UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE HEALTH AND LIFE SCIENCES

School of Medicine

"Beyond Reasonable Doubt - An Analysis of the Uncertainty behind a Positive Test for Growth Hormone Abuse in Sport"

by

Ioulietta Erotokritou-Mulligan

MSc, BSc(Hons)

Thesis for the degree of Doctor of Philosophy

June 2008

ABSTRACT

Growth hormone (GH) is a protein hormone which occurs naturally in the body and it is produced by the pituitary gland. The anabolic properties of GH have made this an attractive substance to be used by patients with catabolic medical problems and athletes that wish to enhance their athletic performance. However, to date, there is no scientific evidence to confirm that its use will enhance physical performance in sport. Aside from the morality issues regarding GH abuse in sports, the use of GH has been associated with a number of potential health risks.

GH is on the International Olympic Committee (IOC) Prohibited Substance list. To date however, there is no established method available for detecting its abuse reliably. In 1999 the GH-2000 research group proposed a GH detection methodology to the IOC. The method uses the concentration values of two GH-dependent markers; Insulin Growth Factor-I (IGF-I) and Procollagen type III peptide (P-III-P). Concerns were raised by the IOC regarding the uncertainty involved in this methodology on areas such as the effect of ethnicity, injuries and the physiological variability on its ability to find an athlete guilty beyond reasonable doubt. The GH-2004 project was set up at the University of Southampton in order to address these issues by undertaking several studies to assess the effect of these potential confounding factors.

This thesis used data collected by the GH-2004 project to examine the issues raised by the IOC. Furthermore, work was undertaken to validate the performance of the proposed method on independent data which will provide further evidence of the reliability of the GH-2000 detection method. An examination into the uncertainty surrounding the use of immunoassay methodology in this method was also carried out.

This thesis demonstrates that the GH-2000 detection method is robust when used for individuals of any ethnic group regardless of any sport injuries experienced by athletes. Furthermore, it is shown that the physiological variability of the concentration of the two GH-dependent markers will not adversely affect the performance of this method. The validity of this method is also successfully demonstrated using independently collected data.

Before this method is officially implemented, there is important laboratory analysis which needs to be undertaken, which will enable us to determine reliable cut-off limits, which will both maximise the sensitivity of the test but also reduce the rate of false positive results.

It has been shown that this methodology is a powerful method and should help to cut down the abuse of GH.

Keywords:

Growth Hormone, Doping, Sports, IGF-I, Injury, Ethnicity, Physiological variability, Immunoassays

ABBREVIATIONS

ALS	Acid-labile subunit
BMI	Body mass index
CV	Coefficient of variation
EF3	Discriminant function for females
EF3b	Age-adjusted Discriminant function for females
EM1	Discriminant function for males
EM1b	Age-adjusted discriminant function for males
FDA	Food and Drug Administration
GH	Growth hormone
hGH	Human growth hormone
ICTP	Carboxyl-terminal cross-linked telopeptide of type I collagen
IGFBP	IGF-binding protein
IGF-I	Insulin growth factor – I
IOC	International Olympic Committee
OLR	Ordinary Least-Squares Regression
Osteo	Osteocalcin
PICP	Carboxyl-terminal propeptide of type I procollagen
P-III-P	Procollagen type III peptide
rhGH	Recombinant human growth hormone
Te	Testosterone
USADA	United States Anti-Doping Agency
WADA	World Anti-Doping Agency

ACKNOWLEDGEMENTS

The principal investigators of the GH-2004 project were Dr. Richard I.G. Holt, Reader of Endocrinology and Metabolism at the University of Southampton and Professor Peter Sönksen, Emeritus Professor at the University of Southampton. Dr. Richard I.G. Holt was responsible for designing all studies covered in this project. Statistical input on the study design was made by Dr. Eryl Bassett, Senior Lecturer in Statistics at the University of Kent. The initial Ethics application for approval of these studies was made by Dr. Richard I.G. Holt. Subsequent revisions were made by the GH-2004 Research Fellow; Dr Cathy McHugh.

Recruitment for study participants was undertaken by members of the GH-2004 project team; myself, Dr Richard I.G. Holt, Dr Cathy McHugh, Ms Clare Hartley (Project Manager 2002-2004), Mr Rod Park (Project Manager 2004-2006) and the medical students: Jenny Bacon, Kate Harrison, Ben Curtis, Luke Smith, Victoria Wells, Barnaby Fontaine and Polly Milward. Furthermore, Dr. Richard Seah recruited athletes with a sporting injury for the GH-2004 Injury Study.

Compliance assessments, sample collection and anthropological measurements were undertaken by Dr Cathy McHugh or under the supervision of Dr McHugh by one of the medical students involved in the GH-2004 project. Data quality assessment was a collective effort by the whole of the GH-2004 project team.

Laboratory Analysis of samples collected by the GH-2004 project was undertaken by the WADA accredited laboratory at King's College in London by Professor David Cowan and Mr Christiaan Bartlett.

I would like to thank the nurses of the Wellcome Trust Clinical Research Facility at the University of Southampton for their assistance in recruitment and follow-up of the subjects.

Furthermore, I would also like to thank Dr. Astrid Kniess and Professor Alessandro Sartorio for sharing the results of their studies for the purposes of the analysis presented in this thesis.

I would also like to thank Professor Caroline Fall for taking the time to read my transfer thesis so thoroughly and offering me with valuable suggestions and comments during my Transfer thesis viva, which enabled me to improve the presentation of this final thesis.

In addition, I would like to acknowledge the support of Dr Richard Budgett, Director of Medical Services for the British Olympic Association. I would also like to acknowledge the legacy of the GH-2000 Project (a European Union 'Biomed 2' Project Contract umber: BMH4 CT950678). The GH-2004 study was funded by the World Anti-Doping Agency and US Anti-Doping Agency without which the studies described in this thesis would not have been possible.

With the oversight of my supervisors, editorial advice has been sought. No changes of intellectual content were made as a result of this advice.

SPECIAL ACKNOWLEDGEMENTS

I would like to genuinely thank my supervisors (in no particular order) Dr Richard Holt, Dr Eryl Bassett and Professor Peter Sönksen for being a true inspiration and for looking after me so well the last few years. I feel incredibly blessed for having such an amazing supervisory team. I could not have done this without their support and help.

I would also like to thank my parents, Erotokritos and Eve for believing in me and supporting me emotionally & financially all these years. Without their love, I would have probably given up a long time ago.

Also, I would like to thank my wonderful brothers Stefanos and Michalis and my wonderful sister Margarita, for sharing with me every sad and happy moment during the time of preparing this thesis, offering me their love, support and encouragement.

Last, but definitely not least, the person I owe all of this to, is my beloved husband Stewart. I would like to thank him for looking after me better than an angel could, with infinite patience and love. I hope I can now finally return some favours. Thank you.

To my angel, Stewart

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND TO THE RESEARCH

Growth Hormone is currently a popular doping agent in sport (Holt et al. 2008) and the current GH detection methodology adopted by the World Anti-Doping Agency can detect abuse for a maximum of 24 hours since last administration (Powrie et al. 2007). The use of GH as a performance enhancing substance carries some serious health risks (Holt et al. 2008; Powrie et al. 1995), gives doped athletes an unfair advantage (Barroso et al. 2008; Graham M.R. et al. 2008) and is against the International Olympic Committee's competition rules (The World Anti-Doping Agency 2006).

In 1996, the GH-2000 research group funded by the European Union set out to develop reliable GH detection methodology (Sonksen 2001). In 1999, this group proposed an important GH detection methodology, which would detect GH abuse up to 28 days since last administration (Sonksen 2001). Following this development the International Olympic Committee requested further validation of this methodology in order to be able to identify "beyond reasonable doubt" GH dopers with minimal risk of false accusation.

In 2003, the GH-2004 project was set up at the University of Southampton with the aim of validating this methodology further (McHugh et al. 2005). The results from the GH-2000 and the GH-2004 research groups, as well as the results from other independently run international research projects have been investigated in this thesis with the key aim of validating the GH detection methodology and ensure that any results from this methodology will be able to give evidence beyond reasonable doubt.

Cheating in sports has been prohibited in some form since ancient times (Yesalis C.E. et al. 2002). Doping is currently considered as cheating in sports, and it is prohibited by the International Olympic Committee (The World Anti-Doping Agency 2006) . The abuse of performance enhancing substances is unfair to fellow competitors and carries some potentially serious health consequences for the doped athlete.

It has been suggested that allowing doping in sport, will ultimately transfer competition from the sporting fields to scientific laboratories (Barroso et al. 2008). In such a situation, it is thought that competition will cease to be between the athletes, but will become one between scientists potentially with no consideration on the long-term health of the athlete, rather on the short-term win of an Olympic gold medal.

It is therefore important to develop and implement reliable procedures that will detect doping in sport. This, in conjunction with strict penal systems, will enable us to discourage and eventually stop doping in sports. This will protect athletes' health, and will allow a fair play between athletes. Currently, in many countries such as in Australia, doping is illegal. However, this has not been universally adopted by all countries yet, with countries such as the United Kingdom still trying to classify doping as an illegal offence (Yesalis C.E. et al. 2002).

Over the decades, a large variety of doping agents and methods has been identified, and in some cases, high profile athletes have been found guilty of substance abuse (Holt et al. 2008). The true extent of doping is not accurately known though, due to its nature. Substances such as steroids, THG, EPO and growth hormone are only a small list of the substances currently abused by athletes.

Growth Hormone (GH) is a popular doping agent amongst athletes, because of its lipolytic and anabolic properties. Athletes believe that these properties will give them an enhanced athletic performance. These properties combined with the fact that to date there is no reliable and robust methodology to detect GH abuse have made GH an ideal doping agent.

GH has been on the International Olympic Committee (IOC) Prohibited Substance list since 1989. The detection of GH abuse has posed a formidable challenge to a number of research groups across the world, and this is primarily due to the physiological properties of GH (Holt et al. 2008). Direct measurement of GH in plasma or urine is an inadequate method of detection. It is not possible to distinguish in plasma or urine between exogenous rhGH and endogenous GH as both types of GH consist of identical

amino acid sequences (Sonksen 2001). Furthermore, it is not possible to set physiological threshold levels of GH in plasma because its levels fluctuate significantly throughout a day and are significantly affected by stress, exercise, sleep and food intake (Kraemer W. et al. 2002). Only small and unpredictable amounts of GH are cleared through urine, and therefore this is also an inadequate method of detection (Saugy et al. 1996).

Immunoassays and blood sampling are required for the detection of these substances (Sonksen 2001). The GH-2000 research group proposed a GH detection methodology to the IOC. The method uses the concentration values of two GH-dependent markers; Insulin Growth Factor-I (IGF-I) and Procollagen type III peptide (P-III-P) (Dall et al. 2000; Longobardi et al. 2000). This method was derived using a double blind GH administration, placebo controlled study using healthy volunteers (Powrie et al. 2007). It is known that the levels of IGF-I and P-III-P in elite athletes are significantly higher than in normal subjects (Healy et al. 2005). Therefore, in order to adjust for these differences, the results from the double-blind study were calibrated against data collected from a cross-sectional study involving elite athlete volunteers within two hours of post-competition (Powrie et al. 2007).

Due to the underlying methodology used to develop the detection method there are a number of potential areas of uncertainty. It is therefore important to ensure that these areas are investigated further to ensure that the methodology is not significantly affected by this.

A limitation of cross-calibrating the GH detection methodology with a cross-sectional study is that volunteers in the cross-sectional study were only measured on one occasion of the sporting season. It is well known that training intensity and volume vary considerably during the season and the same may be true for the athletes' physiological and functional variables, including the biomarkers' concentrations (Felsing et al. 1992).

It has been suggested that GH secretion is significantly affected by acute exercise above certain intensity and the magnitude of the GH response is related to the peak intensity of exercise (Felsing et al. 1992; Kraemer et al. 1995). Furthermore, it is known that elite athletes are able to train at much higher intensities than the normal population, and that during a training season there are significant differences in training intensity.

Since many of the GH-related mediators, binding proteins and markers do not exhibit the same wide fluctuations as GH (Malcovati et al. 2003; Woitge et al. 1998), the purpose of this study was to investigate variations in IGF-I and P-III-P in athletes from selected sports. We can then determine whether the quantified variations in these GH-related markers could adversely affect the outcome of the proposed GH detection methodology.

In 1999, the GH-2000 research group proposed a GH detection methodology to the IOC. Concerns were raised by the IOC regarding the uncertainty involved in this methodology on areas such as the effect of ethnicity, injuries and the physiological variability on its ability to find an athlete guilty beyond reasonable doubt. The GH-2004 project was set up at the University of Southampton in order to address these issues by undertaking several studies to assess the effect of these potential confounding factors.

The primary aim of this thesis is to use data collected by the GH-2000 and GH-2004 projects and examine the issues raised by the IOC as well as other areas, which could raise levels of uncertainty. Work was also undertaken to validate the performance of the proposed method on independent sets of data, which will enhance its reliability (Erotokritou-Mulligan et al. 2007). An examination into the uncertainty surrounding the use of immunoassay methodology in this method was also carried out.

It is concluded that the proposed method can reliably detect GH abuse beyond reasonable doubt, as it is demonstrated that it works reliably in individuals of any ethnic group regardless of any sport injuries incurred by athletes. The test might

provide false positive results in the situation where an athlete has acromegaly, but under current WADA regulations, the onus is on the athlete to prove that their result is falsely positive due to such a condition. The accused has the right to request a second sample to be tested. If GH level is raised with the markers also high in more than one blood sample, that is highly suggestive and one would do something like an oral glucose tolerance test (in endocrine unit) to see if GH levels suppress to zero – which would then exclude the possibility for acromegaly.

Furthermore, it is shown that the physiological variability of the concentration of the two GH-dependent markers will not adversely affect the performance of this method. The validity of this method is also successfully demonstrated using an independent set of data. Before this method is officially implemented there is important laboratory analysis which needs to be undertaken, which will allow enable us to determine reliable cut-off limits which will both maximise the sensitivity of the test but also reduce the rate of false positive results. In conclusion, this thesis shows that the GH-2000 proposed method can reliably identify GH doping in sport beyond reasonable doubt with a high degree of sensitivity.

1.2 JUSTIFICATION FOR THIS RESEARCH

The GH-2000 detection methodology utilises the serum concentration values of IGF-I and P-III-P. The concentration of these markers were used to construct age adjusted gender specific formulae that gave good discrimination between those taking GH and those taking placebo (Powrie et al. 2007; Wallace et al. 2000).

These discriminant functions were calibrated against population specific reference ranges using elite athlete volunteers (Powrie et al. 2007). This was necessary because it is thought that the high training workload and psychophysical stress from competitions may modify their homoeostasis and affect their overall IGF-I and P-III-P levels (Powrie et al. 2007).

In order to ensure that the GH-2000 detection method will convict an athlete beyond reasonable doubt, it is necessary to examine the potential areas of uncertainty, where there is a possibility due to physiological properties of the GH-dependent biomarkers to accuse an athlete incorrectly.

The results of GH-2000 were reviewed by a panel of international experts (including a representative of the Court of Arbitration in Sport) at an IOC-organized Workshop in Rome in March 1999. The consensus was that the proposed GH-2000 detection methodology was an important step towards the detection of GH abuse. However, there were a number of areas in which the level of uncertainty was too high and would require further validation before this methodology could be widely implemented in Olympic Games.

In order to address these issues and any other areas where the test could falsely accuse an innocent athlete of GH doping, the GH-2004 research group was set up at the University of Southampton. The key aim of the GH-2004 research was to assess the significance and impact certain areas of uncertainty would have on the proposed Growth Hormone detection methodology as originally developed by the GH2000 research group.

The areas of uncertainty, which might affect the GH-2000 detection methodology, fall under three broad categories:

- Statistical uncertainty due to the statistical methodology used to determine doping
- Analytical bias due to the analytical methods employed to analyse the concentrations of IGF-I and P-III-P
- Uncertainty which arises because of the physiological properties of IGF-I and P-III-P.

This thesis, examines at the results obtained from both the GH-2000 and the GH-2004 study, as well as the results from research undertaken independently to assess the impact of these three areas of uncertainty.

1.3 METHODOLOGY

In order to be able to prove the validity of the GH-2000 detection methodology, the analysis of the research was separated in five key areas.

1.3.1 AREA OF UNCERTAINTY 1 –DEVELOPMENT OF ASSAY CONVERSION METHODS

WADA regulations require that any detection methodology can be reliably repeated using two independent immunoassay analytical methods. Due to the nature of immunoassays, often, measurements obtained with one method will not be directly comparable to the results obtained by a second method. Additionally, assay manufacturers often update their assays but also some assays might be discontinued. This therefore requires that any proposed GH detection methodology is not dependent on a particular immunoassay, and there is a process in place to ensure that results obtained with any assay method can be converted to a "gold" standard scale.

In the period between the GH-2000 and GH-2004 studies the immunoradiometric (IRMA) IGF-I assay was revised by its manufacturer (Nichols). This meant that IGF-I concentrations values given with the GH-2000 assay were different to the values given by the GH-2004 assay. Furthermore, it has been shown by a number of peer-reviewed publications (Quarmby et al. 1998; Khosravi et al. 2005; Juul 2005; Abellan et al. 2005) that the WHO International Reference Reagent (IRR) for IGF-I Immunoassays (87/518) is unreliable leading to numerical differences in the IGF-I assay results.

Furthermore, it was planned that identical assays would be used for all studies undertaken by the GH-2004 project. However, following successful completion of the GH-2004 cross-sectional study and halfway through the GH-2004 Injury study, in 2005, unrelated healthy and safety issues caused the Food and Drug Administration Authority (FDA) to prohibit Nichols Institute Diagnostics from trading, and it was not possible to use this IGF-I assay methodology as planned. Instead, serum samples

analysed using the DSL-5600 ACTIVE® IGF-I IRMA assay methodology manufactured by DSL Laboratories (Diagnostic Systems Laboratories, Inc., Webster, TX).

This made it necessary to derive assay adjustment methodology. The first reason was that although initially it was planned that identical assays would be used for the GH-2004 study as those used in the GH-2000 study. However, this was not possible and therefore in order to be able to utilise the GH-2000 detection strategy and to be able to compare results between the two projects. Similarly, assay adjustments might be necessary to be able to compare results derived by other independent research groups if they use assay methodology, which is different to the GH-2000 study.

1.3.2 AREA OF UNCERTAINTY 2 – EFFECT OF ETHNICITY

The majority of volunteers examined in GH-2000 research were of white European ethnic background. Therefore, it is necessary to examine whether findings concluded by the GH-2000 research group and in particular the GH-2000 detection methodology can be extrapolated to athletes of non-white European origins. Although currently there is one publication examining the effect of ethnicity on the concentrations of IGF-I and P-III-P, it is not known how ethnicity might affect the performance of the detection methodology. This key issue needs addressing before the GH-2000 method can be implemented at an Olympic Games event.

1.3.3 AREA OF UNCERTAINTY 3 - EFFECT OF PHYSIOLOGICAL INTRA-INDIVIDUAL VARIABILITY

One of the main reasons why direct measurement of serum GH cannot be used to identify GH doping is because it is not possible to set physiological threshold levels of GH in plasma because its levels fluctuate significantly throughout a day and are significantly affected by stress, exercise, sleep and food intake (Kraemer W. et al. 2002). Only small and unpredictable amounts of GH are cleared through urine, and therefore this is also an inadequate method of detection (Saugy et al. 1996).

A limitation of cross calibrating the GH detection methodology with a cross-sectional study is that volunteers in the cross-sectional study were only measured on one occasion of the sporting season. It is well known that training intensity and volume vary considerably during the season and the same may be true for the athletes' physiological and functional variables, including the biomarkers' concentrations (Felsing et al. 1992).

It has been suggested that GH secretion is significantly affected by acute exercise above certain intensity and the magnitude of the GH response is related to the peak intensity of exercise (Felsing et al. 1992; Kraemer et al. 1995). Furthermore, it is known that elite athletes are able to train at much higher intensities than the normal population, and that during a training season there are significant differences in training intensity.

Since many of the GH-related mediators, binding proteins and markers do not exhibit the same fluctuations as GH (Malcovati et al. 2003; Woitge et al. 1998), the purpose of this study was to investigate variations in IGF-I and P-III-P in athletes from selected sports and determine whether the quantified variations in these GH-related markers could adversely affect the outcome of the proposed GH detection methodology.

In order to do this, the intra-individual variation of the concentration of these two GH-sensitive markers were assessed for the GH markers IGF-I and P-III-P, from repeated blood samples collected from healthy trained individuals and elite athletes with respect to time and exercise.

In the present study, the aim was to quantify the intra-individual variability at resting conditions and post-exercise. Intra-individual variability will be attributable partly to inter-assay batch variability, and some will be attributable to the conditions (resting or post-competition) the sample was collected. As the test might be employed in either resting or post-competition conditions, we are not interested in specifically assessing intra-individual variability in resting conditions and intra-individual variability in post-competition conditions. A secondary objective of this study will be

to assess whether the intra-individual estimate could adversely affect the outcome of the proposed method of detection by the GH-2000 group using discriminant functions.

1.3.4 AREA OF UNCERTAINTY 4 – EFFECT OF INJURY

As GH dependent biomarkers used to formulate the GH-2000 detection method occur physiologically, detection of GH abuse must rely on finding concentrations in excess of those found in an established reference range. Although these markers are more stable in serum than GH and are relatively insensitive to the effects of exercise, they can vary widely among individuals, depending on age, gender, body weight, habitual physical activity, diet and sex steroids use (Healy et al. 2005; Kam et al. 2000; Sonksen 2001; Juul 2005). This makes it more difficult to define cut-off levels beyond which GH abuse could be proven (Sonksen 2001). To address this issue and in order to improve the sensitivity and specificity of any test compared with single marker analysis, the GH-2000 team, proposed a test based on the measurement of IGF-I and P-III-P. The values of the concentrations of these markers are used in specific equations, "discriminant functions" derived from the observed changes of these markers during a double-blind placebo-controlled rhGH administration study.

There have been fears that skeletal injury may have an adverse effect on the performance of the GH detection methodology. This is of particular concern as P-III-P is a marker of soft tissue and bone turnover (Haukipuro et al. 1990; Bail et al. 2001). For example, following a tibial shaft fracture, bone turnover increases for at least 24 weeks, as demonstrated by an increase in P-III-P (Veitch et al. 2006).

It is therefore important to examine whether elevations in either of these proteins following a sporting injury could lead to a false accusation of doping with GH. The aim of this study was to assess the effect of a musculo-skeletal injury on serum IGF-I and P-III-P concentrations in amateur and elite sportsmen and women and most importantly assess the effect of injury on the performance of the GH detection method proposed by the GH-2000 group.

1.3.5 AREA OF UNCERTAINTY 5 – VALIDATING THE METHOD USING INDEPENDENT DATA

It is important for any detection method to be shown that it performs reliably on data collected from independently run GH administration. This is necessary because there is always a potential positive bias in validating a methodology on the same data used to derive the methodology (self-fulfilling prophecy). Furthermore, it is not known accurately how athletes abuse GH in real life, so it is valuable to be able to apply the detection method in a variety of settings.

The Institut fur Dopinganalytik und Sportbiochemie in Kreischa undertook a placebo controlled double blind GH administration study and developed an alternative formula based on IGF-I and P-III-P but also including IGFBP-3 (Kniess et al. 2003).

The statistical procedure used to generate the discriminant functions involved splitting the available data into two; a "training" set (used to calculate the discriminant function) and a "confirmatory" set (used to validate the sensitivity and specificity of the discriminant function)(Strike 1991). The confirmatory set was required to ensure the model is applicable to the general population. Ideally, further validation is needed using a completely independent data set.

The aim of this study was to validate the GH-dependent marker approach by assessing whether the GH-2000 and Kreischa formulae could be used reliably to detect those receiving GH when applied to the alternate data set.

1.4 OUTLINE OF THE THESIS

This thesis begins with a literature review into the history of doping and discusses how doping has evolved over the centuries. This is an important part of this thesis as it sets the context of this work. Doping has been around in some form throughout the centuries all around the world. Surprisingly, prohibiting the use of performance enhancing substances is relatively a recent development. This chapter also includes a thorough examination in the reasons why athletes currently abuse GH. To do this well and to allow a correct understanding of the potential benefits and risks of GH abuse, this chapter looks at the role GH plays in the body system, as well as the currently approved clinical uses and purpose of GH use.

Following this, the thesis focuses next on the work that has been undertaken to date for developing reliable and robust methodology to detect GH abuse in sport by various research groups around the world. This section also goes into the current testing requirements as set by the World Anti-Doping Agency.

