable to detect autologous blood doping, is seems that this model would not be useful in the ABP, unless athletes were using larger doses of ESAs.

BioWS variance

Due to the influence of BioWS variance on the width of the limits generated by the Adaptive model, a model using a small BioWS variance value would be expected to yield high sensitivity but a low specificity. Conversely, a model using a large BioWS variance would be expected to yield high specificity but a low sensitivity. In general, the results of the present study conformed to those trends; the $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$ models (BioWS variances: 56.25 – 244 g^2) recorded the highest sensitivities and lowest specificities, whilst the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models (BioWS variances: 550 – 830 g^2) recorded lower sensitivities but higher specificities.

Of the three models that displayed identical sensitivities and similarly high specificities, $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$, it is difficult to decide which, if any, would be most appropriate for inclusion in the ABP. Based on a 100% specificity record at the 99.9% level, the $Hb^{m\ (Morkeberg)}$ model may be the most appropriate. However, if the decision were to be based on sensitivity alone, the $Hb^{m\ (Pottgiesser)}$ model may be a better choice. Despite the identical sensitivities at the 99.9% level of all three models in question, if these models were compared in a larger cohort of doped athletes the $Hb^{m\ (Pottgiesser)}$ model should be superior in its detection rate due to its use of a smaller BioWS variance. Indeed, the $Hb^{m\ (Pottgiesser)}$ model displayed a sensitivity of 55% (flagging 6 of 11 doped athletes) to autologous blood doping in a year-long detection and doping study (Pottgiesser et al. 2012). From another perspective, the most appropriate model may be the $Hb^{m\ (Eastwood)}$ model because here the estimate of BioWS variance is likely to be most representative of the true BioWS variance in an elite athlete population. The population from which Pottgiesser et al. (2012) yielded their estimate of BioWS variance was small (n=10), exercised only at a recreational level, and consisted of males only. The cohort that Morkeberg et al. (2011) used for calculating their estimate of BioWS variance was a group of 58 male elite athletes. Although this group was larger and a better representation of the athletes' competitive level to which the ABP is applied, the authors themselves noted the need for a greater number of observations upon which to base a true estimate of Hb_{mass} within-subject variance. The estimate of within-subject variance derived from the most representative cohort is that of Eastwood et al. (2011b), which was based on ~900 observations in 130 elite and semi-elite athletes, including both males and females. Each of the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models, therefore, may be most suitable for inclusion in the ABP, for different reasons.

Although not examined in detail in this investigation, it is important to consider whether the use of a single absolute value for BioWS variance, as opposed to a percentage, may create a

bias in the sensitivity and specificity of models depending on the magnitude of an athlete's Hb_{mass}. For example, there would be no difference between choosing a biological SD of 3% or 30 g in an athlete whose Hb_{mass} is 1000 g (SD = $\sqrt{\text{variance}}$). But applying the 3% SD to an athlete with an Hb_{mass} of 500 g results in a deviation of ± 15 g, a lower absolute SD compared to the athlete whose Hb_{mass} is 1000 g. Conversely, applying the 30 g SD to the athlete with an Hb_{mass} of 500 g results in a deviation of ± 6% (a lower relative SD compared to the athlete with an Hb_{mass} of 1000 g). These approaches obviously yield different outcomes, but it is not clear which approach is most appropriate for characterising the way in which Hb_{mass} varies biologically. Currently, the ABP software requires the BioWS variance to be entered in the form of an absolute value, in the units of g^2. This is because the scientists who created the Adaptive model found that a single value for BioWS variance best characterised the biological variation of Hb_{mass} in the cohort of athletes studied by Prommer et al. (2008), regardless of the magnitude of the athlete's Hb_{mass} (personal communication, P.E. Sottas, 2012). However, when Eastwood et al. (2011b) estimated the within-subject SDs of Hb_{mass} in male and female athletes, their results were in the form of a percentage. This suggests that a percentage, rather than an absolute value, better characterised the analytical and biological variations of Hb_{mass} in their cohort. This may also be true in a wider population of athletes.

It is not possible to ascertain from my data whether it would be most appropriate to include an estimate of biological variance as an absolute value or in the form of a percentage in the Adaptive model for Hb_{mass}. But the separate estimates of BioWS variance for male and female athletes used in the $Hb^{m\ (Eastwood)}$ model represent an intermediate solution and offer insight into the effects of the different approaches on the sensitivity and specificity of Hb^m models. The $Hb^{m\ (Eastwood)}$ model recorded a lower specificity than the $Hb^{m\ (Morkeberg)}$ model even though the BioWS variance used for males athletes in the $Hb^{m\ (Eastwood)}$ model was the largest of all the models. This was because one injured female athlete flagged the 99.9% limits of the

$Hb^{m\ (Eastwood)}$ (BioWS variance: 573 g^2) and $Hb^{m\ (Pottgiesser)}$ models (550 g^2), but did not flag the limits of the $Hb^{m\ (Morkeberg)}$ model (611 g^2). If the data are reanalysed using separate BioWS variance estimates for males and females in all Hb^m models (results not shown), the only notable difference from the existing results is the occurrence of one false-positive result at the 99.9% level in the $Hb^{m\ (Morkeberg)}$ model. This athlete was the same injured athlete who had registered as a false-positive in the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models in the original results. Therefore, if separate estimates of BioWS variance for males and females were implemented in the Hb^m model chosen for use in the ABP, the sensitivities and specificities of the $Hb^{m\ (Pottgiesser)}$, $Hb^{m\ (Morkeberg)}$ and $Hb^{m\ (Eastwood)}$ models would be identical at the 99.9% level. This finding lends additional support to the appropriateness of any of these three models for inclusion in the ABP. However, it would be necessary for anti-doping authorities to confirm whether an absolute or percentage value for BioWS variance best characterises the pattern of biological variation in Hb_{mass} before Hb^m models could be implemented in the ABP.

Test frequency

In theory, the sensitivity of the Adaptive model to doping would be reduced if the frequency of testing is too high, because the upper and lower limits of the model may simply *follow* the changes in the athlete rather than being sensitive to those changes (Ashenden et al. 2011). The results of the present study contradict this theory, with the sensitivities of all Hb^m models being higher in the 'All tests' analysis than in the 'Fewer tests' analysis. For the $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$ models, this had a marked effect on the sensitivity of the models at the 99.9% level. The sensitivities of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models at the 99% level were also higher when all tests results were included in the analysis. However, the sensitivities of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models at the 99.9% level were unaffected by testing frequency, further confirming the potential suitability of these models for the ABP.

Since the results of the present study suggest that, in general, sensitivity to doping was improved rather than reduced with more frequent testing, it may be advantageous for anti-doping authorities to test athletes as frequently as possible both in and out of competition. Although Hb_{mass} measures were made throughout the 12-13 week doping period in both cohorts of rHuEPO-doped athletes examined here, typically only one result over that entire period was flagged in each detected athlete. This illustrates the narrow window of opportunity that anti-doping authorities may have to detect doping, and therefore the importance of frequent testing to increase the chances of one result being detected that exceeds the limits. It has also been demonstrated previously that the window of opportunity for detection of blood withdrawal may be as little as 12 days (Prommer et al. 2007b). Given unlimited resources, anti-doping authorities would be able to implement more tests than are currently carried out, but this hypothetical approach is limited logistically and financially. To circumvent these limitations, intelligent testing protocols, where the 99% limits may be used to identify 'suspicious' athletes, can be used to direct resources towards particular athletes at the time of year when they are most likely to be engaging in doping (Pottgiesser et al. 2012; Zorzoli and Rossi 2012).

Although unsupported experimentally by our results, my observations from using the ABP software confirm that the Adaptive model does indeed behave in a way that *follows* the athletes' test results. For example, in an athlete who has seven prior Hb_{mass} results recorded in the ABP, a 9% increase in Hb_{mass} in the next test would exceed the 99% upper limit of the $Hb^{m\ (Morkeberg)}$ model, whereas a progressive 9% increase over the space of the next three tests would not exceed the limits. This is because over the course of three tests, the expected range would shift upwards (~1%) with the new upper limit being 10% above the first test result. This example demonstrates the upward shift in the Adaptive model limits as test results are entered successively, and the corresponding shift is even more substantial (~3%) in an athlete

without prior test results recorded in the ABP. These contrasting findings make it difficult to make a recommendation about the effect of testing frequency on Hb_{mass} models. However, in athletes for whom a number of prior test results have been recorded, it appears that the potential benefit of frequent testing to catch doped athletes outweighs the detrimental 'shift' effect of frequent testing.