The next key area of this thesis is a detailed discussion of the GH detection methodology as proposed by the GH-2000 research group. This section discusses in detail how this methodology was developed along with its key strengths but also possible limitations. Additionally, a description of the work undertaken by the GH-2004 research group in order to ensure the validity of this methodology is given.

Following the literature review chapter, the thesis discusses the methodology undertaken to research the key aim of this thesis; which is an analysis of the uncertainty involved in the GH-2000 detection methodology for GH abuse. This "Methods" section discusses all areas related with the actual design of the research trials undertaken as well as the actual implementation of them. Areas such as subject recruitment process, analytical methods used and ethical considerations are all discussed here. Furthermore, a discussion of the statistical methodology employed to carry out the analysis is also given here.

After the "Methods" chapter, the "Results" chapters follow. These chapters give a detailed view of all the results undertaken for each area examined. Figures and Tables relating to each research area are presented here. Conclusions are presented alongside each study chapter discussing all the key findings of this research. Results from this thesis are put into context in reference to the GH-2000 detection methodology.

Lastly, the final section of this thesis is the "Discussion" section; this section discusses the current limitations of this work and offers suggestions for future research.

1.5 SUMMARY OF KEY FINDINGS

This thesis aimed to examine the validity of the proposed GH-2000 detection methodology for detecting GH abuse in sport. This methodology is based on a gender specific statistical scoring function which utilises the measured concentration values of two GH dependent biomarkers; P-III-P and IGF-I. In order to do this reliably and thoroughly, it was necessary to consider both the medical as well as the statistical science involved in developing this detection methodology. In doing so, a number of key areas of uncertainty were identified of either medical or statistical nature.

In particular, this thesis examined the effect the following areas:

1) The development of reliable statistical methodology, which would allow the GH-2000 detection not to be dependent on specific immunoassays. This methodology also overcomes the issue of the current lack of satisfactory international reference standards for IGF-I and P-III-P.
2) The potential effect of ethnicity on IGF-I and P-III-P and its effect on the scoring function
3) The potential effect of physiological variability on IGF-I and P-III-P and its effect on the scoring function
4) Validation of GH detection method on P-III-P and IGF-I values measured from an independent GH administration study
5) The potential effect of injury on IGF-I and P-III-P and its effect on the scoring function.

Overall results from the analysis undertaken in this thesis have shown that the proposed GH detection methodology by the GH-2000 research group has high sensitivity and specificity, and there is strong evidence to suggest that its implementation would enable identification of athletes who have been abusing GH beyond reasonable doubt.

1.6 BACKGROUND – LITERATURE REVIEW

1.6.1 HISTORY OF DOPING IN SPORT

1.6.1.1 OVERVIEW

> ***"You will not surely die," the serpent said to the woman. "For God knows that when you eat of it your eyes will be opened, and you will be like God, knowing good and evil."***
>
> ***Genesis 3:4-3:6***

It is human nature, that when man is placed in a competition setting they will attempt to have an advantage over their opponent, in order to achieve superiority. This is true in any environment, and in the field of sport, this is particularly apparent. Athletes competing at the ancient Games (as early as 668 B.C.) have been reported to have studied the effects of special diets in their performance (Yesalis C.E. et al. 2002).

In sport, an advantage over the opponent can be achieved in various ways. The sporting industry depends significantly on this superiority-driven desire of the athletes and has been capitalising on this for many years now. Nike, Reebok, Adidas, Speedo and other major sport companies, spend millions of dollars a year in improving sporting equipment and apparel, promising athletes improved performance, and most importantly an advantage over their opponent. Shirts that are light and keep athletes cool and dry during competition, the use of flat seaming to reduce abrasion are just some of the new benefits featured in Nike's Pro campaign (Nike 2005). In the same way, Adidas have developed shoes that will "remove heat and sweat, providing optimal cooling during competition and training". The

TORSION® System developed by Adidas, "allows natural rotation between the rear foot and the forefoot while providing mid-foot support" (Adidas 2006).

An advantage can also be sought, by using new forms of training, and there are many sport-training professionals who have published training guides promising to teach athletes the "winning principles". Many of these methods used to achieve superiority will often work, but just as often, these methods will not. It can be considered a certainty however, that athletes will always be seeking out these new methods hoping to gain an advantage, and improve their performance.

When all legitimate methods have been implemented, and training has achieved the athlete's peak performance; some athletes might seek out chemical and pharmacological methods to improve performance even further. The use of these methods will often help athletes reach their goal more easily, superseding their own physical limits.

Defining which drugs and methods constitute doping is a major challenge. There is no universally agreed definition for all sports. Even the origin of the word "doping" is controversial. Some claim that the word doping is originally derived from the African Kaffirs who used local liquor called "dop" as a stimulant (Kent M. 1997; McHugh et al. 2005; Yesalis C.E. et al. 2002). Others claim that the word "doping" originates from the American expression "dope", a slang word for opium since in the days when doping of horses was first attempted opium was the most common used substance (Csaky 1972).

Csaky, argued that the first case of doping could be traced back to Adam and Eve at the Garden of Eden. In the hope of becoming "Godlike", they ate the forbidden apple, despite having being warned against this by God (Csaky 1972). This argument aims to show that humans have always wanted to better themselves since the beginning of the world, as documented in the Holy Bible, despite the consequences. Humans will often go beyond what is right, in order to achieve superiority over an opponent.

In Ancient Greece, from where the Olympic Games originate, there were stone pedestals lining the entrance of the Olympic stadium. Athletes violating Olympic rules had their names inscribed on these stone pedestals, along with the names of their families and the details of their offence. The athletes were banished for life from the games. This "inscription" punishment also served as a warning to athletes wishing to compete in the stadium, of the shame and punishment they would endure if they were to cheat while competing.

The history of doping - cheating in order to excel - is as ancient as sporting competition itself, but at the same time, the hunt for cheaters is just as old. However, cheating offences did not include the use of substances taken with the aim to perform their athletic performance. An example of a cheating offence, found inscribed on these stones was the bribing of an official. Various explanations can be offered for this; the use of substances to enhance performance might not have been considered to offer a significant advantage over opponents or possibly because it would have been nearly impossible with the technology of the time to prove substance abuse.

It appears that the use of performance enhancing substances was considered for the first time as inappropriate as late as the end of World War I (Hoberman 1992). Hoberman reports that until the 1920s there was little attempt to prohibit the use of performance enhancing substances in a sport competition setting, and even less attempt was made to discourage such use. The first evidence of such discouragement can be found in the writings of Dr Otto Rieser from 1933, in his work "Doping and Doping Substances" (Yesalis C.E. et al. 2002).

It was not until 1967 when the International Olympic Committee first voted to adopt a drug-testing policy banning the use of specific drugs. In an article written by B. Gilbert in 1969 in Sports Illustrated, it was reported that despite the vote undertaken by the IOC in 1967, at the time of writing his report, no single US sporting organisation had specific apparatus to enforce any of the bans voted (Yesalis C.E. et al. 2002).

The first reported test to detect the voted banned substances, did not take place until 1982, when the National Football League (NFL; in USA) performed the first drug testing. Over the years, the International Olympic Committee has reviewed its Prohibited Substances list several times, and today this is enforceable at all Olympic events (Yesalis C.E. et al. 2002). However, despite the fact that the Prohibited Substances List is extensive and enforceable at Olympic events, it does not apply at all sporting events. Professional sports in the United States that do test for drugs have programmes that on average are substantially less rigorous than the IOC program.

1.6.1.2 EARLY HISTORY OF DOPING DEVELOPMENTS

There are many examples in the literature of substance use by ancient people. These substances were derived from plants, animals or even humans, and were used by the ancients either as performance enhancing substances or as medication. An extensive report of this can be found in the book edited by Yesalis and Bahrke (2002). A summarised overview of these uses is given here.

One of the first performance-enhancing substances to be tried was testosterone. The ancients first identified the properties of this hormone, by examining the behaviour of their animals following castration.

The ancients have been reported to eat the organs of other animals or humans to improve or heal their own (Yesalis C.E. et al. 2002). The ancients also seemed to understand that diet type had an effect on their performance. In sport competitions, the earliest report describing the use of a special diet to improve performance is of a man called Charmis; the Spartan winner of the stade race (~200 yards [183 metres]) in the Olympic Games of 668 B.C.. Charmis allegedly used a special diet of dried figs in order to improve his performance (Yesalis C.E. et al. 2002).

The Ancient Greeks were also the first to report the use of stimulants as part of their sport training routine. They used to drink potions of brandy and wine, and ate

hallucinogenic mushrooms and sesame seeds, with the aim of improving their performance (Voy 1991). The gladiators in the Roman Coliseum have also been reported to have used unspecified stimulants to overcome fatigue and injury (Voy 1991).

Most of the stimulants used in these early times, came from plants. Examples are found in Boje (Yesalis C.E. et al. 2002) who reported the use of bufotein; a drug derived from fly-garic (Amanita muscaria), a mushroom containing muscarine (a deadly alkaloid). He suggests that the legendary Berserkers of Norse mythology used this to increase their strength. He also discussed the use of Cola acuminita and Cola nitida for running competitions, by West Africans in ancient times. Jokl (Jokl 1968) discusses how Andean Indians of Peru, chewed coca leaves or drunk coca tea to increase endurance and protect themselves against mountain sickness.

1.6.1.3 19TH CENTURY DEVELOPMENTS IN THE USE OF STIMULANTS & ANABOLICS

In the latter half of the 19th century, modern medicine was developing, and as a result, the number and types of drugs and substances used to improve performance escalated. Stimulants were now primarily used as ergogenic aids - means of increasing muscular work capacity. Furthermore, scientists also began investigations into the anabolic effects of substances that were later classified as hormones.

Caffeine is considered to have been the main substance used in this period, as it was recognised to improve brain functioning. Alcoholic drinks were also considered useful, as a stress relief aid. Athletes were now commonly using stimulants to improve their performance and gain an advantage. As there were no rules against using such substances, athletes did not try to conceal this, and as a result, there are good records on doping from this time. Trainers were also developing their own doping recipes, using combinations of various stimulants, such as strychnine tablets, mixtures of brandy and cocaine.

In the 19th century, the "Six Day" bicycle races began - a continuous race of 144 hours. Such races required athletes to have great physical strength and stamina. Apart from the tremendous amounts of training the athletes required for such an event, the cyclists experimented with a variety of doping mixtures to enable them to perform well and have an advantage over their opponents. The French were reported to be using mixtures with caffeine bases, the Belgians used sugar cubes dripped in ether, and others used alcohol–containing cordials, while the sprinters specialised in the use of nitro-glycerine (Prokop 1970; Yesalis C.E. et al. 2002). As the race would progress, the athletes would increase the amounts of strychnine and cocaine added to their caffeine mixtures. It is therefore not surprising, that the first doping fatality occurred during such an event. This happened in 1886 to Arthur Linton an English cyclist, who is said to have overdosed on "tri-methyl" (thought to be a compound containing either caffeine or ether), during a 600km race between Bordeaux and Paris (Prokop 1970). It should be noted that there is dispute over this, as others (Donohoe 1986) suggest that Linton actually won this race and did not die until 10 years later from typhoid fever.

A second sport, which utilised much of this development and experimentation in the area of stimulants, was the "ultramarathon". This was a race in which athletes would walk for 6 days and 6 nights continuously, with the winner being the person who would cover the greatest distance. It may be easily appreciated how the use of stimulants would have been beneficial in this sport, and there are reports that mixtures of champagne, brandy, hot drops of morphine, belladonna and strychnine were used to aid the athletes to have keep high levels of strength and energy (Osler et al. 1979).

The 19th century saw the birth of organotherapy, with Charles Édouard Brown-Sequard, an eminent physiologist and neurologist. He presented his findings to the Society of Biology in Paris, reporting radical findings following a three-week programme of self-injections "first, blood of the testicular veins; secondly semen; and thirdly juice from a testicle . . . from a dog or a guinea pig"(Brown-Sequard C.E. 1889).

Although many believed that he had merely experienced a placebo effect, the idea of hormone replacement was conceived and now Brown-Sequard is considered the "father of modern endocrinology". His research began a series of experiments throughout the western world, investigating this method for rejuvenation purposes as well as a treatment for various diseases.

It is unsurprising that this research was identified as an area that should be investigated as means of improving athletic performance, and in 1894 Oskar Zoth & Fritz Pregl assessed the effect of testicular extracts on muscular strength. With hindsight it is considered unlikely, these testicular extracts were contributing positively to the performance of athletes - however, Zoth and Pregl while investigating the effect on athletic performance, provided the grounds where such research could be furthered.

1.6.1.4 20TH CENTURY DOPING PANDEMIC

Following the research initiated by Charles Édouard Brown-Sequard, the 20th century saw major developments in the world of endocrinology and sport doping. Scientists had now isolated, chemically characterised, synthesized testosterone and identified the basic nature of its anabolic effects. The first recorded case of using testosterone as a means of improving athletic performance, was in 1941, in an 18-year-old horse named Holloway, who following treatment had improved his performance significantly, winning many races (Kearns et al. 1942). Not long after this incident, sport trainers realised that they would be able to improve the performance of their athletes by getting them to use these anabolic hormones. Photographs of body-builders of this period are highly suggestive of testosterone and anabolic steroids uses, with athletes having deformed body shapes with extremely large muscles beyond what would be physiologically possible through extensive training.

In the mid 1930s, the doping world saw an increased use of amphetamines. Although amphetamines were first identified in 1887, it was not until later, that athletes saw the benefits of using these substances when competing. It was reported that amphetamines were systematically used during the Second World War by the British army. A report in the Air Surgeon's Bulletin says: ". . . one pill (Benzedrine) may be worth a Flying Fortress when the man who is flying it can no longer stay awake" (Yesalis C.E. et al. 2002). The benefits experienced by these servicemen, along with the use of amphetamines by college students, encouraged athletes to include the use of amphetamines in their "drug diet".

It appears that apart from bodybuilding, cycling was another sport that had a pivotal role in stimulant use in sport. Prokop in 1970 characterised cycling competitions as "special hotbeds of doping". The 1960s and 1970s have been characterised as the "amphetamine decades" for anyone competing in cycling. During the 1967 Tour de France, the world witnessed the first televised doping fatality, of English cyclist Tom Simpson. The autopsy of his body proved that he had died of high levels of methamphetamine. In the following decade, 1980s, anabolic steroids and cortisone were mainly used, and from there on erythropoietin (EPO), blood transfusion and human growth hormone. Just hours before the 2000 Tour de France was to begin, three cyclists failed a mandatory EPO test and were expelled from competition.

By the early 1980s and beyond, the use of human growth hormone had become well established on the body building community's drug menu. In 1982 Fred Hatfield, wrote a controversial book, naming Human growth hormone as "the 'state of the art' strength and size drug in the free world" (Yesalis C.E. et al. 2002).

1.6.1.5 EAST GERMANY DOPING SYSTEM

Following the collapse of the German Democratic Republic (GDR) in 1990, a number of classified documents were uncovered which described the promotion by the government of the use of drugs in high-performance sports, notably androgenic

steroids. These documents included confidential doctoral theses, scientific reports, progress reports of grants and reports of physicians and scientists who served as unofficial collaborators for the Ministry for State Security ("Stasi"). These documents describe in detail that hundreds of physicians and scientists, performed doping research and administered prescription drugs as well as unapproved experimental preparations since 1966. The number of athletes involved in these trials as discussed in these documents were several thousand and they were treated with androgens for many years, including young athletes. Results recorded in these studies have indicated that treatment with androgens showed this to be highly effective in their sports performance, however with some serious side effects which often required surgical or medical intervention (Franke et al. 1997).

1.6.1.6 DEVELOPMENTS IN THE MODERN OLYMPIC GAMES

In the modern Olympic Games, with the developments of medical science, the availability and the variety of these substances increased. In 1904, Thomas Hicks, winner of the marathon, was administered strychnine and brandy several times during the race. In 1932, at the Los Angeles Olympic Games, Japanese swimmers were said to be "pumped full of oxygen"(Hoberman 1992). As the research in these substances increased, so did the use. During the 1968 Olympic Games in Mexico City, athletes and coaches were only concerned in identifying the most effective drugs – rather than being worried about any potential health hazards or moral issues.

Anabolic steroids were referred to by the then editor of Track and Field News in 1969 as the "breakfast of champions". Another magazine, based on numerous interviews, suggested that "athletes were popping more pills for more purposes than were dreamt of in anybody's philosophy – or pharmacy"(Yesalis C.E. et al. 2002).

In the decades to follow, the use of performance enhancing drugs showed no decreased use, rather an increase. In 1988 Seoul Games, Ben Johnson, was tested positive for stanazol, an anabolic steroid, and subsequently confessed under oath in court that he

had used growth hormone as well. A subsequent investigation by the New York Times concluded that "at least half of the athletes who competed in Seoul used anabolic steroids to enhance their performances..."(Yesalis C.E. et al. 2002). The 1996 Atlanta Olympic Games were referred to as the "Growth Hormone Games" by some of the athletes (Yesalis C.E. et al. 2002).

Bamberger and Yaeger in 1997 concluded following a lengthy investigation, that there were three distinct classes of athletes competing at the Olympic sports:

- In the biggest group were most athletes, using performance enhancing drugs and not detected by the tests of the International Olympic Committee
- A smaller group who did not use these drugs
- A third group comprised of just a handful of athletes who used these performance-enhancing drugs and actually were detected.

The 2000 Sydney Olympic Games were scarred as the "Dirty Games", due to the large number of articles published on athletes taking performance-enhancing drugs.

In 2000, the U.S. Office of National Drug Control Policy concluded that the doping epidemic varied widely from 10% to 90% of the athletes. However, what appeared to be the bigger problem in this picture was that athletes were seeming to be encouraged to use these prohibited substances, due to the "high financial stakes for Olympic athletes, corporate sponsors, the TV broadcast and cable industries and sport governing bodies, coupled with the pharmacopoeia of performance enhancing substances, the athlete's drive to win, and the absence of an effective policing mechanism" (National Center on Addiction and Substance Abuse 2000; Yesalis C.E. et al. 2002).

Recently, at the 2004 Athens Olympic Games 24 doping violations were reported. An article (Bryn Palmer 2004) by the British Broadcasting Cooperation (BBC), reported that at the Athens Olympics, there were twice as many doping violations than there was at the Los Angeles Olympic Games in 1984, which until the Athens Games, was known as the "dirtiest" Olympic Games.

At the 2006 Winter Olympics in Turin, police officials, confiscated suspect medical equipment from the residence of Austrian athletes, which included several packs of drugs, and blood transfusion equipment and doping officials tested 10 of the athletes. Although it has not been determined yet how many athletes were doping at these Games, the message is clear; athletes are still interested in the use of pharmacological tools to enhance their athletic performance.

1.6.1.7 FOOD SUPPLEMENTS

The International Olympic Committee and WADA prohibit many chemical and hormonal substances that offer an unfair advantage to athletes. However, an alternative way to attempt performance enhancement is the use of nutritional supplementation, which has gained popularity as a way to achieve this "legally". Food supplements are widely available, and an individual can buy these from local gyms to health shops.

The United States Congress defined a "Dietary Supplement" in the Dietary Supplement Health and Education Act (DSHEA) of 1994 as a "product taken by mouth that contains a dietary ingredient intended to supplement the diet". The dietary ingredients in these products may include vitamins, minerals, herbs or other botanicals, amino acids and substances such as enzymes, organ tissues, glandulars and metabolites. Dietary supplements can also be extracts or concentrates and be found in many forms such as tablets, capsules, soft-gels, gel-caps, liquids or powders.

In the United States, dietary supplements do not need approval from the Food and Drug Administration body (FDA) before they are marketed, as they are covered under the umbrella of "food" and not "drugs". Drug products must be proven to be safe and effective for their intended use before marketing, but there are no provisions in the U.S. law for the FDA to approve dietary supplements for safety or effectiveness before they reach the consumer. In addition, unlike drug products, manufacturers and distributors of dietary supplements are not currently required by law to record,

investigate or forward to FDA any reports they receive of injuries or illnesses that maybe related to the use of their products. Once the product is marketed, the FDA has the responsibility for showing that a dietary supplement is unsafe before it can take action to restrict the product's use or removal from the marketplace (U.S.Food and Drug Administration - Center for Food Safety and Applied Nutrition 2002). If the labelling of the dietary supplement makes any claims concerning its efficacy, then it might be classed as a medicinal drug.

Despite an apparent lack of scientific evidence to support the alleged benefits of dietary supplements, the retail sales in the USA generated $3.3 billion in 1990 and $13.9 billion in 1998 and this was estimated to be tripled by the end of 2008. Furthermore, there have been suggestions that not only is there no significant evidence to suggest that nutritional supplements will enhance physical performance, but also that the prolonged use of nutritional supplements as well as using these supplements in high doses could pose a significant health risk (Maughan 2005).

Although food supplements are not prohibited by WADA or the IOC, it has been discussed in recent studies that these supplements can often be contaminated with prohibited substances such as anabolic androgenic steroids (including testosterone and nandrolone as well as the pro-hormones of these compounds), ephedrine and caffeine. These substances will not be listed on the package labelling of the supplements and the use of these substances will inevitably lead to an athlete testing positive for use of prohibited substances. The IOC/WADA regulations do not accept the "innocent" use of these substances as a valid excuse for the existence of these substances in an athlete's sample (Striegel et al. 2005; Maughan 2005; Maughan 2004).

1.6.1.8 FUTURE ISSUES IN DOPING: GENE DOPING

Athletes are always seeking to find new ways of improving their physical performance by searching out the most efficient method, and the most undetectable method. Gene

doping is currently considered by the scientific community as being the future of doping. It is known that athletes have always been ahead of the "doping game" with athletes abusing substances like Growth Hormone 10 years before they were even accepted by the medical community as an effective medical treatment (Sonksen 2001).

The World's Anti-Doping Agency has defined gene or cell doping as the "non-therapeutic use of genes, genetic elements and/or cells that have the capacity to enhance athletic performance"(The World Anti-Doping Agency 2006). Gene therapy has recently been shown to be of benefit for the treatment of serious diseases such as in patients with haemophilia (Kay et al. 2000) and patients with x-linked severe combined immunodeficiency disease. Angiogenic gene therapy shows possible positive signs in improving angina (Hanif et al. 2005).

It has been discussed (Haisma et al. 2004) that gene doping could be used (or might already be in use) in a variety of athletic settings. In sport, gene therapy might be used in the future to regenerate the tissue defects following trauma in by an improved process, which would minimise the effect of sport injuries to an athlete. Furthermore, gene therapy has the potential to modify normal metabolism. Erythropoietin (EPO), human growth hormone (hGH), insulin-like growth factor-1 (IGF-I), peroxisome proliferator activated receptor-delta (PPAR δ), and myostatin inhibitor genes have been identified as primary targets for doping (Sweeney 2004; Defrancesco 2004; Unal et al. 2004).

The use of gene therapy carries a number of potential risks, the main risk being the fact that gene therapy has only been carried out in tightly controlled environments, where the gene transfer vectors have been extensively tested for toxicity and safety, so the consequences of using in an uncontrolled environment, without proper training and understanding would be risky. Furthermore, other risks, depending on the specific protein expressed in gene doping will be similar to those of doping using other methods for the concerned protein. So for example if gene doping is used to increase the Erythropoietin (EPO) levels, athletes will run the risk of blood clots caused by the increased density of the blood. In contrast to the injection of synthetic EPO which can

be stopped post competition allowing normal physiology to resume the case of EPO being delivered by gene doping, it is not known how safely the production of EPO may be lost indefinitely after gene doping, thus possibly causing many uncontrolled fatalities (Haisma et al. 2004). Similarly, in the case of gene therapy for increasing IGF-I or GH production, the risks of acromegaly and tumour development exists.

Detection of gene therapy for the purposes of enhancing performance is difficult, mainly because any effects are "naturally" produced by the body and the DNA which is used for gene transfer is of human origin. A possibility for detecting gene doping is using methodologies currently considered for supraphysiological amounts of proteins or by detecting the vector used for the gene therapy.

It is clear that the area of gene doping is associated with significant health risks, some of which are unknown because it is a relatively new area. It is not accurately known what the potential dangers might be in the situation where individuals not qualified in this area experiment with this, particularly in the case of sport doping, where the sacrifices are often huge for the promise of fame, glory and a gold medal.

1.6.2 GROWTH HORMONE DOPING

Human growth hormone (hGH) is one of the newest drugs to be exploited for competitive advantage by athletes. However, it is interesting to note that "[athletes]... were doing so 10 years before endocrinologists recognised and understood its potency as an anabolic agent" (Sonksen 2001). In the mid-1980s, recombinant human growth hormone (rhGH) was developed as a safer alternative to cadaveric growth hormone. As the latter was extracted from the pituitary glands of cadavers, treatment led to the of transmission of potential fatal prion infections such as Creutzfeldt-Jakob disease (CJD) (Brown et al. 2006; Heinemann et al. 2007; Hirst 2005; Narayan et al. 2005).