Combination of Hb_{mass} and %Ret

Contrary to suggestions by previous researchers that the inclusion of %Ret in Hb_{mass} models would enhance the sensitivity of Hb_{mass} to doping, the combination of Hb_{mass} and %Ret as used in the ON^{hm+ret} model was not useful in this case. The ON^{hm+ret} model was devised using raw data from the publication of Parisotto et al. (2000a) where changes in Hb_{mass} and %Ret were measured in subjects who were given relatively large doses of rHuEPO. The results of my analysis suggest that the changes in Hb_{mass} and %Ret were too similar between the microdose rHuEPO recipients and non-doped athletes who had undertaken LHTL altitude training to be differentiated. Or in other words, microdosing with rHuEPO quite closely simulates the body's natural response to altitude (Ashenden et al. 2001). Finally, the negligible %Ret response in the microdose rHuEPO doping study meant that the sensitivity of the ON^{hm+ret} model to microdose rHuEPO doping was actually less than that of Hb_{mass} alone.

The use of models combining Hb_{mass} and %Ret has previously been demonstrated to be successful in the detection of autologous blood doping (Morkeberg et al. 2011; Pottgiesser et al. 2012). The major difference between the models used in those studies and the ON^{hm+ret} model is the way in which the two variables are combined; Morkeberg et al. (2011) and Pottgiesser et al. (2012) both used 'OFF' models that looked for a combination of high Hb_{mass} with low %Ret because this is the characteristic effect that autologous blood doping has on the body. The administration of rHuEPO results in high Hb_{mass} with concomitant high %Ret,

the combination that was included in the ON^{hm+ret} model. It is possible that an OFF model may have been more sensitive to rHuEPO doping in our analysis, looking for the characteristic combination of high Hb_{mass} and low %Ret in the weeks following cessation of rHuEPO doping. However, it was not possible to assess the sensitivity of an OFF model using these data because neither of the rHuEPO studies that I assessed collected Hb_{mass} and %Ret results after the cessation of the injections.

Limitations

The decision to use single values of between-subject variance, population mean and TE for all Hb^m models was to enable the influence of the BioWS variance of Hb_{mass} on the sensitivity and specificity of the models to be assessed in isolation. Each of the studies from which the four different BioWS variance values were derived, also had between-subject variances, population means and TEs available that could have been incorporated into the models. However, the influences of the between-subject variance and population mean are very small compared to the influence of the TE and BioWS variance (Sottas et al. 2010). For example, once a single baseline result is entered for an athlete, the influence of the between-subject variance is already reduced by two-thirds (Sottas et al. 2010). Although it may be important for future investigations to explore the influence of different estimates of between-subject variance and population mean on the sensitivity and specificity of Hb^m models, I would expect the effects to be minimal. As a sub-component of within-subject variance, the TE is equally important as the BioWS variance for influencing the width of the expected range generated by the Adaptive model. However, there is relatively close agreement about the magnitude of the TE for measuring Hb_{mass} (~1.4-2.0%). If Hb_{mass} is used in the ABP, the quality control measures that would need to be put in place (such as, identical equipment and strict adherence to uniform procedures) would likely ensure that the TEs recorded by different laboratories for this procedure would be very similar. Although the 1.7% value used for the

TE in this study may not represent the best-case scenario, it does represent a realistic estimate of current analytical error for CO re-breathing (Pottgiesser et al. 2012).

The abnormally high sequence results in athletes whose first test results fell outside the expected range reduced the utility of this aspect of the analysis for the Hb^{m} models. The population mean and between subject variance values that heavily influence the expected range for an athlete's first result were derived from a population of male and female endurance athletes (mostly cyclists and triathletes) (Prommer et al. 2008). Although these estimates produced an accurate range for most of the athletes involved in the present study, athletes who were at a sub-elite or recreational level, or whose body mass greatly exceeded that of the endurance population on which the estimates were based (such as, rugby league players) did not conform to the expected first Hb_{mass} result. While the ABP is not likely to be implemented in non-elite populations, the ABP is available for use by the international governing bodies of a variety of different sports. It may be necessary to calculate sport-specific estimates of population mean and between-subject variance of Hb_{mass} before the sequence analysis in the ABP can be reliably used in all sports.

Finally, it must be acknowledged that although I have classified all the athletes who took part in the studies not involving rHuEPO as non-doped athletes, it is not possible to be absolutely certain that these athletes were competing 'clean'. None of the athletes in those studies have been sanctioned for drug offences, and at least 60% of the athletes taking part would have been subjected to regular testing by national anti-doping agencies. Unfortunately, this is not a guarantee of ethical behaviour and I recognise this may be a potential limitation of my estimates of specificity for the Hb_{mass} models.

Conclusion

The $Hb^{m\ (Prommer-small)}$, $Hb^{m\ (Prommer-large)}$ and ON^{hm+ret} models are not suitable for use in the ABP because their level of specificity does not satisfy WADA's requirement for a maximum of 1 false-positive in 1000 tests. There are three different models using Hb_{mass} as their sole marker that could be implemented in the ABP with an acceptable specificity, each offering 17% sensitivity to rHuEPO doping at the WADA recommended 99.9% level of specificity. The $Hb^{m\ (Morkeberg)}$ model afforded the best specificity with an equivalent sensitivity to the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models. Alternatively, in a larger population of doped athletes, the $Hb^{m\ (Pottgiesser)}$ model may offer the best sensitivity. Finally, the estimate of BioWS variance included in the $Hb^{m\ (Eastwood)}$ model may be closest to the true BioWS variance of Hb_{mass} in athletes. Therefore, either of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ may be suitable for use in the ABP, but research in larger doped and non-doped athlete populations may be required to ascertain which of these three models possesses the best sensitivity whilst maintaining the necessary level of specificity. One major limitation of the $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models was the false-positive result they generated for one injured athlete whose Hb_{mass} was likely affected by the natural influences of surgery and changes in training load. However, this false-positive would be avoided if WADA's Anti-Doping Administration & Management System could expediently identify events such as surgery, and medical records could be accessed to substantiate such claims.

Acknowledgements

My sincere thanks go to Pierre-Edouard Sottas for granting me access to the ABP software and for his guidance and support in its use.

Summary

In recent years, a number of researchers have suggested that the inclusion of haemoglobin mass (Hb_{mass}) measures in an Adaptive model of the Athlete Biological Passport (ABP) would improve the sensitivity of the ABP (Giraud et al. 2010; Morkeberg et al. 2011; Pottgiesser et al. 2007; Prommer et al. 2008). However, doubts have also been expressed about the suitability of Hb_{mass} measurement for this purpose. Specifically, the lack of a quality control system for Hb_{mass} measurement has been highlighted (Schumacher and Pottgiesser 2010) and doubts have been expressed about whether the perturbations in Hb_{mass} associated with doping would be distinguishable from the normal within-subject variation in Hb_{mass} (Eastwood et al. 2011b; Lundby and Robach 2010). Therefore, the aims of this thesis were:

(i) to develop a quality control system for Hb_{mass} measurement;

(ii) to quantify the potential confounding effects of prolonged exercise, illness, injury, different forms of hypoxia and microdoses of rHuEPO on Hb_{mass}; and,

(iii) to examine the sensitivities and specificities of different Adaptive models based on Hb_{mass} in the ABP.

The investigation described in Chapter 3 demonstrated that custom-made quality control solutions could be used to ensure the equivalency of Hb_{mass} measures made in different laboratories. This 'proof-of-concept' removes one potential barrier to the inclusion of Hb_{mass} measurements in the ABP.

In Chapters 4, 5, 6 and 7, it was confirmed that several factors confounded Hb_{mass} in non-doped athletes. Recent participation in ultra-endurance triathlon, Classic altitude training and

LHTL altitude training led to group mean increases in Hb_{mass} of 3.2%, 3.8% and ~4%, respectively. In contrast, the effects of surgery and reduced training in injured and ill athletes were to decrease Hb_{mass}, by 2.3% and 2.7%, respectively. It is necessary, however, to bear in mind that these estimates were mean effects and that the fluctuations of Hb_{mass} in individual athletes varied substantially. Notable individual deviations in Hb_{mass} included an 8.3% increase following ultra-endurance triathlon, an 8.4% increase after Classic altitude training, an 8.7% increase after LHTL altitude training and a 15.6% decrease in an injured athlete. Acute Intermittent Hypoxic Exposure (AcIHE) did not change Hb_{mass}. The group mean -0.3% change in Hb_{mass} following AcIHE suggested that this form of hypoxia does not constitute an adequate hypoxic stimulus to induce haematological adaptation in athletes. In addition, Chapter 8 demonstrated that a 12-week regimen of rHuEPO microdosing led to a group mean Hb_{mass} increase of 11.0%. The group mean Hb_{mass} change in this doped population was substantially larger than the mean effects of any of the aforementioned confounders in non-doped athletes (11% versus ~4%). However, there was large variation in the individual Hb_{mass} responses to microdose rHuEPO doping, ranging from 3% to 20% over the course of the 12 weeks.

The sensitivities and specificities of six different Adaptive models based on Hb_{mass} were examined in Chapters 8 and 9 of this thesis. The sensitivity and specificity of Adaptive models that used Hb_{mass} as a single marker (Hb^m models) were heavily influenced by the estimate of biological within-subject (BioWS) variance included in model. The $Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ –\ large)}$ models would not be suitable for inclusion in the ABP because each of these two models recorded a substantial number of false-positive results in a sample of 159 non-doped athletes. In contrast, any one of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models may be suitable for use in the ABP due to their high specificity. Each of these three models recorded 17% sensitivity to rHuEPO doping. One key limitation of the Hb^m

$^{(Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models was the false-positive result they recorded for one injured athlete following surgery and a period of detraining. If anti-doping authorities wished to implement Hb_{mass} measures as part of the ABP, they would need to consider ways in which athletes who have experienced injuries or illnesses could submit medical records as evidence to justify any unusual changes in Hb_{mass}. WADA's on-line Anti-Doping Administration & Management System could readily be used to identify events such as surgery. However, any means for these athletes to be given exemptions would have to be policed judiciously to prevent athletes abusing this system.

Finally, in theory, a combination of smaller changes in multiple blood markers may allow higher sensitivity to doping than single markers alone. However, the $ON^{hm + Ret}$ model would not be a useful inclusion in the ABP due to its failure to differentiate accurately between hypoxia-induced and doping-induced increases in Hb_{mass}.

Practical Recommendations

The practical recommendations arising from this thesis include:

1. Quality control solutions spanning the range of 0-10% HbCO should be used to ensure equivalency of Hb_{mass} measurements made in different laboratories.
2. Hb_{mass} measurements should not be made in athletes within 3 hours of finishing ultra-endurance triathlon and possibly other endurance sports.
3. Sports performances within 1 week of returning to sea-level after LHTL or Classic altitude training will likely be compromised. Athletes and coaches should plan their competitive season accordingly. It remains unclear whether these impairments result from training-induced fatigue or altitude, per se.

4. Any performance benefits that arise from AcIHE are unlikely to be related to changes in Hb_{mass}. Therefore scientists should focus their resources on examining non-haematological adaptations to this form of hypoxia.
5. No further consideration should be given to including either the $Hb^{m\ (Prommer\ -\ small)}$, $Hb^{m\ (Prommer\ -\ large)}$ or ON^{hm+ret} models in the ABP.

Future directions

The stability of %HbCO during storage and transport of quality control samples for Hb_{mass} measurement requires further investigation. Furthermore, the quality controls used in my investigations were custom-made and it will be important for future research to confirm that a similar level of precision can be achieved using a commercial manufacturing procedure.

If Hb_{mass} measurement is to be incorporated into the ABP, it will be essential for the time-course of Hb_{mass} changes following exercise in elite athletes to be clarified. Future research should also investigate whether the observed effect of ultra-endurance triathlon on Hb_{mass} is sport-specific or whether restrictions should be applied to the measurement of Hb_{mass} after other sports.

There were indications in the results of Chapter 9 of this thesis that the decision about whether to include separate estimates of BioWS variance for males and females may influence the sensitivity and specificity of Adaptive models based on Hb_{mass}. These results did not offer insight into the true pattern of BioWS variance in a population of athletes. Future research should investigate whether the biological variability of Hb_{mass} is best characterised by a percentage or an absolute value, and in each case whether separate estimates should be used for males and females.

It was demonstrated in Chapter 9 that the specificities of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models were sufficient to allow reliable inclusion of each in the ABP. Assessments of specificity were made using a sample of 159 longitudinal Hb_{mass} profiles of non-doped athletes, however only 18 longitudinal profiles of doped athletes were available for the assessment of sensitivity. The decision about whether to utilise any of these models in the ABP is likely to be dependent on any additional sensitivity they offer over that of the current markers of the ABP. Consequently, future research should assess the sensitivity of the $Hb^{m\ (Morkeberg)}$, $Hb^{m\ (Pottgiesser)}$ and $Hb^{m\ (Eastwood)}$ models in a larger sample of doped athletes.

Despite the failure of the ON^{hm+ret} model, other researchers' descriptions of successful OFF-models indicate that combining Hb_{mass} with other blood parameters may enhance the contribution of Hb_{mass} measurements to the ABP (Morkeberg et al. 2011; Pottgiesser et al. 2012). Future research should investigate whether Hb_{mass} can be combined with any other blood markers to create a model that is sensitive and specific to doping.

Leakage of CO from the re-breathing circuit significantly inflates the Hb_{mass} result (Ryan et al.) and would invalidate measurements made for anti-doping purposes. As Pottgiesser et al. (2012) recently pointed out, the CO re-breathing procedure requires subject cooperation and unwilling athletes could easily feign accidental leakage of CO from around their mouth. The toxicity of CO prevents repetition of the procedure indefinitely, so unscrupulous athletes could avoid participation in testing in this way. Therefore, the use of Hb_{mass} for anti-doping purposes is dependent on a leak-proof system being created. Future research should investigate the feasibility of implementing such a system, otherwise the utility of Hb_{mass} in an anti-doping setting is seriously compromised.

References

Ahlgrim C, Pottgiesser T, Kron J, Duerr H, Baumstark M, Schumacher YO (2009) Relations between haemoglobin mass, cardiac dimensions and aerobic capacity in endurance trained cyclists. *J Sports Med Phys Fitness* 49: 364-371

Alexander A, Garvican LA, Burge CM, Clark SA, Plowman JS, Gore CJ (2011) Standardising analysis of carbon monoxide rebreathing for application in anti-doping. *J Sci Med Sport* 14: 100-105

Ashenden M, Gough CE, Garnham A, Gore CJ, Sharpe K (2011) Current markers of the Athlete Blood Passport do not flag microdose EPO doping. *Eur J Appl Physiol* 111: 2307-2314

Ashenden M, Varlet-Marie E, Lasne F, Audran M (2006) The effects of microdose recombinant human erythropoietin regimens in athletes. *Haematologica* 91: 1143-1144

Ashenden MJ (2002) A strategy to deter blood doping in sport. *Haematologica* 87: 225-232

Ashenden MJ, Fricker PA, Ryan RK, Morrison NK, Dobson GP, Hahn AG (1998) The haematological response to an iron injection amongst female athletes. *Int J Sports Med* 19: 474-478

Ashenden MJ, Gore CJ, Martin DT, Dobson GP, Hahn AG (1999) Effects of a 12-day "live high, train low" camp on reticulocyte production and haemoglobin mass in elite female road cyclists. *Eur J Appl Physiol Occup Physiol* 80: 472-478

Ashenden MJ, Hahn AG, Martin DT, Logan P, Parisotto R, Gore CJ (2001) A comparison of the physiological response to simulated altitude exposure and r-HuEpo administration. *J Sports Sci* 19: 831-837

Ashenden MJ, Sharpe K, Schoch C, Schumacher YO (2004) Effect of pre-competition and altitude training on blood models used to detect erythropoietin abuse by athletes. *Haematologica* 89: 1019-1020

Banister EW, Calvert TW, Savage MV, Bach TM (1975) A systems model of training for athletic performance. *Aust J Sports Med* 7: 57-61

Berglund B (1992) High-altitude training. Aspects of haematological adaptation. *Sports Med* 14: 289-303

Billat LV, Koralsztein JP (1996) Significance of the velocity at VO_{2max} and time to exhaustion at this velocity. *Sports Med* 22: 90-108

Bonetti DL, Hopkins WG (2009) Sea-level exercise performance following adaptation to hypoxia: a meta-analysis. *Sports Med* 39: 107-127

Bonetti DL, Hopkins WG, Kilding AE (2006) High-intensity kayak performance after adaptation to intermittent hypoxia. *Int J Sports Physiol Perform* 1: 246-260

Bonetti DL, Hopkins WG, Lowe TE, Boussana A, Kilding AE (2009) Cycling performance following adaptation to two protocols of acutely intermittent hypoxia. *Int J Sports Physiol Perform* 4: 68-83

Borno A, Aachmann-Andersen NJ, Munch-Andersen T, Hulston CJ, Lundby C (2010) Screening for recombinant human erythropoietin using [Hb], reticulocytes, the OFF(hr score), OFF(z score) and Hb(z score): status of the Blood Passport. *Eur J Appl Physiol* 109: 537-543

Brien AJ, Harris RJ, Simon TL (1989) The effects of an autologous infusion of 400 mL red blood cells on selected haematological parameters and 1500 m race time in highly trained runners. *Bahrain Med Bull* 11: 6-16