Growth Hormone was first promoted as potent performance-enhancing anabolic agent in "The Underground Steroid Handbook" first published in California in early 1982 by Daniel Duchaine (Duchaine D. 1982). The book encouraged athletes to use growth

hormone to enhance their performance. Specifically, Duchaine stated in his book: "Wow, is this great stuff! It is the best for permanent muscle gains....People who use it can expect to gain 30 to 40 pounds of muscle in ten weeks." (Duchaine D. 1982). Duchaine was known in the body building community as the "steroid guru" and was responsible for the development of many doping trends. Through experimentation on him and other body builders he was coaching, he advised athletes on the best substances to use and best combinations for these for best performance results.

Duchaine later retracted this statement, as he could not find any scientific studies that supported his claims. In his later book published in 1993 "Ultimate muscle mass" he stated that "I'd guess that almost 90% of all athletes taking STH [growth hormone] got no anabolic results from it (this includes at least two Mr. Olympia competitors)" (Duchaine D. 1982; Haycock 2006). However, the damage had already been done by that time as the use of growth hormone was proliferating among the doping world.

One of the most well known recorded examples of GH misuse was uncovered following the Dubin enquiry, where Ben Johnson, admitted taking growth hormone to boost his performance, amongst other drugs during his competition in the 1988 Seoul Games. The Dubin enquiry was set up following the 1988 Seoul Olympic Games doping scandal to investigate drug taking in Canadian sport. Justice Charles Dubin, head of this enquiry, conducted one of the most searching investigations in history, into the abuse of drugs in sports. During these hearings, the Canadian sprinter Angella Issanjenko also admitted using hGH along with other drugs. According to the 900-page Dubin inquiry report (Dubin C.J. 1990) it was concluded that the tight regulations of the use of GH had not prevented it from becoming available to athletes.

In 1989 the IOC included growth hormone in its prohibited substance list, as part of a new doping class of "peptide hormones and analogues" despite the lack of a legitimate test for hGH (Kicman et al. 1992). However, this did not seem to deter use of growth hormone as a performance enhancing substance. For many athletes, it became the new drug of choice (Hoberman 1992), with some athletes calling the 1996 Olympics the 'hGH games' (Zorpette 2000).

There are no official incidence figures to show how common GH usage is amongst athletes. As Sönksen states in his 2001 article, most of the evidence is circumstantial. In 1998 a member of the Chinese swimming team on its way to the World Championships in Perth, was found with 13 vials containing growth hormone (Zorpette 2000). The incident of a pharmaceutical importer's loss of 1575 GH vials, six months before the 2000 Olympic Games, mentioned in Sönksen's article (Zorpette 2000), raised concerns over how clean the Sydney Games were.

Over the years, many accusations have been made against athletes claiming that they have been abusing growth hormone as a performance enhancing substance. However, due to the lack of a reliable detection method, the WADA and the IOC have been unable to prove this. Track stars Marion Jones, Tim Montgomery, NFL players such as Bill Romanowski, and sluggers including Barry Bonds, Gary Sheffield and Jason Giambi, are amongst some of the athletes who have been recently accused to be taking growth hormone in the controversial book "Game of Shadows", which was written following undercover investigations of two San Francisco reporters (Williams L. et al. 2006).

One of the most infamous cases of growth hormone abuse came after the raid on the Bay Area Laboratory Co-Operative's (BALCO) headquarters on 3 September 2003. Victor Conte, the owner of the Bay Area Laboratory Co-Operative (BALCO), claimed that he had supplied GH to many high-profile American athletes including Tim Montgomery and Marion Jones. Marion Jones, 5 times Olympic medal winner, admitted in 2007 use of performance enhancing drugs – including the use of growth hormone. She was later sentenced to a 6-month jail sentence for falsely denying administering performance-enhancing substances. Tim Montgomery allegedly admitted to taking GH before a US Federal grand jury and later faced a 2-year ban for doping offences. Conte was imprisoned for 4 months for his role in the scandal. (Williams L. et al. 2006).

It is believed by athletes that the use of growth hormone will improve athletic performance. However, to date there is only a small amount of some published

evidence to support such claims (Holt et al. 2008). It has been shown that normally growth hormone has a key role in regulating body composition. Research studies have shown that growth hormone will divert calories in food towards protein synthesis ('Partitioning Agent') (Dubin C.J. 1990; Sonksen 2001). This means that muscle mass will be increased while fat stores reduce. Athletes believe that increased muscle mass implies increased muscle strength, and that is one of the primary reasons for its use. A second reason, is as mentioned earlier, is the fact that although growth hormone is in the Prohibited List, to date, there is no robust and reliable method that will detect its use but elite athletes would not take the risks inherent in doping if they did not believe that they were obtaining benefit. It is also likely that they can evaluate new doping substances objectively on their own performance more easily than scientists can using the relatively blunt instrument of a randomised controlled trial (RCT).

Through a number of scientific studies, it has been shown that administration of growth hormone will promote protein accretion and it will increase muscle fibre hypertrophy (Berggren et al. 2005; Doessing et al. 2005; Fryburg et al. 1991; Fryburg et al. 1993; Janssen et al. 1999).

Until recently however, published reports from controlled studies suggested no significant performance enhancing effect of hGH administration in young healthy adults (Berggren et al. 2005; Jenkins 1999; Rennie 2003; Yarasheski et al. 1992; Yarasheski et al. 1993; Yarasheski et al. 1995; Zachwieja et al. 1999).

In addition, in the competitive environment, growth hormone will not always be used alone, and it has been reported anecdotally that it is often used in conjunction with other anabolic substances (Kraemer W. et al. 2002; Karila et al. 1998; Dubin C.J. 1990). A recent study by Giannoulis et al involving healthy elderly men showed that there was significant evidence to suggest an additive effect with concomitant administration of testosterone and growth hormone (Giannoulis et al. 2006). Performance enhancing claims for GH by athletes could potentially be attributable to the concomitant use of other anabolic substances although the study by Giannoulis showed GH to be a more powerful anabolic agent than testosterone. A very recent

randomised control trial of rhGH administration to ex-steroid abusers has shown clear performance-enhancing effects (Graham M.R. et al. 2008).

A further attraction to the use of GH by athletes is not just that it is thought to produce a similar effect to anabolic steroids, but also that there is currently no reliable method to detect GH use. The World Anti-Doping Agency, has announced that they implemented a pilot test to detect GH dopers at the 2004 Athens Olympic Games and the 2006 Winter Games in Turin but no positive results were identified.

1.6.3 THE GH-IGF-I AXIS

1.6.3.1 GROWTH HORMONE PHYSIOLOGY

1.6.3.1.1 THE GH MOLECULE

GH is a polypeptide hormone comprising 191 amino acid and it accounts for 10% of the pituitary hormonal content. It is stored and secreted by the somatotrophs located within the lateral aspects of the anterior pituitary (Kraemer W. et al. 2002). The predominant circulating form (75%) has a molecular weight of 22 kDa and has a half-life in the plasma of approximately 15 minutes after secretion or intravenous injection(Veldhuis et al. 2002; Psrker et al. 1962).

Figure 1-1, summarises the normal physiology of GH (Holt 2002). After secretion, a substantial proportion of GH circulates bound to high affinity GH-binding protein (GHBP), which is identical to the extra-cellular domain of the GH receptor. GH exerts its action on target cells by binding to the cell surface of the GH receptor (GHR), which is ubiquitously expressed throughout the body with high levels in liver and adipose tissue. Following the binding of one molecule of GH to two receptors, GH stimulates many metabolic processes throughout the body.

There is a large number of GH-dependent markers produced by the influence of GH, with IGF-I (and its binding proteins) being the best known (Le et al. 2001). IGF-I gene

expression is stimulated by GH in many tissues. Although IGF-I is made ubiquitously, the liver is the main source of circulating IGF-I (Le et al. 1998; Sonksen 2001). GH stimulates the production of IGF-I. As hepatic IGF-I production is regulated by other factors other than GH, serum IGF-I may be considered as a marker of GH action on the liver (Butler et al. 2001; Sonksen 2001).

Growth hormone therapy must be administered by injection, as it is digested to its constituent amino acids if administered by mouth (Sonksen 2001). After subcutaneous or intramuscular administration, blood concentrations of GH reach a peak between 1 and 3 h after injection and fall to undetectable levels after 12-24 h (Refetoff et al. 1970).

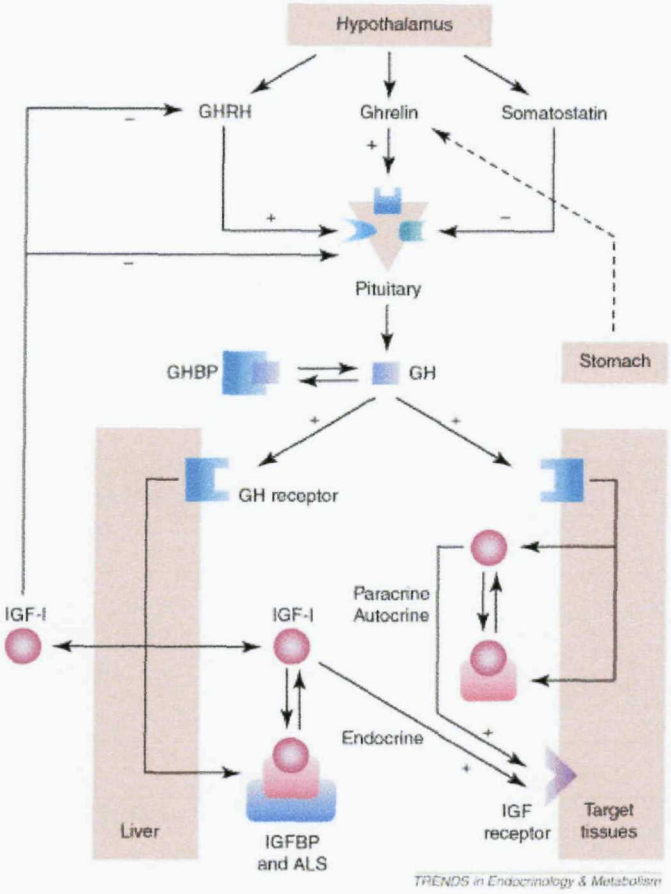

FIGURE 1-1 THE PRODUCTION AND CLEARANCE OF GROWTH HORMONE IN THE BODY SYSTEM AND THE SYNERGY WITH IGF-I. GH PRODUCTION OCCURS AT THE PITUITARY GLAND AND THIS IS CONTROLLED BY HYPOTHALAMIC HORMONES, SOMATOSTATIN AND GHRH, AS WELL AS THE MAINLY GASTRIC GHRELIN. GHRH AND GHRELIN ARE RESPONSIBLE FOR STIMULATING GH SECRETION. SOMATOSTATIN IS RESPONSIBLE FOR DECREASING GH PRODUCTION. GH ACTIONS ARE FACILITATED BY IGF-I WHICH IS PRODUCED IN MANY DIFFERENT TISSUES, BUT PRIMARILY IS PRODUCED IN THE LIVER. IGF-I CIRCULATES BOUND TO IGFBPS AND THESE ARE RESPONSIBLE FOR CARRYING OUT BIOLOGICAL AND ENDOCRINOLOGICAL ACTIONS OF IGF-I. OVER 99% OF CIRCULATING IGF-I IS BOUND IN A TERNARY COMPLEX COMPRISING IGF-I, IGFBP-3 AND AN ALS. IGF-I INHIBITS GHRH AND GH SECRETION IN A NEGATIVE FEEDBACK MECHANISM. COPYRIGHT TRENDS IN ENDOCRINOLOGY AND METABOLISM 2002 (HOLT 2002)

1.6.3.1.2 GROWTH HORMONE RELEASE

The pituitary gland secretes GH in a pulsatile pattern every 3-4 hours and the major stimuli for GH secretion in humans are sleep, exercise and stress (Muller et al. 1999). The vast majority of the daily GH output occurs during sleep particularly during slow wave sleep. Other factors identified as having a modifying effect on hGH concentrations, have been reported to include adiposity, fitness level, stress, gender, age and diet (Harvey 1995). GH secretion is increased by hypoglycaemia, high temperature, stress and deep sleep, and is decreased by obesity, a carbohydrate- rich diet and b-adrenergic agonists.

1.6.3.1.3 PHYSIOLOGICAL INFLUENCES ON THE GH/IGF AXIS

1.6.3.1.3.1GENDER

Young women have approximately 50% higher 24-hour integrated GH concentrations than young men. In addition, in pre-menopausal women, GH production varies according to the phase of the menstrual cycle, approximately doubling in the late follicular compared with the early follicular and mid-luteal phases. This variation is likely to be related to changes in gonadal steroid production as GH pulse amplitude is positively correlated with serum oestradiol and negatively correlated with serum progesterone levels (Faria et al. 1992). After a glucose load, plasma GH levels are normally suppressed to greater extent in young men and young women in the early follicular phase of their menstrual cycle.

1.6.3.1.3.2NUTRITION

Body composition has a significant effect on GH production. Obese individuals have have reduced mean GH concentrations, associated with decreased pulsatile GH secretion and a shorter GH half-life. It has also been shown that the peak in response for GH secretion to GHRH is reduced in obese subjects and this is often reversible with fasting or weight loss. A possible explanation for this is related to hyperinsulinaemia, associated with obesity, may inhibit GH secretion, perhaps via regulation of IGF-I and/or IGFBP concentrations. Ghrelin levels are reduced in obese

patients, but it is unclear whether this mediates any of the reduced GH secretion observed in visceral obesity (Hartman et al. 1992).

Patients with anorexia nervosa by contrast demonstrate an increased basal and GHRH-stimulated GH release and elevated ghrelin levels, reinforcing the importance of nutritional state on GH regulation. Patients with anorexia nervosa have low circulating IGF-I and IGFBP-3 levels as a result of low hepatic insulin activity and can be viewed as developing a state of functional GH resistance.

In healthy individuals, fasting results in increased GH production secondary to an increase in GH pulse frequency and amplitude whereas re-feeding rapidly has an opposite effect causing suppression in fasting-enhanced GH secretion.

Other central and peripherally generated peptides such as leptin and neuropeptide Y, which have important roles in the regulation of food intake and energy balance, are also involved in the control of GH secretion though the relative importance of these factors in GH regulation is not yet apparent.

1.6.3.1.3.3AGE

During puberty, daily GH production rates increase threefold, peaking during late puberty – and associated with maximal linear growth velocities. After puberty, GH production starts to decline at a rate of approximately 14% per decade of adult life (Iranmanesh et al. 1991).

The decline in GH secretion that occurs with increasing age is associated with a reduction in GH storage. By using appropriate stimuli GH pulses generated in elderly subjects were similar to those seen in the younger individuals. Similarly, the frequency of GH pulses was maintained, however the amplitude of these pulses were reduced in comparison with those observed in young adults. This suggests that GHRH activity may be diminished with increasing age. This was supported by observation that GHRH administration in elderly individuals increases spontaneous GH secretion and serum IGF-I, to levels seen in young adults.

In parallel with the fall in GH secretion with advancing age, serum levels of IGF-I and IGFBP-3 also decline (Janssen et al. 1998), though the concentration of the latter falls to a lesser degree than IGF-I and GH secretion (Donahue et al. 1990). Studies have however failed to demonstrate a correlation between GH secretion and IGF-I and IGFBP-3 in elderly individuals, suggesting that GH becomes less important in determining circulating IGF-I concentration with advancing age. Under these circumstances, other factors such as nutritional state, fragility and chronic disease may play a greater role in determining IGF-I and IGFBP-3 concentrations.

1.6.3.1.3.4 EXERCISE

Physical exercise plays an important role in the regulation of the GH–IGF-I axis (Kanaley et al. 1997) with exercise increasing GH secretion. The GH response to exercise, is dependent on the duration and intensity of the exercise session, the fitness of the exercising subject and other environmental factors (Eliakim et al. 1998).

Weltman et al showed that the effect of exercise is greater in young women than postmenopausal women and men of any age. Furthermore, it was shown that age diminishes the GH response to exercise and abolishes the young-adult gender difference (Weltman et al. 1994).

The duration of exercise needed to induce GH secretion after physical exercise is inversely related to the intensity of this exercise (Rigamonti et al. 2005). Maximal GH levels in plasma are present 15–30 min after the start of intense exercise, and then decline to half values in the following 16–20 min (Lassarre et al. 1974). Reciprocally, an important role has been attributed to GH in the regulation of the energy needed to accomplish a muscular performance in humans (Galbo 1982).

Approximately tenfold increases have been observed in GH plasma concentrations during prolonged, moderate exercise and more intensive short-term exercise [accumulation of lactate; VO2max (maximum oxygen consumption) above 70%] (Felsing et al. 1992).

Anaerobic exercise is more stimulatory to GH release than continuous aerobic exercise, even if duration and workload is kept the same for the two types of exercise (VanHelder et al. 1984).

Although exercise is a powerful stimulus to GH release, the GH response to exercise is blunted in older and obese individuals. This suggests that higher relative exercise intensities may be necessary than exercise alone to stimulate adequate GH release in obese subjects.

1.6.3.1.4 GROWTH HORMONE CLEARANCE

Circulating GH is cleared from the bloodstream through degradation, predominantly in the liver and kidney. In the kidney, GH is cleared principally by glomerular filtration, uptake and degradation in the proximal tubule. Only minute quantities of GH appear in the urine and urinary excretion of GH is not a constant proportion of plasma concentrations, the excretion dynamic is influenced by many other factors (Krogsgaard et al. 1994; Sonksen 2001).

1.6.3.2 ACTIONS OF GROWTH HORMONE

The following section discusses in some detail the effect Growth Hormone has on fat metabolism, muscle protein synthesis and on the bone connective tissue. This is an important section, because it provides a deeper understanding into the rationale behind GH abuse in sport. Athletes believe that the effects of naturally occurring GH will be enhanced further with exogenous GH administration. A number of research studies are discussed in this section, which demonstrate the evidence available to date on the effect of GH.

The majority of these studies were undertaken either using GH deficient individuals or health individuals. It is reasonable to assume that the effect of GH administration in GH deficient individuals will be similar to that found in healthy individuals.

1.6.3.2.1 EFFECT OF GROWTH HORMONE ON FAT METABOLISM

An important property of GH is that it exerts potent lipolytic effects (Kraemer W. et al. 2002; Sonksen 2001). There are three potential mechanisms suggested for this; the first is by a direct effect (Sonksen 2001), the second by an increase in other lipolytic hormones such as catecholamines and glucagon and the third by influencing the fat cells responsiveness to other hormones (Kraemer W. et al. 2002). GH can directly decrease the amount of lipids produced, by reducing the number of lipogenic enzymes produced (Sonksen 2001). Indirectly, it has been suggested that GH increases the number of adrenergic receptors and/or inhibitory factors, which would lead to an increased fat utilisation. It was shown in a GH administration study that GH increases body fat utilisation significantly (Crist et al. 1988).

1.6.3.2.2 EFFECT OF GROWTH HORMONE ON MUSCLE PROTEIN SYNTHESIS

One of the most well known effects of growth hormone is its positive effect on muscle protein synthesis. Perhaps this is the main reason why GH users believe that administration of GH will lead to an increased performance. However, despite the fact that in theory GH should be beneficial to sporting performance, the condition of acromegaly (natural excess GH) and negative clinical studies as evidence of a lack of benefit.

Many studies have concluded that GH administration to GH deficient adults and children have a positive effect in limb and skeletal muscle and muscle strength (Cuneo et al. 1991; Jorgensen et al. 1989; Tanner et al. 1977).

GH induces glucose and amino acid uptake, and in general it is thought to stimulate protein synthesis, possibly using the energy derived from its lipolytic activity, by usage of the nutritional calories to initiate protein production (Sonksen 2001). Growth hormone enhances amino acid uptake and transport, thus increasing the capacity for protein synthesis. Growth hormone optimises the efficiency and utilisation of nitrogen by increasing protein synthesis and decreasing protein

degradation. Action is mainly on synthesis and not degradation. Insulin prevents degradation and this is why many athletes abuse both insulin and GH. Mechanistically many of the GH effects on protein are thought to occur indirectly via IGF-I; however direct effects are possible (Kraemer W. et al. 2002).

Reduced muscle mass and slower muscle protein synthesis is thought to be associated with reduced growth hormone and IGF-I (Fryburg et al. 1991; Fryburg et al. 1993; Lange et al. 2001; Welle et al. 1996; Welle 1998).

Despite the few studies that suggest GH administration will increase muscle protein synthesis and strength, there are an important number of studies that suggest that GH administration will not increase muscle strength significantly (Yarasheski et al. 1992; Yarasheski et al. 1993; Yarasheski et al. 1995; Deyssig et al. 1993). The common feature of these studies is they are all short-term GH administration studies, which might be the reason why there are no findings to suggest a positive effect of GH on muscle protein synthesis and strength.

Nevertheless, a critical appraisal of studies investigating the hypothesis that GH improved performance in humans published in 2002 has also suggested that at present there is no significant evidence to suggest such an argument (Dean 2002).

To date, there have been no published studies investigating the effects of GH administration on the performance of athletes (Rigamonti et al. 2005; Kraemer W. et al. 2002) and any claims on the properties of GH have been pragmatic and based on hearsay rather than scientific evidence. Anecdotal evidence suggests that athletes will use higher dose of growth hormone and for a longer period of time than it would be ethically acceptable to mimic in a research study, so this adds to the difficulties of carrying out research trials that would support such claims (Dubin C.J. 1990; Sonksen 2001).

1.6.3.2.2.1RESULTS FROM STUDIES INVOLVING GHD SUBJECTS

A large number of articles published in academic peer review journals, have shown that there is significant evidence to suggest that GH administration significantly

improves functional capacity and quality of life. The results of these studies have been discussed in recent reviews (Gibney et al. 2007; Woodhouse et al. 2006).

The key findings in these studies suggest that maximal oxygen consumption (VO_{2max}; aerobic capacity or the maximum ability to take in and use oxygen) in GHD adults has been consistently shown to be reduced by estimates ranging from 17 to 27% compared with values predicted for age, gender, and height (Cuneo et al. 1991).

Using supraphysiological amounts of GH in trials of GH replacement, it was shown that GHD adults showed an improvement in exercise performance (Cuneo et al. 1991). The trials undertaken by Cuneo et al were some of the first trials to be undertaken in this area and were successful in showing marked increases (and normalization compared with predicted values) in VO_{2max}, maximal power output, and the ventilatory threshold (VeT; lactate threshold) after 6 months of GH replacement in GHD subjects (Cuneo et al. 1991).

A further study that suggests that a longer GH administration period might of benefit in order to observe a significant effect of GH on performance was carried out by Bollerslev in 1996 (Bollerslev et al. 1996). In this placebo-controlled, crossover study involving of 55 adults with adult-onset GH deficiency the effects of GH therapy on exercise capacity and body composition were investigated. GH and placebo were administered for 9 months each, separated by a 4-month washout period. Body fat mass was significantly reduced and lean body mass was increased with GH treatment. Total and low-density cholesterol levels decreased. Absolute maximal oxygen uptake increased by 6%, relative to body weight by 9%, and there was a trend toward increased endurance performance by 7%. The authors concluded that treatment with a low, physiologic dose of GH produced positive effects on body composition and lipids and improved exercise capacity, likely to be of clinical relevance.

The vast majority of studies reported to date have demonstrated increased maximum work rate and VO_{2max} (Cuneo et al. 1991; Jorgensen et al. 1989; Jorgensen et al.

1994b; Jorgensen et al. 1994a) after GH replacement in subjects with both childhood- and adult-onset (AO)-GHD.

Although the majority of the evidence suggest that GH administration has a positive impact in GH deficient patients, in some cases, these trials did not manage to show this result on a statistically significant level. However, this can possibly be explained by the small sample sizes used in these trials. An example of this is the study by (Janssen et al. 1999). This study investigated the changes in muscle volume, strength and bioenergetics during a placebo controlled, study with administration of GH in normal physiological dose involving 28 males with GH deficiency. The objective of this study was based on the evidence that adults with GH deficiency, have been reported to suffer from muscle weakness, due to a reduction in muscle mass. It was examined whether GH therapy would improve significantly the muscle strength and mass. It was concluded that although following GH therapy, muscle mass had increased significantly in the GH treated group, there was no significant evidence to suggest that muscle strength had changed.