Brien AJ, Simon TL (1987) The effects of red blood cell infusion on 10-km race time. *J Am Med Assoc* 257: 2761-2765

Bruce EN, Bruce MC (2003) A multicompartment model of carboxyhemoglobin and carboxymyoglobin responses to inhalation of carbon monoxide. *J Appl Physiol* 95: 1235-1247

Brugniaux JV, Schmitt L, Robach P, Nicolet G, Fouillot JP, Moutereau S, Lasne F, Pialoux V, Saas P, Chorvot MC, Cornolo J, Olsen NV, Richalet JP (2006) Eighteen days of "living high, training low" stimulate erythropoiesis and enhance aerobic performance in elite middle-distance runners. *J Appl Physiol* 100: 203-211

Burge CM, Skinner SL (1995) Determination of hemoglobin mass and blood volume with CO: evaluation and application of a method. *J Appl Physiol* 79: 623-631

Burtscher M, Gatterer H, Faulhaber M, Gerstgrasser W, Schenk K (2010) Effects of intermittent hypoxia on running economy. *Int J Sports Med* 31: 644-650

Butcher JD (1993) Runner's diarrhea and other intestinal problems of athletes. *Am Fam Physician* 48: 623-627

Catlin DH, Fitch KD, Ljungqvist A (2008) Medicine and science in the fight against doping in sport. *J Intern Med* 264: 99-114

Cazzola M (2000) A global strategy for prevention and detection of blood doping with erythropoietin and related drugs. *Haematologica* 85: 561-563

Clark SA, Quod MJ, Clark MA, Martin DT, Saunders PU, Gore CJ (2009) Time course of haemoglobin mass during 21 days live high:train low simulated altitude. *Eur J Appl Physiol* 106: 399-406

Cohen J (1988) *Statistical power analyses for the behavioural sciences*. Lawrence Erlbaum Associates, Hillsdale

Connes P, Caillaud C, Simar D, Villard S, Sicart MT, Audran M (2004) Strengths and weaknesses of established indirect models to detect recombinant human erythropoietin abuse on blood samples collected 48-hr post administration. *Haematologica* 89: 891-892

Costill DL, Thomas R, Robergs RA, Pascoe D, Lambert C, Barr S, Fink WJ (1991) Adaptations to swimming training: influence of training volume. *Med Sci Sports Exerc* 23: 371-377

di Prampero PE (2003) Factors limiting maximal performance in humans. *Eur J Appl Physiol* 90: 420-429

di Prampero PE, Atchou G, Bruckner JC, Moia C (1986) The energetics of endurance running. *Eur J Appl Physiol Occup Physiol* 55: 259-266

Dick FW (1992) Training at altitude in practice. *Int J Sports Med* 13 Suppl 1: S203-206

Eastwood A, Bourdon PC, Norton KI, Lewis NR, Snowden KR, Gore CJ (2011a) No change in hemoglobin mass after 40 days of physical activity in previously untrained adults. *Scand J Med Sci Sports*: (Epub ahead of print)

Eastwood A, Bourdon PC, Snowden KR, Gore CJ (2012) Detraining decreases Hbmass of triathletes. *Int J Sports Med*: (Epub ahead of print)

Eastwood A, Bourdon PC, Withers RT, Gore CJ (2009) Longitudinal changes in haemoglobin mass and VO2max in adolescents. *Eur J Appl Physiol* 105: 715-721

Eastwood A, Hopkins WG, Bourdon PC, Withers RT, Gore CJ (2008) Stability of hemoglobin mass over 100 days in active men. *J Appl Physiol* 104: 982-985

Eastwood A, Sharpe K, Bourdon PC, Woolford SM, Saunders PU, Robertson EY, Clark SA, Gore CJ (2011b) Within subject variation in hemoglobin mass in elite athletes. *Med Sci Sports Exerc* 44: 725-732

Eichner ER (2007) Blood doping : infusions, erythropoietin and artificial blood. *Sports Med* 37: 389-391

Ekblom B, Berglund B (1991) Effect of erythropoietin administration maximal aerobic power. *Scand J Med Sci Sports* 1: 88-93

Fallon KE, Bishop G (2002) Changes in erythropoiesis assessed by reticulocyte parameters during ultralong distance running. *Clin J Sport Med* 12: 172-178

Faria EW, Parker DL, Faria IE (2005) The science of cycling: physiology and training - part 1. *Sports Med* 35: 285-312

Faude O, Meyer T, Scharhag J, Weins F, Urhausen A, Kindermann W (2008) Volume vs. intensity in the training of competitive swimmers. *Int J Sports Med* 29: 906-912

Faude O, Meyer T, Urhausen A, Kindermann W (2009) Recovery training in cyclists: ergometric, hormonal and psychometric findings. *Scand J Med Sci Sports* 19: 433-441

Favret F, Richalet JP, Henderson KK, Germack R, Gonzalez NC (2001) Myocardial adrenergic and cholinergic receptor function in hypoxia: correlation with O(2) transport in exercise. *Am J Physiol Regul Integr Comp Physiol* 280: R730-738

FINA (2009) FINA Points System http://www.fina.org/pool/index.php?option=com_content&view=article&id=7236&Itemid=297. Accessed 20th April 2011

Forster HV, Dempsey JA, Birnbaum ML, Reddan WG, Thoden J, Grover RF, Rankin J (1971) Effect of chronic exposure to hypoxia on ventilatory response to CO 2 and hypoxia. *J Appl Physiol* 31: 586-592

Foster C, Florhaug JA, Franklin J, Gottschall L, Hrovatin LA, Parker S, Doleshal P, Dodge C (2001) A new approach to monitoring exercise training. *J Strength Cond Res* 15: 109-115

Fraser CG, Harris EK (1989) Generation and application of data on biological variation in clinical chemistry. *Crit Rev Clin Lab Sci* 27: 409-437

Friedmann B, Frese F, Menold E, Kauper F, Jost J, Bartsch P (2005) Individual variation in the erythropoietic response to altitude training in elite junior swimmers. *Br J Sports Med* 39: 148-153

Friedmann B, Weller E, Mairbaurl H, Bartsch P (2001) Effects of iron repletion on blood volume and performance capacity in young athletes. *Med Sci Sports Exerc* 33: 741-746

Fuchs U, Reiss M (1990) Ho¨hentraining: das Erfolgskonzept der Ausdauersportarten. *Trainerbibliothek* 27: 128

Gabbett TJ, Ullah S (2012) Relationship between running loads and soft-tissue injury in elite team sport athletes. *J Strength Cond Res* 26: 953-960

Garvican LA, Burge CM, Cox AJ, Clark SA, Martin DT, Gore CJ (2010a) Carbon monoxide uptake kinetics of arterial, venous and capillary blood during CO rebreathing. *Exp Physiol* 95: 1156-1166

Garvican LA, Eastwood A, Martin DT, Ross ML, Gripper A, Gore CJ (2010b) Stability of hemoglobin mass during a 6-day UCI ProTour cycling race. *Clin J Sport Med* 20: 200-204

Garvican LA, Lobigs L, Telford RD, Fallon KE, Gore CJ (2011a) Haemoglobin mass in an anaemic female endurance runner before and after iron supplementation. *Int J Sports Physiol Perform* 6: 137-140

Garvican LA, Martin DT, Clark MA, Quod M, Stephens B, Prommer N, F. SW, Impellizeri FM, Rampinini E, Sassi A, Gore CJ (2012) Time course of the hemoglobin mass response to natural altitude training in elite endurance cyclists. *Scand J Med Sci Sports* 22: 95-103

Garvican LA, Martin DT, McDonald W, Gore CJ (2010c) Seasonal variation of haemoglobin mass in internationally competitive female road cyclists. *Eur J Appl Physiol* 109: 221-231

Garvican LA, Pottgiesser T, Martin DT, Schumacher YO, Barras M, Gore CJ (2011b) The contribution of haemoglobin mass to increases in cycling performance induced by simulated LHTL. *Eur J Appl Physiol* 111: 1089-1101

Gaudard A, Varlet-Marie E, Bressolle F, Audran M (2003) Drugs for increasing oxygen and their potential use in doping: a review. *Sports Med* 33: 187-212

Giraud S, Sottas PE, Robinson N, Saugy M (2010) Blood transfusion in sports. *Handb Exp Pharmacol*: 295-304

Gore CJ, Bourdon PC, Woolford SM, Ostler LM, Eastwood A, Scroop GC (2006a) Time and sample site dependency of the optimized CO-rebreathing method. *Med Sci Sports Exerc* 38: 1187-1193

Gore CJ, Clark SA, Saunders PU (2007) Nonhematological mechanisms of improved sea-level performance after hypoxic exposure. *Med Sci Sports Exerc* 39: 1600-1609

Gore CJ, Hahn AG (2005) Letter to the editors - Re: Heinicke K, Heinicke I, Schmidt W, Wolfarth B. A three-week traditional altitude training increases hemoglobin mass and red cell volume in elite biathlon athletes. Int J Sports Med 2005; 26: 350-355. *Int J Sports Med* 26: 504-505; author reply 506-507

Gore CJ, Hahn AG, Aughey RJ, Martin DT, Ashenden MJ, Clark SA, Garnham AP, Roberts AD, Slater GJ, McKenna MJ (2001) Live high:train low increases muscle buffer capacity and submaximal cycling efficiency. *Acta Physiol Scand* 173: 275-286

Gore CJ, Hopkins WG (2005) Counterpoint: positive effects of intermittent hypoxia (live high:train low) on exercise performance are not mediated primarily by augmented red cell volume. *J Appl Physiol* 99: 2055-2057

Gore CJ, Hopkins WG, Burge CM (2005) Errors of measurement for blood volume parameters: a meta-analysis. *J Appl Physiol* 99: 1745-1758

Gore CJ, Parisotto R, Ashenden MJ, Stray-Gundersen J, Sharpe K, Hopkins W, Emslie KR, Howe C, Trout GJ, Kazlauskas R, Hahn AG (2003) Second-generation blood tests to detect erythropoietin abuse by athletes. *Haematologica* 88: 333-344

Gore CJ, Rodriguez FA, Truijens MJ, Townsend NE, Stray-Gundersen J, Levine BD (2006b) Increased serum erythropoietin but not red cell production after 4 wk of intermittent hypobaric hypoxia (4,000-5,500 m). *J Appl Physiol* 101: 1386-1393

Gough C, Sharpe K, Ashenden M, Anson JM, Saunders PU, Garvican LA, Bonetti DL, Gore CJ, Prommer N (2011) Quality control technique to reduce the variability of longitudinal measurement of hemoglobin mass. *Scand J Med Sci Sports* 21: e365-371

Gough CE, Eastwood A, Saunders PU, Anson JM, Gore CJ (In Press) Spurious Hb mass increases following exercise. *Int J Sports Med*

Gough CE, Saunders PU, Fowlie J, Savage B, Pyne DB, Anson JM, Wachsmuth N, Prommer N, Gore CJ (2012) Influence of altitude training modality on performance and total haemoglobin mass in elite swimmers. *Eur J Appl Physiol*: (Epub ahead of print)

Green HJ, Sutton JR, Coates G, Ali M, Jones S (1991) Response of red cell and plasma volume to prolonged training in humans. *J Appl Physiol* 70: 1810-1815

Grehant M, Quinquard, E. (1882) Mesures du volume du sang contenu dans l'organisme d'un mammifere vivant. *C R Acad Sci Paris* 94: 1450

Hackett PH, Roach R (1995) *High-altitude medicine*. Mosby, St Louis, MO

Hamlin MJ, Marshall HC, Hellemans J, Ainslie PN (2010) Effect of intermittent hypoxia on muscle and cerebral oxygenation during a 20-km time trial in elite athletes: a preliminary report. *Appl Physiol Nutr Metab* 35: 548-559

Hampson NB (2008) Stability of carboxyhemoglobin in stored and mailed blood samples. *Am J Emerg Med* 26: 191-195

Harrison MH (1985) Effects on thermal stress and exercise on blood volume in humans. *Physiol Rev* 65: 149-209

Harriss DJ, Atkinson G (2009) International Journal of Sports Medicine - ethical standards in sport and exercise science research. *Int J Sports Med* 30: 701-702

Heinicke K, Heinicke I, Schmidt W, Wolfarth B (2005) A three-week traditional altitude training increases hemoglobin mass and red cell volume in elite biathlon athletes. *Int J Sports Med* 26: 350-355

Heinicke K, Wolfarth B, Winchenbach P, Biermann B, Schmid A, Huber G, Friedmann B, Schmidt W (2001) Blood volume and hemoglobin mass in elite athletes of different disciplines. *Int J Sports Med* 22: 504-512

Hew-Butler T, Collins M, Bosch A, Sharwood K, Wilson G, Armstrong M, Jennings C, Swart J, Noakes T (2007) Maintenance of plasma volume and serum sodium concentration despite body weight loss in ironman triathletes. *Clin J Sport Med* 17: 116-122

Hochachka PW, Beatty CL, Burelle Y, Trump ME, McKenzie DC, Matheson GO (2002) The lactate paradox in human high-altitude physiological performance. *News Physiol Sci* 17: 122-126

Hopfl G, Ogunshola O, Gassmann M (2003) Hypoxia and high altitude. The molecular response. *Adv Exp Med Biol* 543: 89-115

Hopkins W (2005) Competitive performance of elite track and field athletes: Variability and smallest worthwhile enhancements. *Sportscience* 9: 17-20

Hopkins WG (2000) Measures of reliability in sports medicine and science. *Sports Med* 30: 1-15

Hopkins WG (2004) How to interpret changes in an athletic performance test. *Sportscience* 8: 1-7

Hopkins WG (2006) Spreadsheets for Analysis of Controlled Trials, with Adjustment for a Subject Characteristic. *Sportscience* 10: 46-50

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

Hoppeler H, Weibel ER (2000) Structural and functional limits for oxygen supply to muscle. *Acta Physiol Scand* 168: 445-456

Hsia CC (1998) Respiratory function of hemoglobin. *N Engl J Med* 338: 239-247

Hutler M, Beneke R, Littschwager A, Boning D (2001) Measured fraction of carboxyhaemoglobin depends on oxygen saturation of haemoglobin. *Scand J Clin Lab Invest* 61: 83-87

Jacobs RA, Rasmussen P, Siebenmann C, Diaz V, Gassmann M, Pesta D, Gnaiger E, Nordsborg NB, Robach P, Lundby C (2011) Determinants of time trial performance and maximal incremental exercise in highly trained endurance athletes. *J Appl Physiol* 111: 1422-1430

Julian CG, Gore CJ, Wilber RL, Daniels JT, Fredericson M, Stray-Gundersen J, Hahn AG, Parisotto R, Levine BD (2004) Intermittent normobaric hypoxia does not alter performance or erythropoietic markers in highly trained distance runners. *J Appl Physiol* 96: 1800-1807

Jung K, Shon R (1994) Altitude Training - A summary of the main papers presented at the European Athletics Coaches Association's workshop on altitude training at Belmeken, Bulgaria in May 1994. *An edited translation from Die Lehre der Leichtathletik* 33

Katayama K, Fujita H, Sato K, Ishida K, Iwasaki K, Miyamura M (2005) Effect of a repeated series of intermittent hypoxic exposures on ventilatory response in humans. *High Alt Med Biol* 6: 50-59

Katayama K, Matsuo H, Ishida K, Mori S, Miyamura M (2003) Intermittent hypoxia improves endurance performance and submaximal exercise efficiency. *High Alt Med Biol* 4: 291-304

Kjellberg SR, Rudhe U, Sjostrand T (1949) Increase of the amount of hemoglobin and blood volume in connection with physical training. *Acta Physiol Scand* 19: 148-151

Knaupp W, Khilnani S, Sherwood J, Scharf S, Steinberg H (1992) Erythropoietin response to acute normobaric hypoxia in humans. *J Appl Physiol* 73: 837-840

Koistinen PO, Rusko H, Irjala K, Rajamaki A, Penttinen K, Sarparanta VP, Karpakka J, Leppaluoto J (2000) EPO, red cells, and serum transferrin receptor in continuous and intermittent hypoxia. *Med Sci Sports Exerc* 32: 800-804

Lasne F (2001) Double-blotting: a solution to the problem of non-specific binding of secondary antibodies in immunoblotting procedures. *J Immunol Methods* 253: 125-131

Lasne F, Martin L, Martin JA, de Ceaurriz J (2009) Detection of continuous erythropoietin receptor activator in blood and urine in anti-doping control. *Haematologica* 94: 888-890

Levine BD (2002) Intermittent hypoxic training: fact and fancy. *High Alt Med Biol* 3: 177-193

Levine BD, Stray-Gundersen J (1997) "Living high-training low": effect of moderate-altitude acclimatization with low-altitude training on performance. *J Appl Physiol* 83: 102-112

Levine BD, Stray-Gundersen J (2005) Point: positive effects of intermittent hypoxia (live high:train low) on exercise performance are mediated primarily by augmented red cell volume. *J Appl Physiol* 99: 2053-2055

Lippi G, Banfi G (2006) Blood transfusions in athletes. Old dogmas, new tricks. *Clin Chem Lab Med* 44: 1395-1402