The largest study to date addressing exercise performance in response to GH replacement included 55 patients with AO-GHD in a placebo-controlled, crossover study in which GH therapy was individually dosed to obtain an IGF-I concentration within the normal range for age and sex (Bollerslev et al. 2005). A highly significant effect of GH replacement to increase VO_{2max} by approximately 6% was observed.

The overall body of evidence therefore supports an effect of GH to improve maximum work rate, VO_{2max}, and VeT, with changes in VO_{2max} apparently accounted for by increased LBM (Gibney et al. 2007).

1.6.3.2.2.2RESULTS FROM STUDIES INVOLVING HEALTHY SUBJECTS

In one trial (Deyssig et al. 1993) 22 young athletes were followed receiving GH or placebo for six weeks. It was concluded that despite the increase in GH, IGF-I and IGFBP-3, their body weight and fat mass were not changed significantly. Moreover, strength increased significantly in both groups (GH and placebo). The authors

concluded that the anabolic and lipolytic actions of GH therapy in adults depend on fat mass and the presence of GH deficiency. Furthermore, in highly trained athletes who already have reduced fat mass, GH administration had no effect on body composition or on muscle strength.

Fryburg in 1993 suggested following an 8 hour infusion study, with a sample size n=8, that although GH administration acutely stimulates skeletal muscle, there is no evidence to suggest that whole-body protein synthesis is also stimulated (Fryburg et al. 1993). A later study by Healy et al however, showed that using a longer GH administration period with daily GH injections for a period of 4 weeks examining endurance-trained male athletes (n=11), indeed GH administration showed a net anabolic effect on whole body protein metabolism at rest and during and after exercise (Healy et al. 2003).

A recent article (Berggren et al. 2005), published in 2005, studied a group of young men and women with normal GH insulin-like growth factor I axes with the objective to study the effects of GH-IGF axes on exercise capacity. A double blind placebo controlled study design was used, over a period of 28 days. Subjects randomised in the GH treatment group, received either 0.033mg/kg or 0.1mg/kg every day of the treatment period. Following GH administration, there were no significant differences observed in the power output and oxygen uptake measured between the two treated groups. However, subjects in the GH treatment group had a significantly increased total body weight. It was therefore concluded that administration of supraphysiological hGH during a 4 week period, did not improve power output or oxygen uptake.

Despite these negative studies, several recent studies have shown a performance benefit in healthy adults.

In a recent 6 month GH administration study (Giannoulis et al. 2006), in a population of healthy elderly men GH produced a significant positive effect on performance. In this study, Giannoulis et al (Giannoulis et al. 2006) carried out a 6-month

randomised, double blind placebo-controlled trial involving 80 healthy older men (age, 65-80 years old) in order to assess the effects of near-physiological GH with/without testosterone administration on lean body mass, total body fat, mid-thigh muscle cross-section area, muscle strength, aerobic capacity, condition-specific quality of life, and generic health status of older men. It was found that while co-administration of GH with testosterone produced the most significant changes on the investigated variables the administration of GH or Testosterone alone still produced significant results compared to the placebo treatment.

A recent, a well designed and executed study involving volunteers who were had been in the past abusers of anabolic steroids demonstrated for the first time an ergogenic effect of GH in healthy young athletes (Graham M.R. et al. 2008).

The reason it has been difficult to undertake GH administration studies, which will demonstrate accurately whether GH has the ability to increase muscle bulk and strength simultaneously is because in reality athletes will abuse GH in combination with anabolic steroids and other performance enhancing substances, in varying concentrations during differing training and dietary regimens. This makes it impossible to replicate in a single research trial, and therefore it seems likely that the athletes is probably best placed to address the question whether GH is performance enhancing. This certainly appeared to be a potent weapon when abused by the former East German coaches (Franke et al. 1997). Furthermore, the combinations and amounts of performance enhancing substances anecdotally known to be abused by athletes will never be medically safe and would be immoral and unsafe to replicate for the purposes of medical research (Holt et al. 2008).

1.6.3.2.3 EFFECTS OF GROWTH HORMONE ON BONE AND CONNECTIVE TISSUE

GH plays an important role as an anabolic agent in connective tissue in human skeletal muscle and tendon (Rennie 2003; Doessing et al. 2005). It is currently thought that a major role of connective tissue in muscle and tendon is to provide

means of providing a matrix for transmission of force from individual muscle fibres to the bone. Therefore, a strengthened connective tissue would give a stronger and more strain-resistant muscle and tendon. This could fit with the claimed effect of rhGH on athletic performance. Furthermore, an anabolic effect of rhGH in connective tissue could also suggest a potential for rhGH in treatment of muscle and tendon injuries, a common problem in sports, hence another reason for GH doping.

1.6.3.3 ACROMEGALY & GH DEFICIENCY

1.6.3.3.1 ACROMEGALY

Acromegaly is a disorder characterized by autonomous overproduction of GH caused by a GH-producing pituitary adenoma. Acromegaly is a rare condition, with an estimated incidence of 4 per million per year (Dekkers et al. 2008).

The first description of acromegaly was in 1864, as recorded by Verga (Colao et al. 2004). This case concerned a woman who had suffered typical somatic disfigurement, arrhythmias and osteoarthropathy. A post-mortem examination showed that the woman had a giant pituitary. The skull for this patient is currently kept at the Anatomical Museum of Modena in Italy. This case was incorrectly classified by Verga as early menopause.

Some years later, in 1881, following the autopsy of the Italian actor Ghirlenzoni, Brigidi reported a description clinically consistent with acromegaly, as he was identified with visceromegaly and enlarged hypertrophic pituitary. This case was also incorrectly interpreted as to primary bone disease (Colao et al. 2004).

A number of similar cases followed these two cases, however in none of the cases the pathogenesis of the condition was fully understood, and it was only in 1909 when Harvey Cushing identified the etiology of the diseases through a partial hypophysectomy (the removal of the pituitary gland), which also enabled him to suggest the potential treatment.

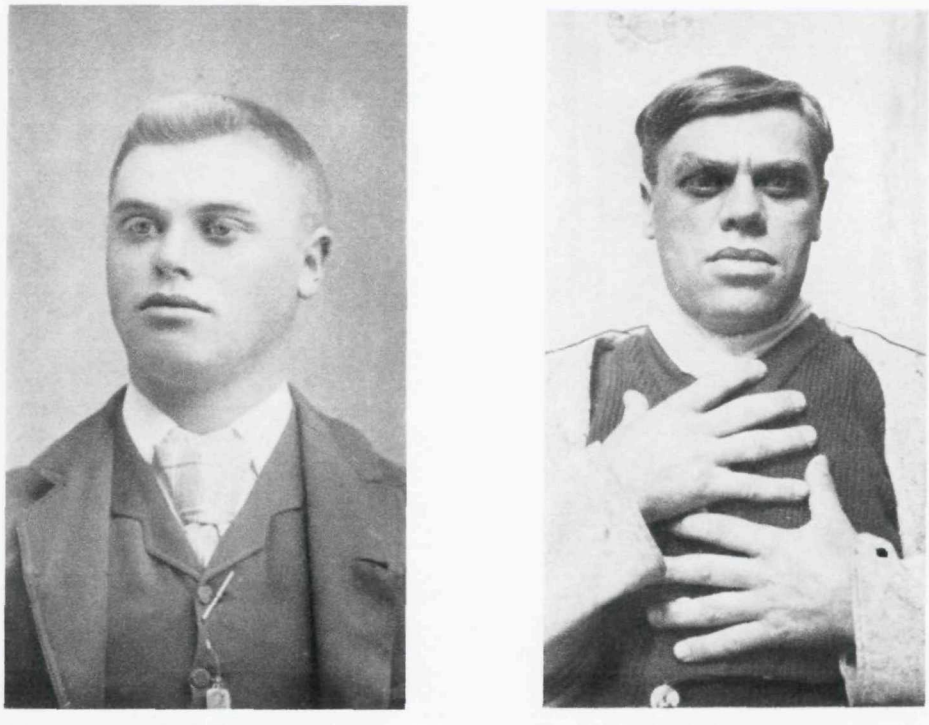

FIGURE 1-2 ON THE LEFT, THE FIRST PATIENT TO BE DIAGNOSED WITH ACROMEGALY BY H. CUSHING, AS HE WAS BEFORE ONSET OF SYMPTOMS. ON THE RIGHT, THERE IS A PHOTOGRAPH OF THE PATIENT AS HE WAS ON DATE OF ADMISSION. (CUSHING 1909)

At diagnosis, patients generally exhibit coarsened facial features, exaggerated growth of hands and feet, and soft tissue hypertrophy. Naturally occurring excess levels of GH in the body are associated with pathological conditions such as hypertension, myocardial hypertrophy, diastolic dysfunction, insulin resistance, sleep apnea, and ventilatory dysfunction. In addition to these disease-specific conditions, mass effects of the pituitary tumour can lead to hypopituitarism and visual field defects.

GH/IGF-I excess affects the body system in five main ways (Colao et al. 2004):

- **Soft tissue and skin changes:** Enlargement of the body's extremities, increased skin thickness and soft tissue hyperplasia, increased sweating, skin tags and acanothosis nigricans (a cosmetic condition strongly associated with insulin resistance in which there is darkening

of the skin in areas where there are creases such as the neck and arm pits)

- **Affecting the cardiovascular system:** Biventricular hypertrophy, increased interventricular septum thickness (eccentric hypertrophy), diastolic dysfunction at rest and/or systolic dysfunction on effort, diastolic heart failure, arrhythmias, hypertension and/or endothelial dysfunction and increased carotid IMT.
- **Metabolic features:** Impaired fasting glucose, impaired glucose tolerance, diabetes mellitus, insulin resistance, reduced total cholesterol and increased triglycerides and increased nitrogen retention.
- **Respiratory features:** Upper airway obstructrion, macroglossia, sleep apnea, and ventilator dysfunction. Increased articular cartilage thickness, upper airway obstruction, macroglossia, sleep apnea and ventilator dysfunction
- Other endocrine consequences include thyrotoxicosis, hypercalciuria and hyperparathyroidism

Acromegaly is associated with considerable morbidity (Nabarro 1987) and increased mortality (Colao et al. 2004). However, when adequately treated, the relative mortality risk can be markedly reduced toward that of the normal population. Treatment for acromegaly involves normalising GH secretion. However, because GH is released in a pulsatile fashion, normal and abnormal ranges are not easily defined. Furthermore, in certain cases, it is impossible to remove completely the underlying pituitary adenoma by surgery and often further treatment is needed after surgery to reduce GH levels.

To diagnose and monitor acromegaly in patients, concentrations of GH and IGF-I are the preferred markers used. As discussed in other sections of this thesis, the use of GH concentrations is challenging because a single value is not be reliable, it is a poor reflector on clinical status and results of dynamic tests are often in disagreement with IGF-I results.

In contrast, IGF-I has a long half-life, allowing a single value to be used for acromegaly diagnosis. This is a useful marker to examine when diagnosing and assessing the state of acromegaly for a particular patient, because IGF-I is involved in all stages of this disease, and IGF-I levels are elevated even in patients with the mildest activity whose plasma GH levels are within the defined normal range.

Plasma GH levels are useful in assessing partial biochemical responses to therapy when IGF-I levels may still be maximally increased, and its dynamic responses may provide information about persistent dysregulation of the somatotrophic axis. However, a persistently high plasma IGF-I level is a clear indication to consider therapy; whereas, abnormal GH responses per se are not (Barkan 2004).

1.6.3.3.2 GH DEFICIENCY

1.6.3.3.2.1SYMPTOMS RELATED TO GHD

GH deficiency may be both, congenital or acquired. In children, GHD is easier to diagnose, as this largely depends on diagnostic features such as a short stature, slow linear growth rate, mildly delayed bone age with reduced bone density, excess adiposity, reduced lean tissue mass and fasting hypoglycaemia (Jorgensen et al. 1994a; Jorgensen et al. 1994b)

However, in individuals with adult onset GHD, diagnosis is much more difficult because these patients have normal stature, and the insidious clinical presentation fails to raise suspicions easily for GHD syndrome. This is also the main reason adult GH deficiency remained undiscovered for many years (late 1980s) after child GHD syndrome was defined (1960s) (Juul et al. 1994; de et al. 1995; Woodhouse et al. 2006).

Adult patients with abnormal levels of GH have a number of somatic impairments, including altered blood biochemistry, metabolism, body composition, muscular and aerobic performance. If this condition is left untreated for prolonged periods, patients have been reported to have reduced levels of energy, decreased muscle strength, and early exhaustion. These symptoms have been shown to result in a worse health related quality of life (HR-QoL) with individuals experiencing functional limitations,

depression, excessive fatigue and diminished productivity (Cuneo et al. 1990; Cuneo et al. 1992; Gibney et al. 2007).

1.6.3.3.2.2GH TREATMENT FOR PATIENTS WITH GHD

Adults with GHD have reduced lean body mass and skeletal muscle mass compared with healthy control subjects, suggesting an underlying abnormality in protein metabolism. The availability of recombinant human GH has resulted in the widespread interest in the use of these treatments in adults with GH deficiency (GHD) (Cuneo et al. 1990; Cuneo et al. 1992; Woodhouse et al. 2006).

A number of research projects have been undertaken in order to determine whether GH replacement increases protein synthesis and consequently resulting in increased muscle mass, or whether GH simply increases water, providing a false impression of enhanced lean body mass. To date most evidence suggests that acute or short-term administration of GH increases protein synthesis, whereas administration that is more chronic appears to result in reduced proteolysis.

Several trials have been undertaken involving elderly adults, patients with HIV wasting or GHD, or GH excess have demonstrated the beneficial effects of normalizing GH status on blood biochemistry and body composition. However, the impact of GH physical performance of functional activities and quality of life is still not known.

Protein molecules contain all nitrogen contained in the human body, therefore the amount of nitrogen can be used to estimate the amount of protein in the body, where an increase in total body nitrogen (TBN) reflects an increase in protein content. Using this as a measure, a number of studies have consistently demonstrated that GH administration enhances nitrogen retention, which is in turn associated with an increase in lean body mass (Woodhouse et al. 2006).

There is general agreement that administration of GH to adults with GHD reduces impairment by increasing circulating levels of its target IGF-I, skin thickness, bone mineral content, and lean body mass, and reducing fat mass. It is not certain whether

an improvement on these impairments translates into clinically meaningful improvements in physical function and health related quality of life (Woodhouse et al. 2006).

Based on reduced absolute lean body and muscle and the tissue changes seen with biopsy data from adults with GHD, a reduction in absolute maximum isometric force generating capacity, isokinetic torque production (particularly at slower velocities), and improved local muscular endurance would be expected (Woodhouse et al. 2006).

A further discussion on the effect GH administration has on GH deficient patients has been given in an earlier section of this thesis (see section 1.6.3.2.2.1). In summary, results to date, have shown that GH treatment for GH deficient individuals has a positive impact on muscle performance.

1.6.3.4 THERAPEUTIC USE OF GROWTH HORMONE

The first peer-reviewed article examining the therapeutic uses for growth hormone was published in 1953 (Ducommum et al. 1953) and investigated the effect of pituitary growth hormone in the prevention of infections caused by a cortisone overdose. Since then, there have been many other proposed therapeutic uses of exogenously administered growth hormone. One of the oldest (Raben 1958) and most widely known uses of growth hormone is to treat children with growth problems. Many research studies have been carried out, identifying growth hormone as beneficial to children with growth problems.

In the UK, growth hormone is currently approved by the National Institute for Health and Clinical Excellence (NICE) for a number of conditions. Similar to the UK, the US Food and Drug Administration body has also approved the use of GH in children for the same conditions. Its application in therapy it is not limited to GH deficient children, but it is now used in other clinical conditions not necessarily related to short stature. The main conditions for which growth hormone administration is used for includes growth hormone deficiency. Results from published clinical trials

(randomised and cohort studies) show that GH therapy is effective in promoting growth in children with GH deficiency.

Another condition for which GH treatment is used for is the Turner syndrome. Turner syndrome is a neurogenetic disorder characterized by partial or complete monosomy-X. It is associated with certain physical and medical features, including oestrogen deficiency, short stature, and increased risk for several diseases, with cardiac conditions being among the most serious.

Furthermore, growth hormone treatment is used to treat growth failure associated with chronic renal failure. Renal failure involves the damage of the kidneys (which can be either temporary or permanent), and this subsequently results in loss of normal kidney function. There are two different types of renal failure - acute and chronic. Acute renal failure has an abrupt onset and is potentially reversible. Chronic failure progresses slowly over at least three months and can lead to permanent renal failure. The causes, symptoms, treatments and outcomes of the two types of renal failure are different.

1.6.4 GROWTH HORMONE ABUSE

1.6.4.1 REASONING FOR GROWTH HORMONE ABUSE

Apart from the clinically accepted uses of exogenous GH, many have also experimented with this hormone, in an attempt to take advantage of its known anabolic properties in sport. Although the performance enhancing properties of exogenously administered growth hormone have only recently been confirmed, many abuse this as part of their training regimes.

Anecdotal evidence suggests that a month's dose of growth hormone will cost more than $2500. However, as there is no evidence so far to support any of the claimed benefits of growth hormone, it has been suggested that growth hormone is just "a fool's gold" (Vogel 2004).

Anabolic-androgen steroid (AAS) hormones were initially the main drugs used as 'doping agents', based on their alleged capacity to increase muscle size and strength (Evans 2004). At present, steroids are still in use by many athletes however it is believed that their use is not as common as it used to be, mainly because other potent, endogenous, anabolic compounds have been made available and gaining prominence in the doping world (Rigamonti et al. 2005) . Furthermore, methods for detecting abuse of AAS are now sophisticated. Among these newer substances, the pituitary derived growth hormone, which has during the past two decades been produced by recombinant DNA technology (rhGH), has gained popularity in the doping world.

Many reasons, besides its undeniable anabolic activity, account for the inappropriate use of GH in sport, including the present availability of unlimited amounts of rhGH and the lack of a reliable International Olympic Committee-accredited method for its detection in body fluids (Rigamonti et al. 2005). The fact that there is currently no official method of detecting GH abuse and a general agreement in athletic communities that GH possesses powerful ergogenic (i.e. performance enhancing) effects, is probably the reason why GH is the "drug of choice" for many athletes.

It is not accurately known how much growth hormone is abused by athletes on a daily basis, as part of their training regimes. It has been suggested that hGH abusers may take up to 0.2 IU/kg/day (1mg = 3 units approximately) which might give different results (Kniess et al. 2003), and might do so during cycles lasting 6-12 weeks; every other day (Sonksen 2001; Kraemer W. et al. 2002) although there are athletes abusing it for even longer periods. Others suggest that this might be between 2 mg three times per week and 5 mg per day for many weeks (Di Pasquale 1990), with this last dose being about tenfold higher than the daily secretion rate of GH in the adult (MacGillivray et al. 1970). Furthermore, it has been reported that GH when used by athletes, will often be taken in combination of other performance enhancing substances such as synthetic testosterone (Kraemer W. et al. 2002) or in sequence of GH treatment followed by anabolic-androgen steroids.

Thus, although there have been several research studies to examine the effects of GH on muscle strength and performance, as stated by Kraemer et al the "real world" presents a complex picture of different combinations of anabolic drugs that anecdotally appear to be more effective than single drug-use". This is perhaps a further reason why athletes remain unconvinced of several scientific claims that GH does not enhance their athletic performance, and are continuing its abuse, despite its potentially fatal consequences, and their disqualification of competitions.

1.6.4.2 RISKS OF GROWTH HORMONE ADMINISTRATION

The actual risks of GH use by athletes are not well appreciated because relevant epidemiological data are lacking. Both hGH and IGF-I are potent mitogenic and anti-apoptotic agents, and several reports have shown an association between IGF-I levels and the incidence of breast, prostate, and colorectal cancers (Allen et al. 2005; Nabarro 1987; Yarasheski 1994).

In medical practice, GH excess can cause insulin resistance, glucose intolerance, peripheral oedema and arthralgias. These adverse effects are also commonly seen in acromegaly described above.

Furthermore, growth hormone abuse with concomitant use of anabolic steroids has been identified as possible aetiological factors in the development of bilateral internal laryngocoeles. This was following an incident involving an elite bodybuilder who admitted to GH abuse in the past prior to his diagnosis as reported by Moor and Khan in 2005 (Moor et al. 2005).

In addition, a report suggested that contrary to earlier studies, GH-induced carpal tunnel syndrome is not an age-related phenomenon and alerted physicians to include GH abuse as a possible aetiology of median neuropathy in athletes (Dickerman et al. 2000; Longobardi et al. 2000). This was following an incident where a male elite bodybuilder suffered median nerve neuropathy during a self-administered course of growth hormone.

The chronic abuse of supraphysiological doses of GH by athletes is likely to be associated with significant side effects, as illustrated by studies of acromegalic patients. A detailed recent review of this is given in Buzzini (Buzzini 2007). A list of these potential side effects is also given here. As there is anecdotal evidence to suggest that GH doses administered are significantly higher than the doses used for GH deficiency treatment by adults of children, it is likely that the side effects exhibited in athletes could be of much greater magnitude than the ones studied so far.

Suggested potential side effects of growth hormone misuse in athletes as reviewed by Buzzini are summarised below (Buzzini 2007):

- Cardiovascular: Hypertension, cardiomyopathy, congestive heart failure, arrhythmia
- Pulmonary: Respiratory failure, sleep apnoea, narcolepsy, sleep disturbances
- Musculoskeletal: Osteoarthritis, carpal tunnel syndrome, slipped capital femoral epiphysis, worsening of existing scoliosis, osteoporosis, arthralgias, muscle weakness, myopathy, avascular necrosis of the femoral head, gigantism
- Endocrine and metabolic: Diabetes mellitus, insulin resistance, glucose intolerance, dyslipidemia, peripheral edema, hypothyroidism, menstrual irregularity, erectile dysfunction, multiple endocrine neoplasia type 1 (for those who have a genetic mutation)
- Neurologic: Idiopathic intracranial hypertension (pseudotumour cerebri), visual field defects, cranial nerve palsy, headache
- Malignancy: Increased risk of leukemia, solid tumours (breast, colon, prostate, and endometrial cancer), and increased risk of second neoplasms Infection HIV/AIDS, hepatitis B and C (nonsterile or contaminated syringes), Creutzfeldt-Jakob disease (GH extracted from human pituitary glands obtained on the "black market")

- Cosmetic: Prognathism and jaw malocclusion, coarsened facial appearance, increased skull circumference, dentition problems, acral overgrowth, frontal bone bossing, gynecomastia
- Visceromegaly: Tongue, thyroid gland, salivary glands, liver, spleen, kidney, prostate
- Skin and gastrointestinal: Hyperhydrosis, oily texture, skin tags, colon polyps
- Injection injuries: Direct and indirect trauma to nerves and soft tissue, abscess
- Other: Counterfeit products might carry serious risks as they do not have to adhere to safety regulations and might have a compromised quality of life.

1.6.5 GROWTH HORMONE ABUSE DETECTION METHODS

1.6.5.1 WADA REGULATIONS FOR GROWTH HORMONE TESTING

The World Anti Doping Agency (WADA) has a number of international standards in place in order to ensure effective testing in sport. The purposes of adopting international standards in sport testing as stated by WADA are (The World Anti-Doping Agency 2006):

- Effective testing,
- Maintaining the integrity and identity of samples, from notifying the athlete to transporting samples for analysis.
- Ensuring production of valid test results and evidentiary data
- Achieving uniform, harmonized results and reporting from all accredited laboratories.

This section will describe what the procedures in sport testing are, focusing on the area of peptide hormones, as Growth Hormone is classified in this section. It should

be noted that these testing procedures will be followed when the sporting event is governed by the WADA/IOC regulations.

The WADA code currently states concerning growth hormone doping that an athlete will be found guilty of doping with growth hormone under these conditions:

"Unless the Athlete can demonstrate that the concentration was due to a physiological or pathological condition, a Sample will be deemed to contain a Prohibited Substance (growth hormone) where the concentration of the Prohibited Substance or its metabolites and/or relevant ratios or markers in the Athlete's Sample so exceeds the range of values normally found in humans that it is unlikely to be consistent with normal endogenous production.

If a laboratory reports, using a reliable analytical method, that a Prohibited Substance is of exogenous origin, the Sample will be deemed to contain a Prohibited Substance and shall be reported as an Adverse Analytical Finding.

The presence of other substances with a similar chemical structure or similar biological effect(s), diagnostic marker(s) or releasing factors of a hormone listed above or of any other finding which indicate(s) that the substance detected is of exogenous origin, will be deemed to reflect the use of Prohibited Substance and shall be reported as an Adverse Analytical Finding."(The World Anti-Doping Agency 2006)

The WADA code makes no provisions to whether the specimen sample will be blood or urine.