Lippi G, Franchini M, Salvagno GL, Guidi GC (2006) Biochemistry, physiology, and complications of blood doping: facts and speculation. *Crit Rev Clin Lab Sci* 43: 349-391

Lundby C, Robach P (2010) Assessment of total haemoglobin mass: can it detect erythropoietin-induced blood manipulations? *Eur J Appl Physiol* 108: 197-200

Lundby C, Robach P, Saltin B (2012) The evolving science of detection of 'blood doping'. *Br J Pharmacol* 165: 1306-1315

Lundby C, Thomsen JJ, Boushel R, Koskolou M, Warberg J, Calbet JA, Robach P (2007) Erythropoietin treatment elevates haemoglobin concentration by increasing red cell volume and depressing plasma volume. *J Physiol* 578: 309-314

Malcovati L, Pascutto C, Cazzola M (2003) Hematologic passport for athletes competing in endurance sports: a feasibility study. *Haematologica* 88: 570-581

Margaritis I, Tessier F, Richard MJ, Marconnet P (1997) No evidence of oxidative stress after a triathlon race in highly trained competitors. *Int J Sports Med* 18: 186-190

Martino M, Gledhill N, Jamnik V (2002) High VO2max with no history of training is primarily due to high blood volume. *Med Sci Sports Exerc* 34: 966-971

McCord JL, Halliwill JR (2006) H1 and H2 receptors mediate postexercise hyperemia in sedentary and endurance exercise-trained men and women. *J Appl Physiol* 101: 1693-1701

McDonald W (ed) (1984) *Butterworth's Medical Dictionary*. Botterworth's, London

McLaughlin JE, Howley ET, Bassett DR, Jr., Thompson DL, Fitzhugh EC (2010) Test of the classic model for predicting endurance running performance. *Med Sci Sports Exerc* 42: 991-997

Meeusen R, Duclos M, Gleeson M, Rietjens G, Steinacker JM, Urhausen A (2006) Prevention, diagnosis and treatment of the Overtraining Syndrome: ECSS Position Statement 'Task Force'. *Eur J Sport Sci* 6: 1-14

Miller BJ, Pate RR, Burgess W (1988) Foot impact force and intravascular hemolysis during distance running. *Int J Sports Med* 9: 56-60

Millet GP, Roels B, Schmitt L, Woorons X, Richalet JP (2010) Combining hypoxic methods for peak performance. *Sports Med* 40: 1-25

Morkeberg J, Belhage B, Damsgaard R (2008) Strategies to combat doping in cycling. *Int SportMed J* 9: 155-161

Morkeberg J, Sharpe K, Belhage B, Damsgaard R, Schmidt W, Prommer N, Gore CJ, Ashenden MJ (2011) Detecting autologous blood transfusions: a comparison of three passport approaches and four blood markers. *Scand J Med Sci Sports* 21: 235-243

Mujika I, Busso T, Lacoste L, Barale F, Geyssant A, Chatard JC (1996) Modeled responses to training and taper in competitive swimmers. *Med Sci Sports Exerc* 28: 251-258

Mujika I, Padilla S, Pyne D (2002) Swimming performance changes during the final 3 weeks of training leading to the Sydney 2000 Olympic Games. *Int J Sports Med* 23: 582-587

Myhre LG, Brown DK, Hall FG, Dill DB (1968) The use of carbon monoxide and T-1824 for determining blood volume. *Clin Chem* 14: 1197-1205

Neubauer O, Konig D, Wagner KH (2008) Recovery after an Ironman triathlon: sustained inflammatory responses and muscular stress. *Eur J Appl Physiol* 104: 417-426

Nissen-Lie G, Birkeland K, Hemmersbach P, Skibeli V (2004) Serum sTfR levels may indicate charge profiling of urinary r-hEPO in doping control. *Med Sci Sports Exerc* 36: 588-593

Nummela A, Rusko H (2000) Acclimatization to altitude and normoxic training improve 400-m running performance at sea level. *J Sports Sci* 18: 411-419

O'Toole ML, Hiller WD, Roalstad MS, Douglas PS (1988) Hemolysis during triathlon races: its relation to race distance. *Med Sci Sports Exerc* 20: 272-275

Parisotto R, Gore CJ, Emslie KR, Ashenden MJ, Brugnara C, Howe C, Martin DT, Trout GJ, Hahn AG (2000a) A novel method utilising markers of altered erythropoiesis for the detection of recombinant human erythropoietin abuse in athletes. *Haematologica* 85: 564-572

Parisotto R, Gore CJ, Hahn AG, Ashenden MJ, Olds TS, Martin DT, Pyne DB, Gawthorn K, Brugnara C (2000b) Reticulocyte parameters as potential discriminators of recombinant human erythropoietin abuse in elite athletes. *Int J Sports Med* 21: 471-479

Parisotto R, Wu M, Ashenden MJ, Emslie KR, Gore CJ, Howe C, Kazlauskas R, Sharpe K, Trout GJ, Xie M (2001) Detection of recombinant human erythropoietin abuse in athletes utilizing markers of altered erythropoiesis. *Haematologica* 86: 128-137

Pinheiro J, Bates D, DebRoy S, Sarkar D, the R core team.(2009). nlme: Linear and Nonlinear Mixed Effects Models Series.

Pottgiesser T, Ahlgrim C, Ruthardt S, Dickhuth HH, Schumacher YO (2009a) Hemoglobin mass after 21 days of conventional altitude training at 1816 m. *J Sci Med Sport* 12: 673-675

Pottgiesser T, Echteler T, Sottas PE, Umhau M, Schumacher YO (2012) Hemoglobin mass and biological passport for the detection of autologous blood doping. *Med Sci Sports Exerc*: (Epub ahead of print)

Pottgiesser T, Sottas PE, Echteler T, Robinson N, Umhau M, Schumacher YO (2011) Detection of autologous blood doping with adaptively evaluated biomarkers of doping: a longitudinal blinded study. *Transfusion* 51: 1707-1715

Pottgiesser T, Specker W, Umhau M, Dickhuth HH, Roecker K, Schumacher YO (2008) Recovery of hemoglobin mass after blood donation. *Transfusion* 48: 1390-1397

Pottgiesser T, Specker W, Umhau M, Roecker K, Schumacher YO (2009b) Post-transfusion stability of haemoglobin mass. *Vox Sang* 96: 119-127

Pottgiesser T, Umhau M, Ahlgrim C, Ruthardt S, Roecker K, Schumacher YO (2007) Hb mass measurement suitable to screen for illicit autologous blood transfusions. *Med Sci Sports Exerc* 39: 1748-1756

Prommer N, Ehrmann U, Schmidt W, Steinacker JM, Radermacher P, Muth CM (2007a) Total haemoglobin mass and spleen contraction: a study on competitive apnea divers, non-diving athletes and untrained control subjects. *Eur J Appl Physiol* 101: 753-759

Prommer N, Heckel A, Schmidt W (2007b) Timeframe to detect blood withdrawal associated with autologous blood doping. *Med Sci Sports Exerc* 39: S3

Prommer N, Schmidt W (2007) Loss of CO from the intravascular bed and its impact on the optimised CO-rebreathing method. *Eur J Appl Physiol* 100: 383-391

Prommer N, Sottas PE, Schoch C, Schumacher YO, Schmidt W (2008) Total hemoglobin mass--a new parameter to detect blood doping? *Med Sci Sports Exerc* 40: 2112-2118

Pyne D, Trewin C, Hopkins W (2004) Progression and variability of competitive performance of Olympic swimmers. *J Sports Sci* 22: 613-620

R Development Core Team (2010) *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria

Reichel C (2011) Recent developments in doping testing for erythropoietin. *Anal Bioanal Chem* 401: 463-481

Robach P, Schmitt L, Brugniaux JV, Roels B, Millet G, Hellard P, Nicolet G, Duvallet A, Fouillot JP, Moutereau S, Lasne F, Pialoux V, Olsen NV, Richalet JP (2006) Living high-training low: effect on erythropoiesis and aerobic performance in highly-trained swimmers. *Eur J Appl Physiol* 96: 423-433

Robertson EY, Aughey RJ, Anson JM, Hopkins WG, Pyne DB (2010a) Effects of simulated and real altitude exposure in elite swimmers. *J Strength Cond Res* 24: 487-493

Robertson EY, Saunders PU, Pyne DB, Aughey RJ, Anson JM, Gore CJ (2010b) Reproducibility of performance changes to simulated live high/train low altitude. *Med Sci Sports Exerc* 42: 394-401

Robertson EY, Saunders PU, Pyne DB, Gore CJ, Anson JM (2010c) Effectiveness of intermittent training in hypoxia combined with live high/train low. *Eur J Appl Physiol* 110: 379-387

Robertson JD, Maughan RJ, Davidson RJ (1987) Faecal blood loss in response to exercise. *Br Med J (Clin Res Ed)* 295: 303-305

Rodriguez FA, Ventura JL, Casas M, Casas H, Pages T, Rama R, Ricart A, Palacios L, Viscor G (2000) Erythropoietin acute reaction and haematological adaptations to short, intermittent hypobaric hypoxia. *Eur J Appl Physiol* 82: 170-177

Rossi F, Banuls O, Zorzoli M (2012) Cycling Anti-Doping Foundation Business Report 2010-2011.

Rusko HK, Tikkanen HO, Peltonen JE (2004) Altitude and endurance training. *J Sports Sci* 22: 928-945

Ryan BJ, Brothers MD, Nelson JL, Doan BK, Zupan MF, Prommer N, Byrnes WC Influence of carbon monoxide leaks on the measurement error of total haemoglobin mass. *Scand J Clin Lab Invest* 71: 523-528

Sanchis-Gomar F, Martinez-Bello VE, Gomez-Cabrera MC, Vina J (2010a) The hybrid algorithm (Hbmr) to fight against blood doping in sports. *Scand J Med Sci Sports* 20: 789-790; author reply 792-783

Sanchis-Gomar F, Martinez-Bello VE, Nascimento AL, Perez-Quilis C, Garcia-Gimenez JL, Vina J, Gomez-Cabrera MC (2010b) Desmopresssin and hemodilution: implications in doping. *Int J Sports Med* 31: 5-9

Saunders PU, Ahlgrim C, Vallance B, Green DJ, Robertson EY, Clark SA, Schumacher YO, Gore CJ (2010a) An attempt to quantify the placebo effect from a three-week simulated altitude training camp in elite race walkers. *Int J Sports Physiol Perform* 5: 521-534

Saunders PU, Cox AJ, Hopkins WG, Pyne DB (2010b) Physiological measures tracking seasonal changes in peak running speed. *Int J Sports Physiol Perform* 5: 230-238

Saunders PU, Pyne DB, Gore CJ (2009a) Endurance training at altitude. *High Alt Med Biol* 10: 135-148

Saunders PU, Pyne DB, Telford RD, Hawley JA (2004) Reliability and variability of running economy in elite distance runners. *Med Sci Sports Exerc* 36: 1972-1976

Saunders PU, Telford RD, Pyne DB, Hahn AG, Gore CJ (2009b) Improved running economy and increased hemoglobin mass in elite runners after extended moderate altitude exposure. *J Sci Med Sport* 12: 67-72

Sawka MN, Young AJ, Pandolf KB, Dennis RC, Valeri CR (1992) Erythrocyte, plasma, and blood volume of healthy young men. *Med Sci Sports Exerc* 24: 447-453

Schmidt W, Prommer N (2005) The optimised CO-rebreathing method: a new tool to determine total haemoglobin mass routinely. *Eur J Appl Physiol* 95: 486-495

Schmidt W, Prommer N (2010) Impact of alterations in total hemoglobin mass on VO 2max. *Exerc Sport Sci Rev* 38: 68-75

Schmitt L, Millet G, Robach P, Nicolet G, Brugniaux JV, Fouillot JP, Richalet JP (2006) Influence of "living high-training low" on aerobic performance and economy of work in elite athletes. *Eur J Appl Physiol* 97: 627-636

Schumacher YO, Ahlgrim C, Pottgiesser T (2008a) Evaluation of anthropometrical reference parameters for hemoglobin mass in endurance athletes. *J Sports Med Phys Fitness* 48: 509-514

Schumacher YO, Ahlgrim C, Ruthardt S, Pottgiesser T (2008b) Hemoglobin mass in an elite endurance athlete before, during, and after injury-related immobility. *Clin J Sport Med* 18: 172-173

Schumacher YO, Pottgiesser T (2010) Comment on "'The hybrid algorithm (Hbmr) to fight against blood doping in sports"' by F. Sanchis-Gomar et al. *Scand J Med Sci Sports* 20: 791; author reply 792-793

Schumacher YO, Pottgiesser T, Ahlgrim C, Ruthardt S, Dickhuth HH, Roecker K (2008c) Haemoglobin mass in cyclists during stage racing. *Int J Sports Med* 29: 372-378

Schumacher YO, Wenning M, Robinson N, Sottas PE, Ruecker G, Pottgiesser T (2010) Diurnal and exercise-related variability of haemoglobin and reticulocytes in athletes. *Int J Sports Med* 31: 225-230

Selby GB, Eichner ER (1986) Endurance swimming, intravascular hemolysis, anemia, and iron depletion. New perspective on athlete's anemia. *Am J Med* 81: 791-794

Sharpe K, Ashenden MJ, Schumacher YO (2006) A third generation approach to detect erythropoietin abuse in athletes. *Haematologica* 91: 356-363

Sharpe K, Hopkins W, Emslie KR, Howe C, Trout GJ, Kazlauskas R, Ashenden MJ, Gore CJ, Parisotto R, Hahn AG (2002) Development of reference ranges in elite athletes for markers of altered erythropoiesis. *Haematologica* 87: 1248-1257

Siebenmann C, Robach P, Jacobs RA, Rasmussen P, Nordsborg N, Diaz V, Christ A, Olsen NV, Maggiorini M, Lundby C (2012) "Live high-train low" using normobaric hypoxia: a double-blinded, placebo-controlled study. *J Appl Physiol* 112: 106-117

Sjostrand T (1948) A method for the determination of the total haemoglobin content of the body *Acta Physiol Scand* 16: 211-231

Sottas PE, Baume N, Saudan C, Schweizer C, Kamber M, Saugy M (2007) Bayesian detection of abnormal values in longitudinal biomarkers with an application to T/E ratio. *Biostatistics* 8: 285-296

Sottas PE, Robinson N, Giraud S, Taroni F, Kamber M, Mangin P, Saugy M (2006) Statistical clasification of abnormal blood profiles in athletes. *Int J Biostat* 2: Article 3

Sottas PE, Robinson N, Saugy M (2010) The athlete's biological passport and indirect markers of blood doping. *Handb Exp Pharmacol* 195: 305-326

Steiner T, Wehrlin JP (2010) Comparability of haemoglobin mass measured with different carbon monoxide-based rebreathing procedures and calculations. *Scand J Clin Lab Invest* 71: 19-29

Stewart IB, McKenzie DC (2002) The human spleen during physiological stress. *Sports Med* 32: 361-369

Stewart IB, Warburton DE, Hodges AN, Lyster DM, McKenzie DC (2003) Cardiovascular and splenic responses to exercise in humans. *J Appl Physiol* 94: 1619-1626

Stray-Gundersen J, Alexander C, Hochstein A, deLemos D, Levine BD (1992) Failure of red cell volume to increase with altitude exposure in iron-deficient runners. *Med Sci Sports Exerc* 24: S90

Suzuki K, Peake J, Nosaka K, Okutsu M, Abbiss CR, Surriano R, Bishop D, Quod MJ, Lee H, Martin DT, Laursen PB (2006) Changes in markers of muscle damage, inflammation and HSP70 after an Ironman Triathlon race. *Eur J Appl Physiol* 98: 525-534

Swiss Laboratory for Doping Analyses (2009) Athlete's passport haematological module interpretation technical document. Lausanne, Switzerland

Tadibi V, Dehnert C, Menold E, Bartsch P (2007) Unchanged anaerobic and aerobic performance after short-term intermittent hypoxia. *Med Sci Sports Exerc* 39: 858-864

Telford RD, Sly GJ, Hahn AG, Cunningham RB, Bryant C, Smith JA (2003) Footstrike is the major cause of hemolysis during running. *J Appl Physiol* 94: 38-42

Thomsen JK, Fogh-Andersen N, Bulow K, Devantier A (1991) Blood and plasma volumes determined by carbon monoxide gas, 99mTc-labelled erythrocytes, 125I-albumin and the T 1824 technique. *Scand J Clin Lab Invest* 51: 185-190

Toussaint HM, Hollander AP (1994) Energetics of competitive swimming. Implications for training programmes. *Sports Med* 18: 384-405

Truijens MJ, Rodriguez FA, Townsend NE, Stray-Gundersen J, Gore CJ, Levine BD (2008) The effect of intermittent hypobaric hypoxic exposure and sea level training on submaximal economy in well-trained swimmers and runners. *J Appl Physiol* 104: 328-337

Ulrich G, Strunz J, Frese F, Bartsch P, Friedmann-Bette B (2011) Dependence of hemoglobin mass estimation with the optimized CO-rebreathing method on different spectrophotometers. *Scand J Med Sci Sports* 22: 224-231