1.6.5.2 TESTING PROTOCOLS

Samples collected for testing in accordance to the WADA regulations can be either urine and/or blood (World Anti-Doping Agency 2004b; World Anti-Doping Agency 2004a). In the case where a urine sample is collected, the athlete is chaperoned at all times while they provide this sample in a collection vessel chosen by the athlete at random from a selection offered by the doping control officers. Following collection of

the urine sample, the athlete is asked to choose a sample collection kit with two bottles; A and B, in which the athlete will pour the urine contents into the two bottles.

In the case where a blood sample is collected, similarly to the urine sample collection, the athlete will choose from a selection of kits. The kit includes two tubes (A and B) in which the doping control officer will collect the necessary amount of blood required for analysis in a laboratory.

Upon collection, urine samples will be transferred to the relevant WADA accredited laboratory where they will be analysed and then frozen under appropriate conditions. For blood samples, where the sample consists of whole blood or blood fractions containing intact cells, they shall be stored at approximately 4 degree Celsius upon reception and should be analysed within 48 hours.

The purpose of collecting a urine or blood into both tubes A and B is to provide a screening and a confirmatory sample. In the case of blood testing, screening and confirmation tests can be performed initially on the same aliquot of sample. The test is also repeated on a fresh aliquot of the same sample tube, to ensure that the initial test results are repeatable from the same sample bottle. All screening assays shall include negative and positive controls in addition to the sample being tested. Following an adverse finding during screening test sample, then the confirmatory test is undertaken within 30 days.

In the instance where an adverse analytical finding is concluded during the screening testing of Sample A, in which the analytes were found to exceed a threshold, then a confirmatory procedure is required to be undertaken using a new aliquot from sample A. In the case where an immunoassay is used for the screening testing process, then the immunoassay to be used for confirmation is required to use a different antibody that will recognise a different epitope of the peptide/protein than the assay used for screening (World Anti-Doping Agency 2004a; World Anti-Doping Agency 2004b).

In the case where all tests from sample A have shown an adverse analytical finding, then a confirmation test from sample B is undertaken, at the same laboratory where

sample A was analysed using a different analyst. The B Sample result must confirm the A Sample conclusion for the adverse analytical finding to be valid. The mean value for the B sample finding for threshold substances is required to exceed that threshold including a consideration of uncertainty.

The ideal test method, as defined in (Rigamonti et al. 2005) should have a "high degree of analytical reliability", it should be capable of identifying the substance misused "with a long time window opportunity for detection".

1.6.5.3 PROBLEMS WITH DETECTING GROWTH HORMONE ABUSE

Even though many efforts have been made to develop suitable assay methodology to distinguish between endogenous and exogenous GH, it has proven difficult to develop a reliable and robust method using such measurements (Bidlingmaier et al. 2003). Detecting GH abuse through the direct measurement of GH directly is difficult because of the pulsatile release and its sensitivity to a number of external influences. GH concentrations in a normal healthy individual can vary significantly during the day, and are greatly influenced by sleep and exercise. Little and inconsistent amounts of GH are excreted in the urine, making urinary GH testing an unviable method. This section discusses in more detail, the issues and challenges pertaining the development of a reliable GH detection method.

The physiological properties of GH have made it difficult to develop reliable methodology to detect its abuse in sports. The primary reason for the lack of an accepted test method is the amino acid sequence identity between the main fraction of pituitary derived hGH and recombinant hGH, which makes it difficult to discriminate between endogenous and exogenous hGH. The structure of rhGH is indistinguishable from endogenous GH; rhGH is identical to the 22 kDa major isoform secreted by the pituitary gland (Jenkins 2001).

In addition, hGH is known to have a short half-life time in circulation of around 15 min (Bidlingmaier et al. 2003). Furthermore, hGH is a natural substance secreted in a

pulsatile manner and thus elevated random measurements may merely reflect a spontaneous peak (Veldhuis 1998). Therefore, single baseline or random measurements of circulating hGH concentrations will be of little diagnostic significance. Levels higher than the reference range are transiently present in healthy subjects of either sex during any 24 h period (De Palo et al. 2001).

Consequently baseline or random hGH measurements with values above the normal reference range, should not be considered as proof of hGH doping (Healy et al. 1997). This was demonstrated using blood samples collected from 500 elite athletes, candidates for the Sydney Olympics, as part of the campaign "Io non rischio la salute" (I'm not risking my health) conducted by the Italian Olympic Committee (CONI) in 1999. There were a large number of blood samples, including one from a gold medallist that were found to have increased GH values above the normal reference ranges. Following assessment of these findings using the serum concentrations of biochemical markers of the GH function it was concluded that there was no evidence to suggest hGH abuse (Armanini et al. 2002).

Developing a detection method for GH abuse using urinary GH measurements has proven difficult, despite many efforts. This is because the amount of GH in the urine does not accurately reflect the blood concentration and can rise dramatically with exercise due to increased excretion of albumin and other proteins that enhance the amount of GH that appears in the urine. This therefore will adversely affect the result of urinary GH doping test, as a high urinary GH value would not necessarily be indicating GH doping.

In particular, Leger in 1995 concluded that there is too wide a variation in the urinary excretion of hGH, as assessed in normally growing and growth hormone-deficient children, thus limiting the clinical usefulness of this hGH quantification method (Leger et al. 1995). Furthermore, attempts to develop an hGH detection test using hGH measurements in urine by Saugy et al (Saugy et al. 1996) have also proven not to be reliable. The primary reason for the difficulties involved with the measurement of urinary hGH lies in the fact that the concentration of hGH in urine is much lower

than in blood, because only 0.001–0.010% of pituitary GH secretion is excreted in urine. Circulating GH is cleared from the bloodstream through degradation, predominantly in the liver and kidney. In the kidney, GH is cleared principally by glomerular filtration, uptake and degradation in the proximal tubule, and only minute quantities of intact GH appear in the urine. Furthermore, the complex molecular changes resulting from the renal excretion process are not fully understood (Rigamonti et al. 2005) and the relationship between blood and urine GH is so variable that it could never be used in the forensic environment of doping control.

In order to overcome the problems involved with measuring GH reliably in urine samples, two main methodological approaches are currently being considered which are based on the use of serum samples. One method directly determines the circulating levels of different molecular isoforms of GH (Wu et al. 1999). The second method relies on the measurement of the circulating concentrations of one or more of GH-dependent proteins, in relation to the normal reference ranges for the competing athletes (Wallace et al. 1999).

A further problem with direct measurement of serum GH as a method of detecting GH abuse is the fact that exercise is a strong stimulus to hGH release (Jenkins 2001). Furthermore levels of the GH dependent marker IGF-I, has been suggested to be associated with elevated baseline levels with chronic exercise (Sartorio et al. 2004b). It has been shown by Sartorio et al that volunteer elite athletes from different sport disciplines (particularly females), who were not receiving any drugs, can have increased baseline GH concentrations, which are not related to the normal IGF-I levels that are seen (Sartorio et al. 2004b). Furthermore, GH release may be influenced by variations in nutritional intake and the nutritional supplements favoured by athletes. Therefore, in the field of sports it would be difficult to prove GH abuse in an athlete, where athletes undergo intensive training regimes, and follow special food diet.

1.6.6 DESCRIPTION OF GROWTH HORMONE DETECTION METHODOLOGIES

A description of the methods currently under development for detecting GH abuse is provided in this section. At present, there are two different approaches used for developing methodology to detect GH abuse in sports; one that has been misnamed as the "direct" approach and the other the "marker" approach as proposed by the GH-2000 group. This section is split in two parts and discusses all methodology to date developed using these two approaches.

1.6.6.1 THE GH ISOFORM METHOD

The first method to be discussed is the one that was selected by WADA and the IOC for priority development and was implemented for the first time at the Athens Olympic Games in 2004 (Saugy et al. 2006). This method involves the analysis of GH isoforms. This method was originally referred to as the "direct method", because it involves assessment of the various forms of GH. More recently, this method has been more accurately referred to as the "isoform assay method" as it is itself an indirect method.

The pituitary gland secretes a number of GH isoforms: the most abundant is the 22 kDa isoform but there are other minor products of gene transcription and post-transcription modification. The most prominent of these is a 20 kDa isomer that is normally present in the circulation at concentrations averaging around 10% of the 22 kDa isomer. The presence of different isoforms provided the basis for the formulation of the isoform assay method, developed by Wu, Bidlingmaier and Strasburger. This methodology attempts to distinguish between circulating endogenous GH secretion and exogenous recombinant human GH (but not pituitary-derived GH) administration by measuring the ratio between the concentration of different isoforms. If a blood sample is obtained during a peak of endogenous GH secretion, then at least 10% of the immunoassayable GH should be 20 kDa GH. Therefore, if the amount of 20 kDA

measured is significantly lower than 10% of the immunoassayble GH, a doping case is suspected.

Exogenous rhGH administration leads to a marked decrease in the endogenous pituitary-derived non–22- kDa isoforms by negative feedback mechanisms as indicated by Hashimoto in 2000 (Hashimoto et al. 2000).

Work undertaken by the research group developing this "Isoform test" method, concluded that following screening of a large panel of monoclonal antibodies (mAbs) raised against either recombinant or pituitary derived growth hormone some of the antibodies investigated, recognised all isoforms whereas others recognised mostly the 22kDa isoform. Using this information, two different sandwich type immunoassays were developed. Assay I (capture mAb 5D7, tracer mAb biotinylated 10A7) for measuring 22 kDa GH, whereas the "permissive" assay 2 (capture mAb 1H6, tracer biotinylated mAb 10A7) is used for measuring all of the isoforms (Bidlingmaier et al. 2000; Bidlingmaier et al. 2003). Measuring serum samples using both assays allows calculation of the relative abundance of the 22kDa GH. For each sample, the result (in ng/ml) obtained for assay I is divided by the result (ng/ml) obtained for assay 2. Hence, a high ratio of 22- to non–22-kDa has been proposed as a mechanism of detecting exogenous GH usage (Veldhuis 1998; Wu et al. 1999).

In accordance with WADA guidelines, it is required that each ratio of the values should be obtained twice (for screening and confirmatory tests), using different assays. To determine an adverse finding, two ratios must therefore be calculated, one for the screening test and one for the confirmatory test. An adverse finding will only be reported if the screening test and confirmatory test ratios both exceed a specified value.

Age, sex, sporting discipline, ethnicity and pathological states do not seem to affect the relative proportions of GH isoforms (Erotokritou-Mulligan et al. 2007).

The isoform test has a number of limitations that may affect its utility particularly in the post-competition setting. Banfi et al reported that blood samples obtained from

athletes performing a marathon at 4000 m of altitude, had significant increases in the 22 kDA isoform levels following the examination of samples collected immediately post-exercise (Banfi et al. 1994). This increase may decrease the sensitivity of the isoform test when applied to samples collected post-exercise (Wallace et al. 2000; Wallace et al. 2001).

A further disadvantage of this method is the fact that GH isoforms have short half-lives and spontaneous GH secretion returns to baseline values 48 hours after the last dose of GH treatment. This means that the window of opportunity for detection is short, up to 36 hours. In addition, the isoform method cannot detect pituitary-derived GH doping (pituitary-derived GH from both animals and humans is still commercially available) or the abuse of GH secretagogues or IGF-I. Furthermore, the isoform test method has low sensitivity because the lowest limit of detection for this method is currently relatively high in comparison to what would be expected to be found in samples in the event of GH abuse.

Application of this method to a set of data by the research group that developed this method, from a GH administration study (n=30) comparing to data collected after the stimulation of endogenous GH secretion (n=125), showed that the calculated ratios for each of the two samples did not overlap (Bidlingmaier et al. 2000). The only independent publication validating this GH detection methodology found the method had a low sensitivity in samples taken 24 hours after the last GH injection (Graham M.R. et al. 2008).

1.6.6.2 THE GH-DEPENDENT BIOMARKER APPROACH

The second methodology currently being used for developing a test to detect hGH abuse is known as the "GH marker" approach or the GH biomarker approach. This method involves the measurement of circulating concentrations of two GH-sensitive substances. A positive test would need values exceeding physiological variability. The marker approach has a major advantage over the "direct" approach in that concentrations of the biomarkers are more stable in serum than growth hormone and

therefore provide a wider "window of opportunity" for detecting an athlete who has been abusing GH.

To be able to determine reliable and robust physiological ranges it is necessary that the selected GH-dependent markers to be considered in developing a method to detect GH abuse are stable across the population group and are easily and reliably measured. The properties the ideal marker(s) should possess for consideration in the development of a test are:

- The effect of acute endurance-type exercise should not be significant. If an effect is present, this should be transient.
- The markers should be highly responsive to GH administration (at rest and after acute endurance-type exercise).
- The washout or disappearance kinetics after cessation of GH administration (at rest and after endurance-type exercise) should be long, thus allowing detection of the markers for a long period.
- The intra-individual variation (day-to-day stability) of the measured concentration of these markers should be small in placebo treated individuals, thus allowing the determination of reliable physiological ranges.
- The concentration of the chosen GH-dependent markers should be unaffected by ethnicity, illnesses, injuries, heavy training or competition.

1.6.6.2.1 IDENTIFYING SUITABLE GH-DEPENDENT BIOMARKERS

The GH-2000 project considered a large number of GH-dependent markers; particularly from two groups; the insulin-like growth factor axis (liver-produced) (Wallace et al. 1999) and markers of the bone and collagen turnover (Wallace et al. 1999; Longobardi et al. 2000). This was done in order to identify the best markers to be used in developing a method to detect GH abuse. A large number of GH-dependent markers were considered in order to identify the most suitable markers to be considered for developing a reliable detection method. Following a GH administration

washout study, the list of variables to be considered for this method was shortlisted to 13 markers.

These GH dependent markers were further investigated by the GH-2000 group and are given in Table 1-1.

GH-Dependent Markers Considered
Serum GH (Growth Hormone) and 22K GH
Serum GH Binding Protein (GHBP)
Serum Insulin Growth Factor-I (IGF-I)
Serum free IGF-I
IGFBP-1 (Insulin Growth Factor Binding Protein-1)
IGFBP-2 (Insulin Growth Factor Binding Protein-2)
IGFBP-3 (Insulin Growth Factor Binding Protein-3)
ALS (Acid Labile Subunit)
Osteocalcin
BS-ALP (Bone Specific Alkaline Phosphatase)
PICP (Carboxyterminal Propeptide of type I procollagen)
P-III-P (Pro-collagen III N-terminal extension peptide)
ICTP (Carboxyterminal cross-linked Telopeptide of type I Collagen)

TABLE 1-1 - GROWTH HORMONE DEPENDENT BLOOD AND URINE MARKERS CONSIDERED BY THE GH-2000 RESEARCH GROUP IN ORDER TO IDENTIFY THE MOST SUITABLE MARKERS TO BE USED IN DEVELOPING A TEST TO DETECT GH ABUSE.

Studies undertaken by the GH-2000 group examined the concentrations of these markers in either trained or non-trained individuals during and following a GH administration period. The effects of acute exercise were also examined on the concentrations of these markers (Wallace et al. 1999; Wallace et al. 2000). This was done in order to allow identification of the most stable markers that would be minimally affected by exercise, be highly responsive to GH administration and have a long disappearance life, thus enabling detection of GH administration for long periods (Table 1-1). A thorough examination of the considered markers resulted in some of them being rejected. The markers that were not rejected were put forward for further investigation in a full study (Longobardi et al. 2000; Dall et al. 2000).

Conclusions from the initial GH-2000 pilot study (Wallace et al. 1999; Wallace et al. 2000), was that acute exercise transiently increased all components of the IGF

ternary complex (IGF-I, IGFBP-3, and ALS) and GH pre-administration further increased this response to acute exercise but that these changes were trivial when compared with the responses to GH. Total IGF-I, IGFBP-3, and ALS were markedly increased by supraphysiological doses of GH.

It was proposed that total IGF-I, IGFBP-3, and ALS might be suitable markers of GH abuse based on the considerations that each analyte is GH regulated and the elevations of serum IGF-I, IGFBP-3, and ALS can only occur in a limited number of physiological conditions, namely puberty, late pregnancy, and acromegaly. Concentrations of these markers were found to be relatively stable throughout the day, with minor reductions during the early hours of the morning. Small changes were observed during acute exercise. These three markers showed substantial increments after hGH administration (greatest for IGF-I and least for IGFBP-3); and the disappearance kinetics suggested that detection might be possible for several days after cessation of GH treatment (Wallace et al. 1999).

The use of free IGF-I as a marker for GH abuse was difficult because its concentration levels are too variable. Concentrations of IGFBP-1 showed this marker to be too variable to be of any use. Serum IGFBP-2 showed a minor fall in the response to GH treatment, unlike all the other markers that rose, confirming the suggestion of Kicman et al that it is a possible marker for GH detection. However, there was only a small change observed following GH treatment and there was a large degree of overlap between individuals before or after exercise, and also there was an unexpected rebound to values higher than baseline after the end of GH treatment suggesting that IGFB-2 would not be a useful marker (Wallace et al. 1999).

Acute endurance-type exercise stimulation of markers of bone and soft tissue formation and bone resorption and GH administration resulted in a much larger increase in these markers. Such changes were found to persist in some cases for at least 96 h. This is considered an advantage to using these markers, as it allows for the possibility of prolonged detection of GH. The most promising markers included P-III-P and ICTP, as they showed little sensitivity to acute exercise and large stable changes

in response to treatment. In addition, these markers showed little variation in placebo-treated individuals over time. The day to day stability for each individual for these two markers was good and using the serum concentration data, it was shown that there was good separation of GH- and placebo-treated individuals in up to 87.5% of cases (Wallace et al. 1999).

Recommendations from the two articles (Wallace et al. 1999; Wallace et al. 2000) suggested that a test using the combined strategy of markers of GH action in both the IGF/IGFBP axis and bone markers would improve the sensitivity of either approach alone. The discriminatory power of these potential markers was further examined in a large multi-centred double blind placebo controlled study. This was done in order to assess the usefulness of bone and collagen markers and markers of the GH/IGF-I axis in the development of a GH doping test. The study was carried out on a cohort of more than 100 healthy, fit, young adults of both sexes. The effects of GH administration were assessed at 2 supraphysiological doses on serum concentrations of selected markers of GH action on the IGF-I/BP axis (Dall et al. 2000), and also on the serum concentrations of osteocalcin, PICP, ICTP, and P-III-P (Longobardi et al. 2000).

1.6.6.3 THE STATISTICAL METHODOLOGY OF THE GH-2000 METHOD

1.6.6.3.1 DISCRIMINANT ANALYSIS

Discriminant analysis is a statistical classification technique which is used when the classification groups are know a priori and the aim is to devise rules which can allocate previously unclassified objects or individuals into these groups in the most optimum way. In order to do this reliably, two sets of multivariate observations are used; a training sample and the test sample. The training sample is the one for which group membership is known already, where the group membership of the test sample is unknown and for which group membership has to be assigned as accurately as possible (Dunn G. et al. 2001).

The methodology of discriminant analysis allows the derivation of statistical functions. These functions are based on a collection of explanatory variables that are predictive of a response variable. When the response variable is categorical and unordered, the statistical problem is known as discriminant analysis. Discriminant analysis is commonly used in problems where the response variable denotes group membership. The discriminant functions are used to predict group membership from a set of predictors.

Discriminant function analysis aims to find a transform of the explanatory variables, which will give the maximum ratio of difference between a pair of group multivariate means to the multivariate variance within the two groups.

In order to evaluate the performance of a derived discriminant function, a commonly used approach is the "leaving-one-out method". Using this method the discriminant function is derived from just n-1 members of the sample and then used to classify the member not included. This process is carried out n times, leaving out each sample member in urn.

Discriminant analysis is used in a variety of settings; ranging from the financial world, geology, environmental studies and medicine. The type of problem this statistical method is applied to is when it is necessary to predict group membership for a subject from a set of predictor variables. A possible application of discriminant analysis is the use of this methodology to predict whether an individual has a disease or not, e.g. lung cancer. Possible predictor variables to be used for determining this could be number of cigarettes smoked a day, coughing frequency coughing intensity etc.

1.6.6.3.2 THE GH-2000 DETECTION METHOD

Results from a randomised, double blind placebo controlled study, administrating two different doses of hGH to normal subjects were used for the purposes of assessing the effects of nine potential markers of GH abuse, and to study the behaviour of the these markers after GH treatment was withdrawn (Longobardi et al. 2000; Dall et al. 2000).

Since the test is to be applied to the population of competitive athletes, it would have been more appropriate for the data on which a test is based, to be obtained from that population. However, it is considered inappropriate for ethical reasons to administer GH to competing athletes. Hence, a double blind trial involving "normal" aerobically trained amateur athletes was used to simulate the process.

In order to be able to apply the results to the elite athlete population, results from this double-blind trial were calibrated against the cross-sectional trial performed on elite athletes (Healy et al. 2005). This was a separate study which collected information from blood samples from 813 elite athletes within 2 hours of a major National or International event, for the purposes of establishing reference ranges for this group (Healy et al. 2005).

The statistical analysis performed for the purposes of developing a methodology to detect GH abuse was done in two parts.

Using the results from the double blind study, the most suitable markers and combination of markers were identified by assessing the discrimination between the placebo and the GH treated participants of the double blind study (Longobardi et al. 2000; Dall et al. 2000).

The properties of the marker combinations identified as the most suitable, were studied in the elite athlete population sample (Healy et al. 2005), in order to assess whether any subgroups of this population might be at an increased risk of a false positive result. Information derived from examination of these markers in the elite population, were taken into account when devising the statistical methodology, by adjusting the methodology in such a way to eliminate the risk of false positive identification.

Using statistical analysis, the general levels of GH sensitive markers were found to be significantly different between the "normal" volunteers used in the double blind study and the elite athletes for whom the test will be used (Longobardi et al. 2000; Dall et al. 2000).

The statistical techniques used to devise suitable marker combination included discriminant analysis, binary and ordinal logistic regression as well as many univariate and multivariate diagnostic and model building methods (Powrie et al. 2007). The concentration of IGF-I and P-III-P is log-normally distributed across age and in line with best statistical practice, all markers were expressed on the natural logarithmic scale. It was found that the use of a combination of markers, proved to be a better method of discrimination, instead of using information collected from a single GH marker.

1.6.6.3.3 DERIVING THE DISCRIMINANT FUNCTION FORMULAE

From the GH administration double blind study carried out by the GH-200o research group (Dall et al. 2000; Longobardi et al. 2000) it was found that the patterns of marker levels changed over time. Several markers were elevated on days 21 and 28 while the drug was still being administered, but the rate at which the markers declined over the washout period varied for each marker. This suggested that some combinations of markers would discriminate well during the treatment phase (days up to 28), while others would be discriminating better during the washout phase. It was also determined that statistical discrimination procedures should be made on day 21 of the double blind trial, when those taking the active drug would be likely to have levels of the various markers most clearly distinguishable from placebo levels (GH-2000 group 1999; Powrie et al. 2007). The reason Day 21 was used to devise the statistical methodology, was that there was some evidence of non-compliance for some of the subjects on Day 28. Therefore, it was considered more appropriate to use values from visit days where there was no doubt on the reliability of the samples considered.

The analysis was concentrated on the Day 21 data, carrying this out for Males and Females separately, concentrating on discriminating between the placebo and the combined HGH groups, since it was thought that on this day best discrimination between the treatment groups was achieved. The marker combinations EM1 (for males) and EF3 (for females) were the best ones found from the discriminant analyses concentrating on Day 21 results. The naming used to denote the two discriminant

functions used a informal naming convention. The letter "E" stands for the word "Equation" and the letters "M" and "F" stand for the words "Males" and "Females". Since there were a number of discriminant functions developed until EM1 and EF3 were concluded as the best discriminant functions, each functions was numbered EM1, EM2, EM3 and so on.

It was found that for both males and females the markers that were consistently of value in discriminating between the treatment groups in the double blind trial were IGF-I and P-III-P. With the marker levels considered using the logarithmic scale, the formulae calculated from the statistical analysis were (GH-2000 group 1999):

For Males: $EM1 = -2.269 + 0.7207 \times log(P\text{-}III\text{-}P) + 0.5210 \times log(IGF\text{-}I)$

For Females: $EF3 = -4.973 + 1.1317 \times log(P\text{-}III\text{-}P) + 1.0125 \times log(IGF\text{-}I)$

The interpretation of these formulae, is that the larger the calculated value using the individual's concentration values of P-III-P and IGF-I, the more likely it is that the individual was involved in receiving the active drug.

As previously discussed, the distribution of the individual GH-dependent markers is not the same for "normal" individuals and elite athletes. Following the calculation of the discriminant formulae, it was therefore necessary to calibrate these formulae using data obtained from the elite athlete population. It is known that both P-III-P and IGF-I are significantly affected by the age of an individual (Juul et al. 1994; Hertel et al. 1993), and it is therefore to be expected that the calculated scores from the discriminant formulae will also be affected by the age of an individual.