Videman T, Lereim I, Hemmingsson P, Turner MS, Rousseau-Bianchi MP, Jenoure P, Raas E, Schonhuber H, Rusko H, Stray-Gundersen J (2000) Changes in hemoglobin values in elite cross-country skiers from 1987-1999. *Scand J Med Sci Sports* 10: 98-102

Wagner PD (2006) Counterpoint: in health and in normoxic environment VO2max is limited primarily by cardiac output and locomotor muscle blood flow. *J Appl Physiol* 100: 745-747; discussion 747-748

Wehrlin JP, Zuest P, Hallen J, Marti B (2006) Live high-train low for 24 days increases hemoglobin mass and red cell volume in elite endurance athletes. *J Appl Physiol* 100: 1938-1945

Weight LM, Byrne MJ, Jacobs P (1991) Haemolytic effects of exercise. *Clin Sci* 81: 147-152

Wilber RL (2007) Application of altitude/hypoxic training by elite athletes. *Med Sci Sports Exerc* 39: 1610-1624

Wilber RL, Stray-Gundersen J, Levine BD (2007) Effect of hypoxic "dose" on physiological responses and sea-level performance. *Med Sci Sports Exerc* 39: 1590-1599

Withers RT, Gore CJ, Mackay MH, Berry MN (1991) Some aspects of metabolism following a 35 km road run. *Eur J Appl Physiol Occup Physiol* 63: 436-443

Wood MR, Dowson, M.N., Hopkins, W.G. (2006) Running performance after adaptation to acutely intermittent hypoxia. *Eur J Sport Sci* 6: 163-172

World Anti-Doping Agency (2009) Blood analytical requirements for the Athlete Biological Passport http://www.wada-ama.org/Documents/World_Anti-Doping_Program/WADP-IS-Laboratories/WADA_TD2010_BAR_EN.pdf. Accessed 23rd February

World Anti-Doping Agency (2010) Athlete Biological Passport operating guidelines and compilation of required elements http://www.wada-ama.org/Documents/Science_Medicine/Athlete_Biological_Passport/WADA_ABP_OperatingGuidelines_version_3.0.pdf. Accessed 23rd February

World Anti-Doping Agency (2012) The 2012 prohibited list: International standard http://www.wada-ama.org/Documents/World_Anti-Doping_Program/WADP-Prohibited-list/2012/WADA_Prohibited_List_2012_EN.pdf. Accessed 23rd February

Zhang GQ, Zhang DM, Song CJ, Yun KM (2007) The stability of carbon monoxide in stored blood samples. *Fa Yi Xue Za Zhi* 23: 60-62

Zorzoli M, Rossi F (2010) Implementation of the biological passport: The experience of the International Cycling Union. *Drug Test Anal* 2: 542-547

Zorzoli M, Rossi F (2012) Case studies on ESA-doping as revealed by the Biological Passport. *Drug Test Anal*: (Epub ahead of print)

Appendix 1

Estimates of BioWS variance included in five Adaptive models based on Hb_{mass}

Four studies have estimated the magnitude of within-subject variability of Hb_{mass} in athletes (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008). There is, however, inconsistency in the way in which this variability has been reported. Hb_{mass} is measured in grams (g). The estimates of Hb_{mass} variability published in the aforementioned studies have been either in the form of the *variance* (units: g^2), or *standard deviation* (SD: g), or *percent coefficient of variation* (%CV). These different units of variability are related mathematically, as follows:

$$SD = \sqrt{variance}$$

$$\%CV = (SD \div mean) \times 100$$

Furthermore, there is inconsistency between the fours studies (Eastwood et al. 2011b; Morkeberg et al. 2011; Pottgiesser et al. 2012; Prommer et al. 2008) in the extent to which they have broken down within-subject variability into its sub-components: analytical and biological variability.

$$Within\text{-}subject\ SD = \sqrt{(Analytical\ SD)^2 + (Biological\ SD)^2}$$

Whilst all four studies published estimates of the analytical variability, only two studies (Pottgiesser et al. 2012; Prommer et al. 2008) published estimates of the biological variability.

In the Athlete Biological Passport (ABP), separate estimates of the analytical and biological variability of Hb_{mass} must be entered into the Adaptive model. The analytical component must

be entered in the form of a %CV (specifically termed the Typical Error (TE)). The biological within-subject component must be entered in the form of a variance (termed the BioWS variance: g^2).

In this Appendix, the above equations are used to derive estimates of BioWS variance from the within-subject variabilities published in each of the four key studies. These estimates of BioWS variance are entered into five Adaptive models for use in the ABP, the sensitivities and specificities of which are examined in Chapters 8 and 9 of this thesis.

Study 1: Prommer et al. (2008)

Two different estimates of BioWS variance emerged from the data within this study. It was not clear which of these two estimates best represented the biological variability of Hb_{mass} in this cohort of athletes. Therefore, two models were created ($Hb^{m\ (Prommer\ -\ small)}$ and $Hb^{m\ (Prommer\ -\ large)}$), which contained different estimates of BioWS variance.

The first estimate of BioWS variance is the authors' calculated biological SD of 7.5g. This estimate is equivalent to a BioWS variance of 56.25 g^2 ($7.5^2 = 56.25$). ***Hence, 56.25 g^2 was the estimate of BioWS variance included in the $Hb^{m\ (Prommer\ -\ small)}$ model.***

The second estimate of BioWS variance from this study was derived from the within-subject variance of 408 g^2 shown in Table 4 of the publication (Prommer et al. 2008). The mean Hb_{mass} in this population was 912.7 g and the TE was 1.4%. The BioWS variance was derived as follows:

(i) Variance converted to SD: $\sqrt{408}\ g^2 = 20.199\ g$

(ii) SD converted to %CV: $(20.199 \div 912.7\ g) \times 100 = 2.213\%$

(iii) Biological %CV calculated: $2.213\% = \sqrt{(1.4)^2 + (Biological\ \%CV)^2}$

$$Biological\ \%CV = 1.714\%$$

(iv) %CV converted to SD: $1.714\% \times 912.7\ g = 15.644\ g$

(v) SD converted to variance: $15.644^2 = 244\ g^2$

244 g^2 was the estimate of BioWS variance included in the $Hb^{m\ (Prommer\ –large)}$ model.

Study 2: Pottgiesser et al. (2012)

Pottgiesser and colleagues (2012) published an estimate of BioWS variance of 550 g^2.

550 g^2 was the estimate of BioWS variance included in the $Hb^{m\ (Pottgiesser)}$ model.

Study 3: Morkeberg et al. (2011)

These authors published the within-subject variance of log(Hb_{mass}) as 0.001. A more precise value of 0.000912 was gathered from personal communication with Jakob Morkeberg (2012), the lead author of this publication. The mean Hb_{mass} in this population was 1064 g and the TE was 2.0%. The BioWS variance was derived as follows:

(i) Variance converted to SD: $\sqrt{0.000914} = 0.0302$

(ii) Back-transformed to %CV: $(e^{0.0302} - 1) \times 100 = 3.066\%$

(iii) Biological %CV calculated: $3.066\% = \sqrt{(2.0)^2 + (Biological\ \%CV)^2}$

$$Biological\ \%CV = 2.324\%$$

(iv) %CV converted to SD: $2.324\% \times 1064\ g = 24.723\ g$

(v) SD converted to variance: $24.723^2 = 611\ g^2$

611 g^2 was the estimate of BioWS variance included in the $Hb^{m\ (Morkeberg)}$ model.

Study 4: Eastwood et al. (2011b)

Eastwood et al. (2011b) published separate estimates of within-subject SD and analytical SD for male and female athletes. Within-subject SDs were 3.4% and 4.0% for males and females, respectively, and analytical SDs were 1.9% and 2.0% for males and females, respectively. The mean values of Hb_{mass} for males and females in this population were 1022 g and 691 g, respectively. The BioWS variances were derived as follows:

Male

(i) Biological %CV calculated: $3.4\% = \sqrt{(1.9)^2 + (Biological\ \%CV)^2}$

$Biological\ \%CV = 2.820\%$

(ii) %CV converted to SD: $2.820\% \times 1022\ g = 28.816g$

(iii) SD converted to variance: $28.816^2 = 830\ g^2\ BioWS\ variance$

830 g^2 was the estimate of BioWS variance included in the $Hb^{m\ (Eastwood)}$ model for male athletes.

Female

(i) Biological %CV calculated: $4.0\% = \sqrt{(2.0)^2 + (Biological\ \%CV)^2}$

$Biological\ \%CV = 3.464\%$

(ii) %CV converted to SD: $3.464\% \times 691\ g = 23.937g$

(iii) SD converted to variance: $23.937^2 = 573\ g^2\ BioWS\ variance$

573 g^2 was the estimate of BioWS variance included in the $Hb^{m\ (Eastwood)}$ model for female athletes.