The performance of the formulae is judged by comparing their predictions with the actual assignment of volunteers to groups in the Double-Blind trial. To assess performance of the markers, all non-GH treated observations derived from volunteers on placebo treatment, and all observations derived from baseline (Day 0) measurements, were grouped together. This included the placebo group: all visits (Males-142, Females 132), plus the GH groups: Day 0 (M 29, F 29).

It is necessary to adjust for the effect of age when considering for the effect of age, as legal and practical problems could arise in the situation where some athletes are in greater risk of being found guilty of GH doping than others. This would be the case if athletes of different ages have distributions for the doping test criterion, which differed from the majority.

As indicated in published reports, the dependence on age for IGF-I and P-III-P has a non-linear form. The statistical procedures used to adjust for this age was regression based. It was found that the most successful analysis involved age expressed as a reciprocal of age, which was applicable for both genders.

The resulting age adjusted function for males, which removes the dependence on age is:

*GH-2000 Detection Method=-6.586+2.905*log(P-III-P)+2.100*log(IGF-I)-101.737/age*

where age is measured in years.

This age-adjusted marker has been standardised to have mean 0 and standard deviation 1 over the sample of elite male athletes.

Similar calibration was undertaken for female elite athletes, leading to the age-adjusted derived method for females:

*GH-2000 Detection Method=-8.459+2.454 *log(P-III-P)+2.195* log(IGF-I)-73.666/age*

which also has mean 0 and standard deviation 1 over the sample of elite female athletes.

It should be noted, that because the adjustment employed here is monotonic (strictly increases) with age, it would be invalid to use these derived markers to detect GH abuse in pubertal or pre-pubertal athletes, or the elderly population, as it is known that the effect of age in these groups is different than that compared to adults (Juul et al. 1994; Hertel et al. 1993). Appropriate age adjustments can be calculated for these markers, by considering samples of the relevant populations of younger or older ages.

Using the data available from the cross-sectional study involving elite athletes, the physiological reference ranges were calculated for the two markers. The distributions of both age-adjusted derived markers were taken to be of the standardised normal distribution. A 99% reference range can therefore be taken as the range (- 2.576, +2.576). These have been presented in (Healy et al. 2005). The purpose of these reference ranges would be to calibrate the results from the discriminant functions against these reference ranges, as it is known that elite athletes will have higher P-III-P and IGF-I serum concentrations, compared to the healthy volunteers who gave the data necessary to develop the functions (Sartorio et al. 2004b).

The choice of a cut-off value beyond which an athlete would be found guilty of doping, requires arguments based on legal interpretations of phrases such as "beyond reasonable doubt". For the standardised normal distribution, only about 1 observation in 32,000 exceeds a value of 4, while the proportion exceeding a value of 4.5 is less than 1 in 3,000,000. The precise cut-off point however depends on legal requirements for avoidance of false positives.

Sensitivity and specificity of these tests can only be roughly estimated. This is because in order to calculate the exact precision of these tests it is required to assume that those doping with hGH in a real life scenario would follow the protocol used in GH-2000 Double Blind trial. The following is a table of the estimated sensitivity the two gender-specific discriminant functions, presenting the success rate of the functions, under various cut-off limits.

Cut-off	3.5sd	3.7sd	4.0sd	4.5sd
Approximate Specificity	1 in 5000	1 in 10,000	1 in 32,000	1 in 340,000
Success rates and estimated average sensitivity – Males				
Day 21	82%	82%	79%	61%
Day 28	86%	86%	86%	68%
Day 30	63%	54%	50%	29%
Success rates and estimated average sensitivity - Females				
Day 21	41%	33%	30%	19%
Day 28	38%	38%	35%	23%
Day 30	30%	26%	19%	11%

TABLE 1-2 - ESTIMATED SENSITIVITY THE TWO GENDER-SPECIFIC DISCRIMINANT FUNCTIONS, PRESENTING THE SUCCESS RATE OF THE FUNCTIONS, UNDER VARIOUS CUT-OFF LIMITS (GH-2000 GROUP 1999).

It is clear from Table 1-2 that the higher the cut-off point is set; the less discrimination is achieved with the proposed discriminant functions proposed.

1.6.6.3.4 ADVANTAGES AND DISADVANTAGES OF THE GH-2000 DETECTION METHOD

One potential disadvantage of this approach however is that concentrations of these indirect biomarkers such as IGF-I or IGFBPs is that they, like GH secretion decline with age. However, it has been shown that these markers are largely independent of factors such as body weight, BMI, physical activity, diet and sex steroids, which allows these to be used in the development of a robust GH detection methodology. For example the concomitant use of rhGH and anabolic steroids may reverse the GH-induced increase in IGFBP-3, even when IGF-I is increased significantly (Karila et al. 1998) although more extensive studies by Ho and colleagues in Australia do not show this effect (Ho 2007 personal communication). It has been suggested that athletes often use growth hormone in conjunction with other performance enhancing substances, such as anabolic steroids (Kraemer W. et al. 2002). Therefore, the use of IGFBP-3 as means of detection of hGH may not be advisable (Rigamonti et al. 2005).

A further problem identified with the use of GH-dependent markers, is the use of oestrogen when taken orally or transdermally at high doses. It was found by examining the effect of oestrogen in postmenopausal females on the serum concentration levels of IGF-I, IGFBP-3 and ALS, that this led to a reduction in the concentrations of these marker concentrations (Kam et al. 2000). Rigamonti et al suggested that in order to eliminate the possible confounding effect of oestrogen use, when these biomarkers are to be used to detect hGH abuse, it would be better to use the ratios of IGF-I:IGFBP-2 and IGFBP-3:IGFBP-2, which were shown by Kicman et al to increase after acute administration of hGH (Kicman et al. 1997; Rigamonti et al. 2005). However, the study by Kicman et al showed that the increase in these ratios only lasted for up to 30 hours, following hGH administration and other studies have failed to confirm the GH induced rise in IGFBP-2 (Healy et al. 2005).

These issues mean that the determination of normal physiological ranges of the proposed biomarkers are crucially important, as a doping offence can only be considered in individuals whose biomarkers exceeds the normal range.

It is useful to note that 'False Positives' occur through values exceeding the threshold just by chance only at the probability level chosen. Therefore, if the cut-off limit has been set at 1 in 10,000 it is then expected that the number of false positives, which would occur purely by coincidence, would be 1 in 10,000.

1.6.6.4 ALTERNATIVE USE OF THE BIOMARKER APPROACH

Based on the findings proposed by the GH-2000 project, Rigamonti et al proposed a different method of detection. In order to avoid the various limitations that might affect the variation in the markers, such as inter-individual variations, intensity, type of training and sport discipline, they suggested that a longitudinal study is carried out, in which each subject would act as their own control. Furthermore, they suggested that such an approach could be used for testing purposes, by collecting

blood samples from athletes, from which measurements of the biomarkers would be derived, and these would act as the athlete's "endocrine passport".

Rigamonti et al proposed a series of first level tests, in which the measurements of the two groups of biomarkers would be taken and if doping is suspected a second level test would be performed, which would be GH profiling of the athlete. Using this profiling method, they suggest that it would be easy to detect GH dopers because of the mechanisms undergoing in the body system following GH administration. It is expected that either GH administration would disrupt the physiological GH profile and will produce raised stable concentrations of the hormone or, more likely, suppress GH secretion since the pulsatile release of GH would be stopped due to the activation of feedback mechanisms, and thus be able to decide whether GH doping occurred.

The disadvantage of this method would be in the first level testing, in which athletes might be doping when this would occur, and carry on throughout the whole period the "endocrine passport" is established. Furthermore, in order to be able to assess whether an athlete has indeed got abnormal values in his "endocrine" passport, confounding factors that might affect the interpretation of these values, leading to a possible false accusation of athletes, such as nutrition, injury recovery, training method need to be investigated. If athletes are wrongly accused, then there would be a waste of resources, and furthermore it will cause athletes to lose trust in the doping detection system. Lastly, in order for this method to be proven reliable, it is needed to show that the concentrations of the concerning markers are expected to be stable with little variability over time.

1.6.6.5 THE SERUM PROTEOMIC PROFILING METHOD

Serum proteomic profiling uses "surface-enhanced laser desorption/ionization time-of-flight mass spectrometry" (SELDI-TOF MS), a type of serum proteomic profiling to detect proteomic patterns of GH administration. In order to achieve this, protein profiles are obtained from control and patient groups in order to identify differences in protein expression.

Chung et al reported that using the measurements of the haemoglobin alpha-chain from blood samples collected by the GH-2000 group in a double blind, placebo controlled, hGH administration study, there was good discrimination between the two treatment groups (Chung et al. 2006). This study demonstrates that protein mass profiling is a possible tool for the detection of GH administration and suggests that measurement of haemoglobin alpha-chain may have utility as a novel serum biomarkers of GH action. This study has not been repeated and there are no indications that it has adequate sensitivity or specificity for the job.

1.6.6.6 FURTHER WORK ON THE GH-2000 TEST: THE GH-2004 PROJECT

The results of the GH-2000 project were reviewed by a panel of international experts (including a representative of the Court of Arbitration in Sport – CAS) at an IOC organised Workshop in Rome in April 1999. This review produced a number of key issues that needed addressing before the experts felt that there would be viable test suitable for implementation at Olympic Games (GH-2000 group 1999). The main issue raised concerned possible ethnic effects on the GH sensitive markers proposed, since the majority of volunteers in GH-2000 were white Caucasians.

In response to the issues raised at the Rome workshop, the GH-2000 team put together a further submission to the European Union and WADA. This proposal was rejected by the European Union; due to the long length and large cost of the submitted proposal. However, WADA proposed some changes to this proposed research by the GH-2000 group, mainly by reducing the length and size of the proposed research, thus making it a less expensive study.

Following funding and support from WADA and USADA the GH-2004 project was formed based at the University of Southampton. The GH-2004 project made plans to measure only two markers; IGF-I and P-III-P, since these were shown by the GH-2000 project to provide the best sensitivity and high specificity when used to detect GH administration. Four studies would be involved, three of which would replicate the

previous "Double-Blind" and the "Cross-Sectional" studies from GH-2000, but this time concentrating on volunteers from the African and Caribbean, Indo-Asian and Oriental communities within the UK. Furthermore, two further studies were undertaken to obtain further information on the possible effects of injury on the GH-dependent serum markers – the "Injury" study, and to obtain information on the physiological variability of IGF-I and P-III-P, in order to ensure that the magnitude of this would not adversely affect the outcome of the GH-2000 detection method.

1.7 IMMUNOASSAY STANDARDISATION AND THE GH-2000 TEST

The GH-dependent biomarker test approach researched by the GH-2000 group, utilises the concentration values of two GH-sensitive markers; IGF-I and P-III-P that are used in a discriminant function relationship to identify GH abuse. The quantification of the concentration of these two markers in serum is by the use of radioimmunoassay (RIAs).

The reference ranges produced by the GH-2000 group, as well as the discriminant functions "are specific to the assays and reagents used in the GH-2000 project. It is the nature of immunoassays that measurements are not absolute and inevitably a function of the reagents and conditions used in the assay"(Healy et al. 2005). Healy et al suggest that if RIAs are to be used in a test for GH abuse that may lead to sanctions on an elite athlete, it will be necessary for the laboratory concerned to be able to demonstrate validation of their assays with the results reported by the GH-2000 group. Similarly, in order to make direct comparisons of the new results concluded by the GH-2004 project, it is necessary to demonstrate validation of the assays used by the GH-2004 group, with the results reported by the GH-2000 group. Failure to do so, would suggest that any additional findings on areas such as the effect of ethnicity, are meaningless, as they cannot be tested against the results shown by the GH-2000 group.

Furthermore, it is required that any adopted GH detection methodology which utilises immunoassays must use a procedure with a different antibody that should recognise a different epitope of the protein than the assay used for screening. As a result this make it necessary that the GH-2000 detection method can be employed using IGF-I and P-III-P concentrations measured from different immunoassay methods. Therefore, it is necessary that necessary methodology is derived to ensure that the GH-2000 detection methodology is not specific to a particular immunoassay.

At the planning stages of the GH-2004 project, in order to make immediate comparisons of their results with those results derived by the GH-2000 project, it was

planned to use the same RIAs as the GH-2000 project. However, the development of reliable methodology to detect GH abuse has been a lengthy process. During the elapsed period between the GH-2000 and GH-2004 research, some of the RIA manufacturers have updated their assays and/or reagents, and others have stop production of RIAs.

The specific purpose of this study is to describe the necessary methodology required for the calculation of reliable "assay-adjusting" conversion equations to be applied on data collected from the various research groups involved in the development of a GH doping test. This is important, because at present there are no international reference standards for the immunoassay measuring concentrations of the P-III-P biomarker, and for IGF-I even though a WHO reference standard exists, this standard has been heavily criticised as being inappropriate and unreliable (Abellan et al. 2005; Ranke et al. 2003).

Furthermore, by employing the methodology developed in this study, it will become possible for results produced by other research groups utilising different RIAs or other methodology to make direct comparisons with the findings of the GH-2000/GH-2004 study.

A similar situation occurs with the assaying HbA_{1c}, which started since the mid 1980s. HbA1C methods produced inconsistent results, which varied between assay types and between manufacturers. This was mainly because glycohaemoglobins are heterogeneous and that numerous different methods measured different glycated species with different reliability, including lot-to-lot variability, and often without any form of standardisation. Collaborations between the IFCC, IUPAC and International Diabetes Associations are currently undertaken with the purpose of establishing true reference methodology and pure HbA_{1c} calibration material. This material when internationally used, together with an agreement between the on the choice of one unit and one numerical value throughout the world, will be the final significant stage in the global use and true value of the glycohaemoglobin assay (Goodall 2005). Similar

work needs to be undertaken for the assay methodology currently being employed to measure IGF-I and P-III-P.

1.7.1.1 THE NEED FOR ASSAY STANDARDISATION

Values measured from a particular immunoassay are specific to the assays and reagents used at that particular point in time. As previously discussed, it is the nature of immunoassays that measurements are not absolute and inevitably are a function of the reagents and conditions used in the assay. Therefore, the ideal scenario would be if every immunoassay used could be calibrated against a known amount of the protein being measured. This would ensure that measurements from each immunoassay are directly comparable across immunoassays and manufacturers.

The issue of standardisation of protein hormones dates back to the 1920s when the first International Standard for insulin was established. With the increasing utilisation of other protein hormones, significant developments in standardisation were made, including the establishment of standards for growth hormone, gonadotrophins, ACTH and calcitonin. Bioassays are comparative procedures in which the activity of a test is compared to that of a reference; the activity is expressed in activity units rather than in absolute terms (Bristow A.F. 1998).

For biological materials, such as proteins and glycoproteins, the inherent variability of different preparations necessitated the availability of a single preparation to which all other standards could be referenced. With the development of radioimmunoassay in the 1960s it was found that the same problems involved in biological assays were also found in the development of immunoassays. Preparations used as assay standards were variable and impure, and the same principles of assay standardisation were therefore applied to the calibration of immunoassays. The importance of International Standards was therefore recognised. For many of these protein immunoassays international standards were then established. However, to date not all immunoassays have International Standards available.

1.7.1.2 REQUIREMENT FOR AN INTERNATIONAL REFERENCE METHOD

Assessing the accuracy of an immunoassay method is in effect a relative exercise. Accuracy is quantified by measuring the extent to which different assay centres agree or disagree in terms of their test results on a given test specimen, which in turn depends on the extent to which those centres conform to a given set of assay specifications. In practice, even in ideal laboratory conditions, there are variations in assay conditions, between different test centres and within any given test centre over a period of time, as absolute control is not possible.

The use of a standard reference material in immunoassays can accommodate such variations. By assaying a reference material in parallel with the unknown test preparation, the measurement unit for the assay is then expressed in terms of the ratio of the test and standard effects or activities, i.e. units of relative activity.

In standardising immunoassays, it is important to adopt uniform assay specifications, the role of the standard reference preparation being to regulate the effects of those methodological variables that do not submit to experimental control. The provision of stable reference materials for the standardisation of immunoassays is a major undertaking, typically organised at a national or even international (WHO) level.

The reference material is assigned an unit of activity, and the test/reference activity ratio can be expressed as a multiple fraction of the reference unit value. The reference material is a good way for regulating minor methodological discrepancies, as it holds the true value of the immunoassay. Accuracy therefore becomes directly traceable to a base unit that can be physically realised in any given test centre.

When the active test material is a single well-characterised molecular species that elicits a unique response in the assay system then the provision of an adequate reference material is greatly simplified, subject to the requirement of stability and the provision of a functionally inert reference matrix. In this ideal situation, it may be attempted to establish a direct correspondence between the activity of the reference

material and its mass in order to facilitate the translation of functional activity units into gravimetric terms (mass/volume). The results of such a translation might not always be valid as there is not always a necessary correspondence between units of activity and units of mass.

If the active material is ill defined in composition and/or molecularly heterogeneous, the validation of the assay rests on the condition of similarity, stated above.

1.7.1.3 ASSAY METHODOLOGY CONSIDERED

For the laboratory analysis of P-III-P, only two RIA kits appear to be commercially available; the RIA-gnost P-III-P from CIS-BioInternational (Yvette Cedex, France) and Intact P-III-P RIA kit (Orion Diagnostica, Espoo, Finland).

A number of immunoassays exist for measuring total IGF-I in serum. Many of them combine an immunoassay with a sample preparation method with the objective of removing IGFBPs. Until recently, comprehensive validation and comparison of results over the whole concentration range did not exist for these kits, nor a comparison of results obtained in different laboratories (Abellan et al. 2005). Although there are WHO International standards (87/518) to calibrate against when using these assays, recently published reports (Abellan et al. 2005) suggest that this method does not work. Juul (Juul 2001) offers an explanation to this possibly being attributable to the fact that the standard has been reported to have low purity (44%).

Two of the assays considered in this study for P-III-P were manufactured by the same manufacturer the RIA-gnost from CIS-BioInternational (Yvette Cedex, France) but were produced at different times. The third assay was the Orion assay, the P-III-P RIA kit (Orion Diagnostica, Espoo, Finland).

Similarly for IGF-I, two of the investigated assays were produced by the same manufacturer; Nichols Institute Diagnostics (San Juan Capistrano, California, USA)

at different points in time, during which time the assay appears to have been updated. The third RIA assay used for IGF-I measurements was manufactured by the DSL Laboratories (Diagnostic Systems Laboratories, Inc., Webster, TX).

Assay manufacturers update their immunoassays over time and there is no standardisation process for the development of immunoassays by different manufacturers. These are key areas that need to be clarified before introduction of the "marker" test. There have been calls for urgent standardization of calibrants and standards as well lower heterogeneity of antibodies and assay reagents for comparison (McHugh et al. 2005). The comparability of P-III-P and IGF-I values obtained by different methods in different laboratories and different countries is limited since there is no internationally agreed reference method available.

CHAPTER 2: METHODS

Overview

In order to address the aims of this thesis, the results from 8 research trials were considered. In this chapter the study design, the study protocol and the analytical methods employed in each research trial considered is given.

The statistical methods used to examine the objectives of this thesis are also discussed in this section.

2.1 SUBJECTS AND STUDY DESIGN

2.1.1 GH-2000 CROSS-SECTIONAL STUDY

2.1.1.1 SUBJECTS

The aim of this study was to determine the physiological range of these GH-dependent variables in elite athletes after a competitive event to determine whether such values differ from resting values in normal and athletic subjects. In addition, this study was undertaken in order to establish whether any adjustments to this range are required based on age, gender, demographic characteristics, or the nature of the exercise performed.

Samples were collected from 781 volunteer elite athletes (athletes competing at the national or international level) within 2 hours of a major competitive event at the regional, national, or international level. In summary 516 males and 265 females were recruited, age range: 17-64 yr from 15 sporting disciplines. Further details can also be found in Healy et al (Healy et al. 2005). Most of the volunteers were of Caucasian background; there were 57 non-white European male subjects and 21 non-white European female subjects.

Male athletes were older (26.5 ± 5.3 *vs.* 25.1 ± 5.4 yr; $P < 0.02$), taller (182.3 ± 8.5 *vs.* 171.1 ± 7.2 cm; $P < 0.001$), and had a greater BMI (24.1 ± 3 *vs.* 21.8 ± 2.5 kg/m$_2$; $P <$

0.001) than female athletes. Age, height and BMI varied widely between the sports, reflecting the different nature of each discipline.

	Gender		Total
Ethnic Group	**Male**	**Female**	
Afro-Caribbean	37	2	39
Caucasian	459	244	703
Other	4	0	4
Oriental	16	19	35
Total	516	265	781

TABLE 2-1 - FREQUENCY NUMBER OF ELITE VOLUNTEERS RECRUITED BY THE GH-2000 GROUP FOR EACH ETHNIC GROUP BY GENDER.

Volunteers competed in one of 7 sporting categories. The frequency breakdown for each sporting category by gender is given in Table 2-2.

Sporting Category	**Sport**	**GH2000**	
		Male	**Female**
Athletics	**Track and Field, Relay**	95	49
Combat sports	**Judo, Boxing**	26	.
Endurance sports	**Swimming, Marathon**	8	7
Power sports	**Weightlifting**	10	7
Power/endurance	**Power-lifting, Canoeing, Rowing**	192	126
Skiing	**Cross-country skiing, Alpine skiing**	37	33
Team ball sports	**Basketball, Football, Ice Hockey, Handball**	148	43
Total		516	265

TABLE 2-2 - FREQUENCY NUMBERS FOR THE GH2000 CROSS-SECTIONAL STUDY BY SPORTING DISCIPLINE AND GENDER.

2.1.1.2 STUDY DESIGN

The results of this research have been published Healy et al (Healy et al. 2005). The nature of the project was explained to the athletes before their participation in their event, and written informed consent was obtained immediately before blood sampling. All samples were taken by a member of the GH-2000 team and collected according to a standardized protocol at the sporting venue where the event had taken place. Before sample collection, the following demographic data were recorded: gender, race, height (self-reported), weight (measured using the Tanita TB7–305 bioimpedance analyzer, Tanita, Tokyo, Japan), age and sporting category.

During sampling, volunteers were seated, and blood was drawn from a vein in the antecubital fossa using a Vacutainer (BD Biosciences, Franklin Lakes, NJ) and 20-gauge needle into two 5-ml bottles containing SST clot activator gel. Samples were left to clot for 15 min at room temperature before being centrifuged at 1500 rpm in a portable centrifuge. Two-milliliter aliquots of serum were then transferred by pipette into small storage tubes (Cryotubes) and immediately placed on dry ice. They were subsequently transferred on dry ice to storage at 80C in the four participating GH-2000 centres. The West Lambeth Health Authority gave ethics approval for this study.

All samples were coded and kept at 80C until analyzed at one of two central reference laboratories: Sahlgrenska Hospital (Gothenburg, Sweden) and Kolling Institute (Sydney, Australia).

2.1.2 GH-2004 CROSS-SECTIONAL STUDY

2.1.2.1 SUBJECTS

The aim of this study was to assess the effect of ethnicity on the serum concentrations of IGF-I and P-III-P, and validate the physiological ranges determined by the GH-2000 cross-sectional study in the elite non-white European athlete population.

The majority of volunteers examined in GH-2000 cross-sectional study were of white European ethnic background. Therefore, it is necessary to examine whether findings concluded by the GH-2000 research group and in particular the GH-2000 detection methodology can be extrapolated to athletes of non-white European origins. Although currently there is one publication examining the effect of ethnicity on the concentrations of IGF-I and P-III-P, it is not known how ethnicity might affect the performance of the detection methodology. This key issue needs addressing before it can be implemented at an Olympic Games event.

Post-competition samples were collected from 304 athletes at 9 international sporting events from 10 different sporting categories. Samples were collected from 242 male (aged 29.3 ± 0.5, 16 – 62 yrs) and 62 female elite athletes (29.3 ± 1.3, 16 – 52 yrs) from 50 different nationalities, in a setting that is comparable to post-competition anti-doping procedures. The ethnic group breakdown of the volunteers is given in Table 2-3. Blood samples were collected within 2.5 hours of a major competitive event at the regional, national or international level, with 81.4% of the observations collected within 30 minutes, 15.2% observations between 30-60 minutes and the rest within 120 minutes. There were 10 different sporting categories with 19 disciplines, some of which were not represented in the original GH-2000 study. The breakdown of athletes and sporting disciplines is given in Table 2-4.

	Gender		**Total**
Ethnic Group	**Male**	**Female**	
Afro-Caribbean	155	39	194
Arab	20	0	20
Caucasian	34	10	44
Indo Asian	19	6	25
Mixed	5	4	9
Oriental	9	3	12
Total	242	62	304

TABLE 2-3 - FREQUENCY NUMBER OF ELITE VOLUNTEERS RECRUITED BY THE GH-2004 GROUP FOR EACH ETHNIC GROUP BY GENDER.

Sporting Category	Sport	GH2004	
		Male	Female
Athletics	**Track and Field, Relay**	62	30
Combat sports	**Judo, Boxing**	29	.
Endurance sports	**Swimming, Marathon**	91	18
Racket sport	**Badminton**	22	8
Skill sports	**Shooting**	5	4
Team ball sports	**Basketball, Football, Ice Hockey, Handball**	6	1
Wheelchair Team ball	**Wheelchair basketball**	24	.
Total		241	61

TABLE 2-4 - FREQUENCY NUMBERS FOR THE GH2004 CROSS-SECTIONAL STUDY BY SPORTING DISCIPLINE AND GENDER.

2.1.2.2 STUDY DESIGN

Organizers of 68 international and national events were contacted to obtain permission to recruit subjects at the event. Permission was obtained for the research team to attend the following events during 2003 and 2004: World Badminton Championships (Birmingham, UK), World Athletics Championships (Paris, France), World Cup Shooting Finals (Milan, Italy), World Men's Team Squash Championships (Vienna, Austria), Asian Boxing Championships and Athens Olympic Qualification, Karachi, Pakistan, National road relay (Johannesburg, South Africa), Basketball national trials (Paris, France), Wheelchair basketball national championships (Johannesburg, South Africa) and Soweto Marathon (South Africa).

Despite using techniques that were identical to those used in GH-2000 and letters of support from IOC and WADA, we have been disappointed by the lack of support we have had from some of the international sporting community in allowing us access to their events and athletes.

The following organisations either did not reply or replied to say that they were not interested in taking part:

1. All African Games
2. Afro-Asian Games
3. 2003 World University Games
4. FINA World Swimming Championships
5. All continents Gymnastics Olympic Qualifiers (Europe, Asian, Oceania, America, Africa) Trampolining Olympic Qualifiers
6. Pan- Am Games
7. Handball European and World Cup Championships
8. Judo – British National and World Cup series
9. Netball 2003 World Cup, England Netball internationals
10. Rugby World Cup
11. Rugby Sevens
12. British Premiership rugby union teams
13. British Premiership rugby league teams
14. British Premiership football league
15. World Table Tennis championships and Olympic Qualifiers
16. International Cricket Test Matches
17. Karate European Championships
18. Karate British Championships

19. US National Basketball League,
20. US Women's National Basketball Team
21. British Basketball League
22. London, LA, Munich, Dublin Marathons
23. World Weightlifting Championships 2003, Toronto
24. World and European Wrestling Championships
25. Australian masters Championships
26. Boxing Olympic Qualifiers and national tournaments
27. International Triathlon Union World Championships
28. Modern Pentathlon World Cup Series
29. Hockey British Championships and Olympic Qualifiers
30. British Universities Sports Association
31. World Cycling Championships
32. Singapore Sports Commission
33. Chinese Olympic Committee
34. Thailand Olympic Committee
35. Indian Sports Association
36. Japanese Olympic Committee
37. World Skating Union
38. World Ski and Snow Board Committee

39. Japanese Rugby Football Union

40. Malaysian Ministry for Youth and Sports

41. Malaysian Football Association

42. The Olympic Council for Malaysia

43. Titan Games (Boxing, wrestling, weightlifting, judo, shot putt)

44. Medical Commission for the International Equestrian Federation

45. Medical Commissioner for the International Fencing Federation

46. International Paralympic Committee

47. Pan American Wrestling Championships

48. International Federation of Body Builders

49. International Aikido Federation

50. Grand National Archery Society

51. Lawn Tennis Association

52. British Canoe Union

53. World Cup Bobsleigh

54. International Korfball Federation

55. FIFA - Women's World Cup

56. FIFA - EURO 2004

57. British Baseball and Softball association

58. International Netball Association

Once permission had been obtained with an event organizer, the GH-2004 team liaised with the organizer to ensure that a room in close proximity to the sporting arena was made available for the research. A member of the GH-2004 team attended the medical briefing prior to the event to inform the athlete coaches of the aims and objectives of the research project. There was opportunity for those present to ask questions. At several events, the organizers included details of the GH-2004 project in the athletes' briefing pack.

Following the event, a GH-2004 team member approached athletes as they were leaving the arena to invite them to participate in the research. The athletes were informed of the purpose of the research and were given a written information sheet that had been previously translated into appropriate languages. Written informed consent was obtained. No financial compensation was offered for their participation in the study.

The response rate from the athletes was extremely variable. For example, at the Paris Athletics Championships only 1 in 30 athletes who were approached agreed to take part in the study. This contrasted with the World Men's Squash Championship and the Asian Boxing- Athens Qualification tournament where approximately 1 in 3 athletes agreed to participate.

Subjects were required to confirm in writing that they had not taken performance-enhancing drugs prior to the venepuncture.

The Southampton and South West Hampshire Research Ethics Committee gave ethics approval for this study.

All blood samples were taken by a member of the GH-2004 team and collected according to a standardised protocol at the sporting venue where the event had taken place. The following demographic data were recorded and subsequently anonymised: gender, race, height (self-reported), weight (self-reported), age and sporting category. Data collected on the ethnic group background from the GH-2004 volunteers were classified in one of five groups; White-European, Afro-Caribbean, Indo-Asian, Oriental

and Other. The "Other" classification included volunteers from a mixed ethnic background and Arabic athletes. This was done because there were insufficient samples from each gender available for each combination of mixed ethnic background and from Arabic athletes (n<5).

Blood samples were collected and were centrifuged within 4 hours of collection using a portable centrifuge. Most samples were centrifuged within two hours. At the marathon in South Africa, however, the finishing time for the athletes was spread over several hours. There were no facilities to centrifuge the samples at the event and so had to be transported and separated after the event. All samples were kept on ice until centrifuged as IGF-I and P-III-P are stable for up to 5 days under these conditions. Aliquots of serum were transferred into small storage tubes and immediately frozen. These were then placed on dry ice prior to transport to and stored at -80C at the GH-2004 centre (University of Southampton). All samples were coded and kept at -80 °C until transferred later on dry ice to the WADA accredited laboratory at the Drug Control Centre, King's College in London for analysis of serum IGF-I and pro-collagen type III concentrations. All assay analysis was performed in duplicate.

2.1.3 GH-2004 INJURY STUDY

There have been fears that skeletal injury may have an adverse effect on the performance of the GH detection methodology. This is of particular concern as P-III-P is a marker of soft tissue and bone turnover (Haukipuro et al. 1990; Bail et al. 2001). For example, following a tibial shaft fracture, bone turnover increases for at least 24 weeks, as demonstrated by an increase in P-III-P (Veitch et al. 2006).

It is therefore important to examine whether elevations in either of these proteins following a sporting injury could lead to a false accusation of doping with GH. The aim of this study was to assess the effect of a musculo-skeletal injury on serum IGF-I and P-III-P concentrations in amateur and elite sportsmen and women and most

importantly assess the effect of injury on the performance of the GH detection method proposed by the GH-2000 group.

2.1.3.1 SUBJECTS

In this study 127 male (age 29.6 ± 0.9 yrs, range 17 – 68 yrs) and 30 female (age 32.0 ± 2.1 yrs, range 19 – 63) amateur sportsmen and women were recruited from the Accident & Emergency outpatient clinic or orthopaedic fracture clinic at the Southampton University Hospitals Trust, UK. 16 male (age 23.8 ± 1.4, 17-68 yrs) and 10 female (24.7 ± 1.6, 17-63 yrs) professional athletes were also recruited from the British Olympic Medical Centre, Northwick Park Hospital, UK. All subjects were recruited within 10 days of a bony or soft tissue musculoskeletal injury.

The ethnic background of the recruited volunteers were mostly Caucasian (n=174). There were some Afro-Caribbean individuals (n=5), 1 volunteer of Indo-Asian background, 2 of mixed ethnic origin and 1 of Oriental background. Recruited volunteers had recently sustained either a soft-tissue injury (n=75) or a bony injury (n=108). Some volunteers had both soft tissue and bony injuries, and these were classified under the bony injury group. The severity of the injury sustained by the volunteers was classified as either mild (n=92), moderate (n=48) or severe (n=43) (Table 2-5).

Soft Tissue Injury (n= 75)		Bony Injury (n= 108)	
Mild	N=41	**Mild**	N=51
Bruising or haematoma	n= 4	Fractured digits	n= 24
Dislocated digit	n=2	Fractured metacarpal	n= 12
Ligament sprain	n=19	Fractured metatarsal	n=17
Non-specific back injury	n=4		
Tendon injury	n=10		
Moderate	N=19	**Moderate**	N=33
Acute lumbar Annular tear	n=4	Fractured forefoot	n=4
Anterior cruciate damage	n=10	Fractured patella	n=2
Recurrent epicondylitis	n=1	Fractured radius	n=12
Rotator cuff injury	n=3	Fractured scaphoid	n=11
Ruptured hamstring	n=1		
Major	N=19	**Major**	N=24
Achilles tendon rupture	n= 6	Fractured ankle	n=8
Cruciate Ligament Rupture	n=3	Fractured clavicle	n=9
Dislocated shoulder	n=9	Fractured humerus	n=5
Dislocated elbow	n=1	Fractured tibia	n=2

TABLE 2-5 - FREQUENCIES OF SUBJECTS PARTICIPATING IN THE GH-2004 INJURY STUDY IN EACH INJURED GROUP BY TYPE OF INJURY

2.1.3.2 STUDY DESIGN

The inclusion criteria for the study were volunteers who were actively involved in playing sports, had sustained a musculoskeletal injury within 10 days and were willing to take part in a 3-month study. Exclusion criteria included neoplastic disease, diabetes, pregnancy or lactation and any condition likely to affect the functioning of the GH-IGF axis.

The protocol was approved by the Southampton and South West Hampshire Local Research Ethics Committee. All subjects gave written informed consent and the study was conducted in accordance with the ethical principles of the Declaration of Helsinki and the guidelines of good clinical practice.

A brief medical and sporting history was obtained along with the following data: nature of injury, severity, date of injury, gender, race, age and sporting category. The nature of the injury was classified as either soft-tissue or bony.

Volunteers with bone fractures, cracks or chips were included in the bony injury group. Volunteers with joint dislocations, ligament or tendon injury, sprain injury or bruises were included in the soft-tissue injury group. The severity of the injury was assessed by a physician who was blinded to the IGF-I and P-III-P. The injuries were classified into severe, moderate or mild as follows: a severe injury involved fracture of a major long bone or disruption of a major joint or multiple injuries e.g. fracture of the tibia or dislocation of the knee, rupture of cruciate ligaments. A moderate injury involved fracture of a cancellous bone, wrist or forefoot. Soft tissue injuries included tendon or ligament damage. A mild injury involved fracture of a digit or soft tissue injury only.

A 10ml venous blood sample was obtained from each volunteer. Follow up blood samples were then scheduled on 7 ± 3, 14 ± 3, 21 ± 3, 28 ± 3, 42 ± 7, 84 ± 14 days after the injury. The follow up samples were taken either at the subject's home or in the Southampton Wellcome Trust Clinical Research Facility. In practice, it was not possible to obtain samples from all of these time points for each subject.

Blood samples were collected and were centrifuged within 12 hours of collection. Aliquots of serum were transferred into small storage tubes. They were then coded and immediately transferred to a -80C freezer at the GH-2004 centre (University of Southampton) prior to analysis at the WADA accredited laboratory at the Drug Control Centre, King's College London, UK. Samples taken outside the hospital were stored and transported on dry ice prior to transfer to the freezer.

2.1.4 GH-2000 LONGITUDINAL STUDY – ELITE ATHLETES

2.1.4.1 SUBJECTS

The aim of this longitudinal study was to examine the stability of a number of GH dependent biomarkers using samples collected from elite athletes over a 12 month period in- and out-of-competition, across a variety of sporting disciplines.

The aim of this longitudinal study was to examine the stability of a number of GH dependent biomarkers using samples collected from elite athletes over a 12 month period in- and out-of-competition, across a variety of sporting disciplines.

Up to 4 fasting blood samples were obtained from 175 male and 83 female elite athletes over a period of up to 1 year. The athletes were recruited at national or international sporting events. The vast majority of the athletes were white European with only 4 Afro-Caribbean and 1 Oriental subject. The mean age of the men and women was 25.9 ± 0.4 yrs and 24.7 ± 0.5 yrs respectively. Volunteers recruited competed in a diverse area of sport. The breakdown of number of athletes per category was as following: alpine skiing (n=26), cross-country skiing (n=28), long-distance cycling (n=12), sprint cycling (n=4), football (n=10), rowing (n=90), tennis (n=5), swimming (n=28), triathlon (n=8), track and field (n=32) and weight lifting (n=18). The volunteers in this study were competing at a national or international level at the time of the study.

2.1.4.2 STUDY DESIGN

Athletes recruited in this study were followed for one year, with blood samples taken under different training conditions during the year. From each volunteer up to four samples were taken at rest and one sample after a competition.

Demographic information was collected on age, ethnic origin, sport, level of sport, medication and medical history. Also at each sampling situation, specific details were registered such as time since last exercise. Resting blood samples were taken at an ideal three months interval over a 12 months training cycle. Post competition samples

were collected following a national or international level. Samples were taken within one hour after the event.

Following sample collection, blood samples were allowed to clot and then centrifuged at 15 minutes at 3000rpm before the serum was separated and frozen at -80C prior to analysis.

2.1.5 GH-2004 LONGITUDINAL STUDY – AMATEUR ATHLETES

2.1.5.1 SUBJECTS

The aim of this longitudinal study was to examine the stability of a number of GH dependent biomarkers using samples collected from amateur athletes over a 12 month period in- and out-of-competition, across a variety of sporting disciplines.

In total 41 amateur athletes (20 men and 21 women aged 18-42 yrs) were recruited through the University of Southampton. Most volunteers were members of sports clubs and participated in team or individual sports. 2 subjects dropped out of the study after the first visit; one because of injury and the other for personal reasons. Exercise testing was performed on 35 volunteers.

The majority of volunteers were of Caucasian background. 39 volunteers were of white European background and 2 of Indo-Asian background.

Male and female volunteers were of similar ages (20.9 ± 0.76 vs. 20.8 ± 0.37yr). Men were taller (183.2 ± 1.9 vs. 171.1 ± 1.6 cm), and had a greater weight than female athletes (80.0 ± 2.0 vs. 67.6 ± 2.4 kg). There were not enough samples in each sporting discipline to compare age, weight and height measurements across sport (Table 2-6).

Sporting Category	Sport	Frequency
Racket Sports	Badminton	1
Athletics	Athletics, Triathlon	2
Martial arts	Jujitsu club, Taekwondo	3
Not known	Not known	3
Racket Sports	Squash	8
Team Ball Sports	Football, Hockey, Netball, Rugby	11
Water sports	Canoeing, Rowing, Swimming	11
Total		**39**

TABLE 2-6 – FREQUENCY COUNTS OF VOLUNTEERS COLLECTED IN THE GH-2004 LONGITUDINAL STUDY FOR EACH SPORTING CATEGORY.

2.1.5.2 STUDY DESIGN

Each athlete attended the Southampton Wellcome Trust Clinical Research Facility on four occasions over a four-month period for a blood sample. In addition, on one occasion, body composition was measured by anthropometry and bioimpedance (described in more detail in section 2.3) Each subject underwent a maximal exercise test on a treadmill to ascertain the effect of exercise on the GH2000 testing score.

During sampling, volunteers were seared and blood was drawn from a vein in the antecubital fossa. Samples were left to clot into two 5ml bottles containing SST clot activator gel, and were centrifuged within 12 hours of being collected at 1500 rpm in a portable centrifuge. Two-millilitre aliquots of serum were then transferred by pipette into small storage tubes (Cryotubes) and immediately placed on dry ice to storage at -80 C at the GH-2004 centre (University of Southampton). Samples were transferred on dry ice to the WADA accredited laboratory at King's College in London.

2.1.6 ITALIAN LONGITUDINAL STUDY –ELITE ATHLETES

2.1.6.1 SUBJECTS

The aim of this study was to assess the effects of the sporting season (entailing periods of training, competition, recovery, resting) on GH dependent variables in male and female elite athletes from different sporting disciplines.

IGF-I and P-III-P had been previously measured in a longitudinal follow-up study of 25 male and 22 female elite Italian athletes from 9 different sporting disciplines (Sartorio et al. 2006; Sartorio et al. 2004a). The mean age of the men and women was 22.6 ± 0.2 yrs and 22.5 ± 0.2 yrs respectively, 18 -34 yrs. Males had an average weight of 71 ± 8kg (55-83.4 kg) and average height of 179±7cm (168-192cm). Females had an average weight of 56 ± 6kg (41-67kg) and average height of 166±6cm (150-175cm). Volunteers of this study were recruited from either national or regional events. There were 3 rowers, 5 swimmers, 7 alpine skiers, 3 soccer players, 7 middle distance runners, 14 sprinters, 4 triathletes, 1 road walker and 3 cyclists. Volunteers had regularly trained at least 2-3 hours a day for 4-5 days a week during the years preceding the present study and they had not taken medications, apart from oral contraceptives.

2.1.6.2 STUDY DESIGN

A full physical examination and a preliminary screening assessment were devised to detect risk factors that might impede participation in the study. Written informed consent was obtained from every athlete. Approval was obtained by the Ethical Committee of the Italian Institute for Auxology, Milan, Italy.

All athletes were followed during a 6 month period of their period of their sporting season (entailing periods of training, competition, recovery, resting), with blood samples being taken every two months. Each athlete had blood samples collected at the same hour in the morning (10:00 – 12:00 h) or in the afternoon (17:00 – 19:00 h). Basal blood samples were drawn after the subjects had rested quietly for 20-30

minutes in a semirecumbent position. No relevant injuries were recorded during the investigation period. This study is described in more detail in the publication by Sartorio et al (Sartorio et al. 2006).

2.1.7 GH-2000 PLACEBO CONTROLLED DOUBLE BLIND STUDY– AMATEUR ATHLETES

2.1.7.1 SUBJECTS

The aim of this study was to evaluate the effects of GH administration on GH dependent biomarkers in healthy volunteers.

In total, 99 healthy amateur athletes were recruited in this study; 50 male and 49 female, from 4 different countries: Denmark (n=31), Italy (n=27), Sweden (n=30), and the United Kingdom (n=11). Volunteers recruited in this study were aged between 18–40 years old. No remarkable medical history was found in any of the recruited volunteers and none of them received any medications. All volunteers exercised at least 2 times per week for at least 1 year.

Pregnancy tests were performed on all female volunteers to ensure that they were not pregnant prior to entry and were confirmed to be using safe contraception. Nine women used oral contraceptives (1 in the placebo group, 3 in the 0.1 IU/kg/day GH group, and 5 in the 0.2 IU/kg/day group), and the remaining females used barrier methods. Treatment with GH/placebo started randomly with respect to the menstrual cycle. In each country, the protocol was approved by the ethical committee system and the national health authorities.

Oral and written consents were obtained from each subject in accordance with the principles stated in the Declaration of Helsinki. Further details for this study can be found in the publication by Dall et al (Dall et al. 2000).

2.1.7.2 STUDY DESIGN

The GH-2000 research group undertook a multi-centre, randomised, double blind, placebo controlled study, which has been reported previously (Longobardi et al. 2000; Dall et al. 2000). The study involved three treatment arms:

- Dose of 0.1 IU/kg/day (GH 0.1; maximum dose, 9.5 IU/day; n=30; 15 women and 15 men)
- Dose of 0.2 IU/kg/day GH (GH 0.2; maximum dose, 19 IU/day; n=29; 15 women and 14 men),
- Placebo (PLA; n=40; 19 women and 21 men).

GH (Genotropin, Pharmacia & Upjohn, Inc., Stockholm, Sweden; or Norditropin, Novo Nordisk, Copenhagen, Denmark) and placebo were administered as daily sc self-injections in the evening. To minimize side effects, only 50% of the target dose was given during the first week. In case of side effects, the dosage was reduced by 50%, and treatment was discontinued if the complaints persisted for more than 1 week. Compliance was monitored by collection of vials and reports about missing injections.

At baseline, the subjects attended the hospital for 1 day for blood sampling, interview and physical examination. Female subjects underwent a pregnancy test. Blood samples were subsequently collected at the end of each week during the treatment phase (days 7, 14, 21, and 28) and on days 30, 33, 42, and 84 in the washout period.

	Treatment		
	Placebo	**GH 0.1 IU/kg/day**	**GH 0.2 IU/kg/day**
No. of Subjects	40	29	30
Age (years)	25.4 ± 4.5	25.6 ± 4.2	25.8 ± 3.3
Sex	21 Males, 19 Females	15 Males, 14 Females	15 Males, 15 Females
BMI (kg/m²)	22.7 ± 3.2	23.1 ± 2.7	22.5 ± 2

TABLE 2-7 - BASELINE CHARACTERISTICS OF THE SUBJECTS (N = 99) PARTICIPATE IN THE GH-2000 DOUBLE BLIND STUDY

2.1.8 KREISCHA PLACEBO CONTROLLED DOUBLE BLIND STUDY– AMATEUR ATHLETES

The aim of this GH administration study involving non-competitive athletes was to investigate selected serum variables from different processes affected by hGH and identify the ones which would provide best discrimination amongst the GH treated group and Placebo. To do this, a number of biological parameters were measured in human serum in view of their response to GH administration and their intra- and inter-individual variations. The objective of the original study was to develop a preliminary discriminant function for distinguishing of GH- and placebo-treated subjects. The biological markers considered in the original study (Kniess et al. 2003) were IGF-I, IGFBP-3, N-terminal propeptides of type I and III procollagen (PINP, PIIINP), osteocalcin and leptin.

This study is used in this thesis to provide an independent set of data on which the GH-2000 detection methodology can be validated. Furthermore, the proposed methodology as proposed by Kniess et al using the results from this study is validated against the results collected from the GH-2000 double blind study.

2.1.8.1 SUBJECTS

Fifteen healthy male, non-competitive athletes (age range 21–33 years, mean 24 years, body mass index 23–27 kg/m2) were assigned randomly to either placebo (n=5) or 0.06 IU/kg body weight /day rhGH (n=10).

2.1.8.2 STUDY DESIGN

Treatment was administrated daily by medical staff to the athletes by subcutaneous injection for 14 days. Blood samples were taken for analysis of IGF-I, IGFBP-3 and P-III-P fasting at baseline (day 0) and on days 2, 4, 6, 8, 10, 12, 15 and 18 at 6pm (Kniess et al. 2003).

All subjects gave written informed consent and the studies were conducted in accordance with the ethical principles of the Declaration of Helsinki. The study protocols were approved by the Ethics Committee of the centres involved.

2.2 ANALYTICAL METHODS

All studies undertaken by the GH-2000 group were analysed using the same analytical methods. All projects undertaken by the GH-2004 research group were analysed using the same P-III-P assay, however two IGF-I assays were used to analyse samples.

2.2.1 GH-2000 RESEARCH STUDIES

All serum samples were stored at -80 C, and all samples from one subject were analyzed in the same run.

Serum IGF-I was measured by RIA using a monoclonal antibody after acid-ethanol extraction with within-assay CVs of 6.6%, 4.4%, and 2.3%, and between-assay CVs of 9.7%, 7.0%, and 4.6% at 104, 281, and 1324 ng/ml. Results were compared with previously published population-based reference data from the same laboratory. The analyses used on both occasions were carried out using the same technique and the same antisera.

Serum PIIIP was measured by a two-stage sandwich RIA; total assay CVs (within- plus between assay) were 9.1%, 5.7%, and 6.8% at 0.62, 0.95, and 1.18 mg/liter.

The results from these studies have been previously published (Longobardi et al. 2000; Dall et al. 2000; Healy et al. 2005).

IGFBP-3 was measured using an in-house RIA and polyclonal antibodies. The intra-assay CVs were 3.4%, 3.3% and 3.4% at serum concentrations of 60, 245 and 502 nmol/L, respectively.

2.2.2 GH-2004 RESEARCH STUDIES

All samples were analysed in duplicate at the Drug Control Centre, King's College London, UK by Christiaan Bartlett and under the directorship of Professor David Cowan.

For the GH-2004 cross-sectional study, serum IGF-I was measured using the IGF-I Nichols immunoradiometric (IRMA) assay. This is a non-extraction assay with the addition of acid to disassociate IGF-I from its binding proteins and subsequent flooding with IGF-II to prevent reattachment. The assay utilises 2 region-restricted polyclonal antibodies: an antibody to the C terminal amino acids 62 -70 is biotinylated for capture and a radio-labelled antibody to the amino acid sequences 1-23, and 42-61 for detection purposes. During an incubation step, the IGF-I in the sample forms a sandwich complex with the antibodies on an avidin-coated bead. After washing, the radioactivity bound to the solid phase is measured in a gamma counter, the amount of radioactivity present being directly proportional to the amount of IGF-I in the sample. The reported intra-assay coefficient of variation (CV) at concentrations of 61 ng/mL, 292.5 ng/mL and 547.9 ng/mL are 4.6%, 3.3% and 4.1% respectively with an inter assay CV of 15.8%, 10.3%, and 9.3% at concentrations of 60.1 ng/mL, 312.1 ng/mL and 594.3 ng/mL respectively.

At the planning stages of the GH-2004 research, it was planned that the Nichols IRMA IGF-I immunoassay was going to be used for all parts of this study. For the GH-2004 Cross-sectional study this immunoassay was used however for all other GH-2004 studies (GH-2004 Injury study and the GH-2004 Longitudinal study) serum IGF-I serum concentration was measured by the DSL-5600 ACTIVE® IGF-I IRMA manufactured by DSL Laboratories (Diagnostic Systems Laboratories, Inc., Webster, TX). This was because F.D.A. stopped Nichols from manufacturing for unrelated health and safety reasons. Hence, it was not possible to carry on the measurement of IGF-I using the Nichols IRMA assay for the rest of the research.

The DSL-5600 IRMA procedure employs a two-site immunoradiometric assay principle after an extraction step in which IGF-I is separated from its serum binding proteins. The intra-assay precision of the assay as reported by the manufacturer was 3.4%, 3.0% and 1.5% at 9.4, 55.4 and 263.6 ng/mL respectively. The inter-assay precision of the assay was 8.2%, 1.5% and 3.7% at 0.9, 0.8 and 9.6 ng/mL respectively. The manufacturer reported that the IGF-I Standards & Controls were calibrated to

the World Health Organization International Reference Reagent for IGF-I (code 87/518) and performance was verified by immunoassay.

For all GH-2004 studies serum P-III-P was measured by a two-stage sandwich RIA (CIS Biointernational; Oris Industries, Gif-Sur-Yvette Cedex, France). This is a two-stage sandwich radioimmunoassay based on the formation of a complex between solid-phase anti-PIIIP monoclonal mouse antibodies. This assay has a lower limit of detection of 0.1 units of PIIIP/mL and upper limit of 14 units of PIIIP/mL The manufacturer's reported intra assay CV; at 0.8, 1.5, 4.0 U/mL is 2.9%, 2.9% and 4.0% respectively. The inter-assay CV at 0.25, 1.5 and 5.6 U/mL is 11.3%, 7.8%, and 9.3% respectively. Sample pre-treatment was not needed. The QC samples supplied by the manufacturer were used (acceptable range 1.83-2.47 U/mL).

2.2.3 ITALIAN LONGITUDINAL STUDY

IGF-I concentrations were determined by using a commercial immunoassay kit (Mediagnost Gmbh, Tübingen, Germany). The intra- and inter-assay coefficients of variation were 3·5% and 7% for IGF-I, respectively. P-III-P levels were determined using the Orion Diagnostica RIA kits (Oy, Espoo, Finland). Intra- and inter-assay coefficients of variation were 4·3% and 5·3% for PIIINP; the sensitivity was 0·2g/l for P-III-P.

2.2.4 KREISCHA GH ADMINISTRATION PLACEBO CONTROLLED STUDY

Serum IGF-I was analysed after hydrochloric acid-ethanol extraction by competitive fluorescence immunoassay (Kniess et al. 2003). The intra-assay coefficients of variation (CVs) of the IGF-I assay were 19.2%, 8.7% and 6.6% at serum concentrations of 26, 149 and 714 ng/mL respectively.

P-III-P was analysed by radioimmunoassay (RIA) using rabbit antisera (Orion Diagnostica, Espoo, Finland). The intra-assay CVs were 6.4%, 4.4% and 5.1% at PIIINP concentrations of 4.4, 5.9 and 8.8 ng/mL.

The IGFBP-3 concentrations were determined by using a commercial sandwich enzyme-linked immunosorbent assay (ELISA) with two polyclonal IGFBP-3 goat antibodies (DSL, Webster, Texas, USA) (intra-assay CVs 9.6%, 9.5% and 7.3% at serum concentrations of 4.6, 27.4 and 74.4 ng/mL).

2.3 ANTHROPOMETRY MEASUREMENTS

Weight measurements were measured to the nearest 25g using electronic scales. Standing height was measured using Holtain stadiometer with the patient's head in the Frankfurt Plane accurate to 1cm.

Body fat measurements (accurate to 0.1cm) were taken at 4 sites in accordance with accepted body composition equation. These 4 sites were located at bony landmarks where the skin can be easily lifted from underlying muscle; at the bicep, tricep, subscapularis, and suprailiac areas.

Mid upper arm circumference was taken at half way between acromium process and olecranon.

Waist measurements were taken at minimum circumference of mid-section of the body. Hip measurements were at the maximum circumference of the mid-section of the body.

For the GH-2004 longitudinal study, further body composition measurements were taken at the Welcome Trust Clinical Research facility (WT-CRF). Body density was calculated using the equation density=mass/volume. Volume was measured in the Bod Pod, a device used to measure volume as calculated by air displacement. Once density was known, body composition was calculated using scientifically derived equations.

In addition, for the GH-2004 longitudinal study bioimpedance analysis was carried out to assess the body composition by means of electrical conductivity of lean tissue (bone, muscle, water) being much greater than that of adipose tissue.

2.4 STATISTICAL METHODS

The statistical analysis undertaken in this thesis had two purposes:

1. Derive assay conversion methodology, which will allow comparisons of IGF-I and P-III-P concentration between 9 research trials considered in this thesis.
2. To examine the four key aims of this thesis which are:
    a. Determine the effect of ethnicity on serum concentrations of IGF-I, P-III-P and the GH detection method.
    b. Determine the effect of sporting injury on serum concentrations of IGF-I, P-III-P and the GH detection method.
    c. Determine the magnitude of the physiological intra-individual variability of IGF-I, P-III-P and the impact of this on the GH detection method.
    d. Validate the proposed GH detection method on independent data collected from a GH administration study.

2.4.1 CALCULATING ASSAY ADJUSTMENTS

The major challenge for the marker approach is to ensure harmonisation between the different assays used to measure P-III-P and IGF-I. Although there are no established methodologies to adjust measured P-III-P and IGF-I from one assay to another, this problem is not insoluble as a similar problem has arisen for many assays including glycosylated haemoglobin (Goodall 2005). The establishment of international reference preparation and quality control schemes has led to harmonisation of assays within the clinical arena. As there is no international reference preparation for P-III-P, an alternative is to use adjustment factors based on the values measured for normal individuals by different assays.

It is necessary to use statistical analysis to produce an adjustment factor to allow reliable and reproducible comparisons of IGF-I concentrations measured by the different IGF-I assays (Dunn 2004; Strike 1991).

A decision was made to define the GH-2000 IGF-I measurement scale as the "gold standard" to maintain consistency across previously published research by the GH-2000 group. Therefore the measured GH-2004 IGF-I concentrations were transformed to the scales used by the GH-2000 IGF-I assay. It is useful to note though that using the same statistical methodology it is as easily possible to convert the GH-2000 results to any alternative assay scales.

In order to obtain a mathematical conversion factor between two IGF-I assays, the assumption of constant proportional bias (i.e. linear relationship) between the two assay methodologies was made. This assumption was based on the fact that both assays would be measuring the same antibody. This assumption was confirmed by assessing the significance of the intercept in each of the regression models fitted. For all assays considered, it was found that there was no significant evidence to suggest that this assumption would not be valid. Furthermore, the assumption of no constant bias between the two assays was also made (the regression line crosses the y-axis at zero). This allowed the calculation of regression based assay conversion factors.

In the case where paired data was available - i.e. same sample assayed by two different assay methodologies, it was possible to calculate a conversion factor using a regression model and assuming no constant bias between assay methodologies. In the situation where no paired data is available, then an indirect assay conversion methodology was used as described in the section 2.5.2.

The withdrawal of the original Nichols RIA and subsequent IRMA used in this study meant that all the samples in the GH-2000 and GH-2004 studies could not be measured using the same assay. In fact, it was necessary to use a third assay to derive a conversion factor. By using a direct comparison between assays, we were able to do this with few assumptions. The main assumption was that there was a linear proportional relationship between the two assays and the regression line passed through zero on the x and y-axis. It also assumes that there is no degradation in the freezer with time. This is a reasonable assumption because previous studies have samples stored at -80 C to be very stable (Evans et al. 2001). The weakness of the

approach is that by necessity, its uses a relatively small number of samples to derive the conversion adjustment.

2.5 ASSUMPTION OF NO CONSTANT BIAS BETWEEN ASSAYS

It is known from the laboratory validation analysis undertaken by Professor David Cowan and Mr Christiaan Bartlett at King's College London for the IGF-I assay, that the main reason for differences in the assays numbers is caused by differences in the potency of the standards used. When international reference preparations were used in validation work, much of the difference disappeared.

The figure below shows that regression line for the Nichols RIA against the DSL assay together with the line of unity with no assumptions made for the size of the intercept (size of the constant bias).

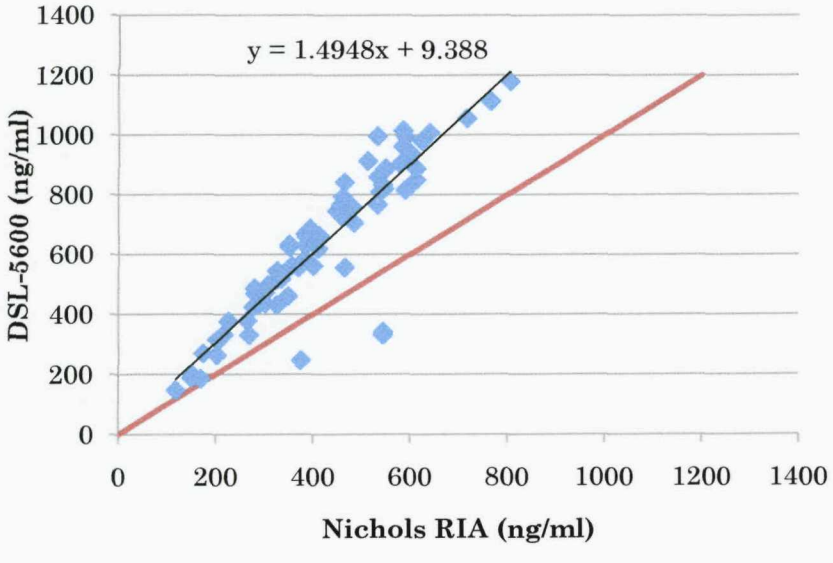

FIGURE 2-1 COMPARISON OF PAIRED IGF-I MEASUREMENTS MEASURED BY BOTH THE DSL-5 600 AND THE NICHOLS IRMA METHOD WITH A FITTED REGRESSION LINE WITH NO FIXED INTERCEPT. THE RED LINE REPRESENTS THE LINE OF EQUALITY.

Assuming a constant proportional bias the regression equation is y = 1.5146x. When no constant bias is assumed the regression equation is y = 1.4948 x + 9.388. As it can be seen the value for the intercept is small and not significantly different from zero.

The next figure shows that regression line for the Nichols IRMA against the DSL assay together with the line of unity.

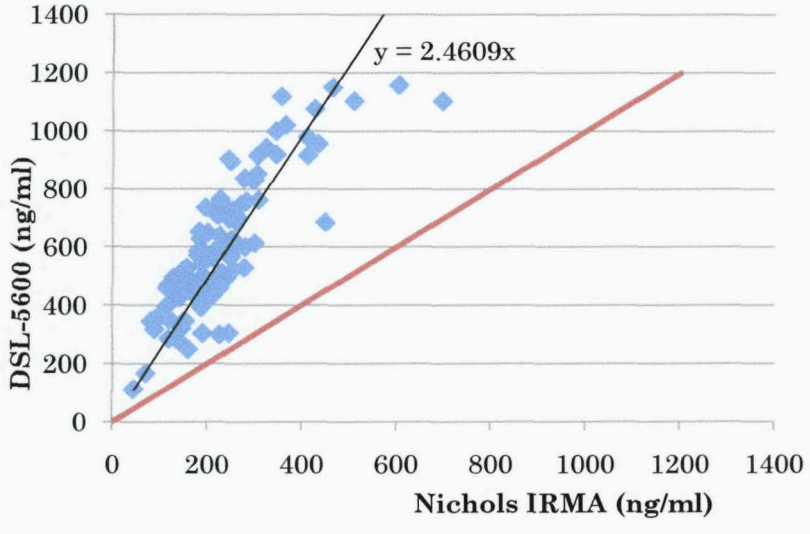

FIGURE 2-2 COMPARISON OF PAIRED IGF-I MEASUREMENTS MEASURED BY BOTH THE DSL-5 600 AND THE NICHOLS IRMA METHOD WITH THE ASSUMPTION OF CONSTANT BIAS=0. THE RED LINE REPRESENTS THE LINE OF EQUALITY.

Apart from the inevitable scatter, it is clear that there is non-linearity when DSL reads above 1000ng/ml (equivalent to approximately 400ng/ml on the Nichols assay). If samples above 1000ng/ml (DSL) are ignored the data appear linear and when no constant bias is assumed, the intercept is small and not significantly different from zero.

Therefore, given this evidence, it can be seen that the approach of proportional bias and no constant bias is justified.

2.5.1 CALCULATING ASSAY CONVERSIONS USING PAIRED DATA

The following gives details of the statistical methodology used to convert one assay to another when paired data were available. Each sample has been assayed by the two methods (assay A and assay B) and therefore we have paired data. Paired data was available for the analysis, where each sample was assayed with the GH-2000 assay methodology but also the GH-2004 methodology.

A number of n samples was used to perform this analysis in the format of (x_i , y_i), $i = 1, 2, \ldots, n$ on n subjects. The x values are from Assay A ones, and the y values are corresponding Assay B ones.

The linear regression model is $y_i = \alpha + \beta x_i + \varepsilon_i$, $i = 1, 2, \ldots, n$, where $\varepsilon_i \sim N(0,\sigma^2)$. It was assumed that proportional model is suitable for calculating a valid conversion method, $\alpha = 0$. The first aim was therefore to estimate β but there were important secondary aims; it was important to check whether the proportional model is realistic and whether the error terms ε_i have constant variance.

If the model $y_i = \alpha + \beta x_i + \varepsilon_i$, with $\varepsilon_i \sim N(0,\sigma^2)$ is correct, then the LS estimator of β is $\beta=\Sigma x_i\ y_i\ /\Sigma x_i^2$ and by standard theory this will also be the maximum likelihood estimate.

2.5.1.1 CALCULATING A CONVERSION FACTOR FOR GH-2004 IGF-I ASSAYS

There were 73 serum samples collected by the GH-2004 research group which were analysed using the Nichols IRMA assay and were subsequently also re-analysed using a second assay; DSL-5600 ACTIVE® IGF-I IRMA manufactured by DSL Laboratories

(Diagnostic Systems Laboratories, Inc., Webster, TX). Using these paired observations, a conversion was estimated as follows:

DSL = 2.461 x Nichols IRMA

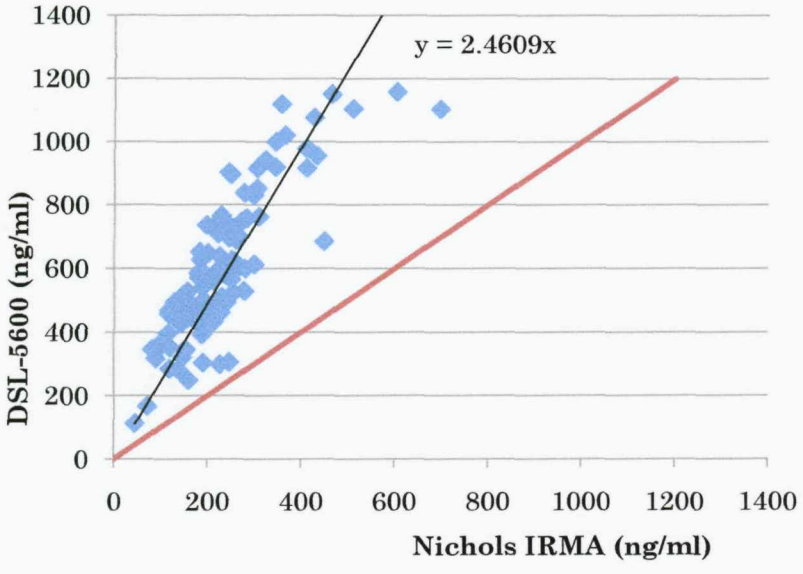

FIGURE 2-3 – COMPARISON OF PAIRED IGF-I MEASUREMENTS MEASURED BY BOTH THE DSL-5600 AND THE NICHOLS IRMA METHOD. THE RED LINE REPRESENTS THE LINE OF EQUALITY.

In order to obtain a mathematical conversion factor between the two Nichols assays, the assumption was made that there was linear relation (proportional bias) between the two assay methodologies. It was also assumed that the regression line would pass through zero (no constant bias), as assays are measuring the same substance and therefore the regression line should pass through zero. The regression with some constant bias was also examined. This did not significantly change the conversion equation, as the value for intercept was small and not significantly different from zero. This method allowed the calculation of regression based assay conversion factors.

In addition, samples that were previously analysed using the Nichols RIA method were re-analysed using the DSL-5600 IRMA assay. Using these paired observations, a conversion was estimated as follows:

DSL = 1.515 x Nichols RIA

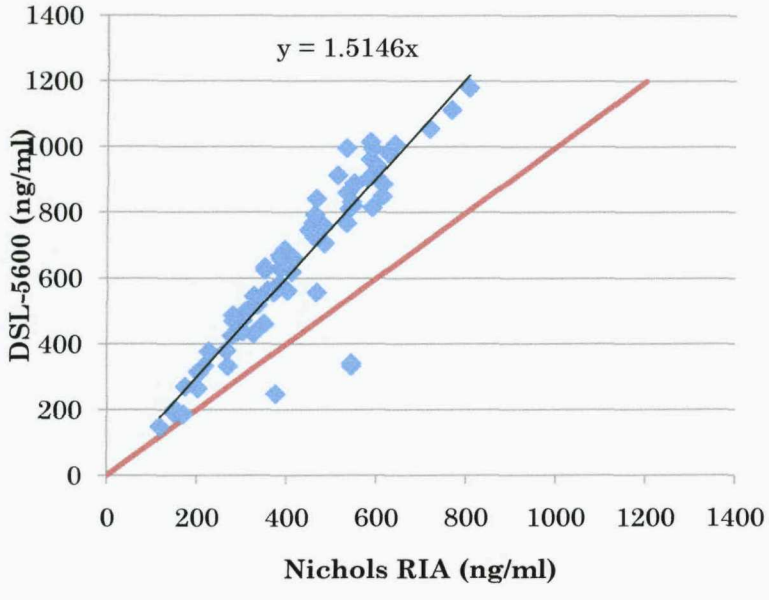

FIGURE 2-4 - COMPARISON OF PAIRED IGF-I MEASUREMENTS MEASURED BY BOTH THE DSL-5 600 AND THE NICHOLS RIA METHOD. THE RED LINE REPRESENTS THE LINE OF EQUALITY.

Using mathematical substitution, a new conversion factor was then derived to convert Nichols IRMA measured concentrations to match the range of concentrations used by the Nichols RIA assay. The derived conversion formula was then estimated as follows:

Nichols 2000=Nichols 2004 x 1.6247.

2.5.1.2 ASSAY CONVERSION OF P-III-P ASSAY

The P-III-P assay for both GH-2000 and GH-2004 studies was manufactured by CIS BioInternational. Slight modifications were made to the assay to improve its

sensitivity and these also lead to improved intra assay variability (CIS BioInternational 2007). Aside from these changes, no significant adjustments were made to the assay structure and methodology and Quality Control data from the manufacturers showed no sign of assay drift over time and so it was therefore not considered necessary to carry out any assay adjustments on the P-III-P measured concentration values when comparing the results from the two cross-sectional studies.

2.5.2 CALCULATING ASSAY CONVERSIONS WITH NO PAIRED DATA AVAILABLE

During the period of this analysis it was not possibly to obtain adequate serum samples to perform paired data comparisons. This meant that it was not possible to re-assay GH-2000 samples with the other assays.

2.5.2.1 ASSAY CONVERSIONS TO USE WITH KREISCHA DATA

The GH-2000 and Kreischa research groups used different assay methodology to measure concentrations of IGF-I, P-III-P and IGFBP-3. Furthermore, the assay used to measure P-III-P concentrations reported in different measurement units; the CIS Bio-International assay (GH-2000) reports in arbitrary units (U/ml) while the Orion RIA-Gnost assay (Kreischa) reports in ng/ml (of a proprietary standard). In order to test the discriminant functions on the alternate dataset, it was therefore necessary to convert measurements of each research group to the assay scales used by the other research group.

While an adjustment equation has been published to convert between the Orion and CIS P-III-P assay (Abellan et al. 2005), this did not work satisfactorily with our data and there are no established methods to do this for IGF-I. A novel method for assay adjustments was developed based on the data available from the placebo and pre-treatment samples from each study, with the assumption that there would be no difference between the two populations of volunteers. Both groups were European Caucasian men with similar ages ($p=0.89$) and BMI ($p=0.78$). This therefore enabled

the calculation of the assay adjustment without the need for an age adjustment. The method used to calculate the adjustment factor for each marker, was by aligning the means and the standard deviations of the data collected by each group (for each marker).

Below are the details of the statistical method developed to align the results from the GH-2000 and Kreischa studies. It is based on the assumption that the means and standard deviations of the P-III-P and IGF-I values of the 'untreated' samples (baseline samples from active and placebo group and all the other placebo samples) of the GH-2000 and Kreischa studies are not different. This is considered a reasonable assumption, as both groups are predominantly White European men of similar ages and BMI. The method aligns the means and standard deviations and then rescales the values to a common scale.

If X_{GH} represents an observation from the GH-2000 trial, and let X_{KR} represent an observation from the Kreischa trial. Then the equation used to convert Kreischa values to the GH-2000 assay scales is:

$$X_{REV\_KR} = \left(\frac{X_{KR} - \overline{X_{KR}}}{s_{KR}}\right) \times s_{GH} + \overline{X_{GH}}$$

where X_{REV_KR} is the revised Kreischa value, on the assay scales used by the GH-2000 group. Similarly, the equivalent equation used to convert GH-2000 values to the Kreischa assay scales is therefore:

$$X_{REV\_GH} = \left(\frac{X_{GH} - \overline{X_{GH}}}{s_{GH}}\right) \times s_{KR} + \overline{X_{KR}}$$

Where X_{REV_GH} is the revised GH-2000 value on the assay scales used by the Kreischa group.

In these expressions, σ_{GH} denotes the sample standard deviation for the measurements collected by the GH-2000 group and similarly σ_{KR} denotes the sample standard deviation for the measurements collected by the Kreischa group. Also, $\overline{X_{GH}}$

denotes the sample mean for the measurements collected by the GH-2000 group and similarly $\overline{X_{KR}}$ denotes the sample mean for the measurements collected by the Kreischa group.

The resulting adjustment equations used to convert data collected by the Kreischa group to the scales of assays used by the GH-2000 group were:

For IGF-I: $X_{REV\_KR} = 1.491 \times X_{KR} - 31.552$

For P-III-P: $X_{REV\_KR} = 0.1214 \times X_{KR} + 0.00067$

For IGFBP-3: $X_{REV\_KR} = 0.1335 \times X_{KR} - 1.638$

where X_{REV_KR} is the revised Kreischa value. The corresponding adjustment factors to convert data collected by the GH-2000 group to the assay scale used by the Kreischa group assays were:

For IGF-I: $X_{REV\_GH} = 0.670 \times X_{GH} + 21.15$

For P-III-P: $X_{REV\_GH} = 8.234 \times X_{GH} - 0.044$

For IGFBP-3: $X_{REV\_GH} = 7.490 \times X_{GH} + 12.273$

where X_{REV_GH} is the revised Kreischa value.

For all three assay comparisons, the significance of the intercept values was assessed, and as seen from the magnitude of these parameters, the magnitude of the intercept was not significant.

For the purposes of validating the discriminant function developed by the GH-2000 group, using data collected by the Kreischa group, the revised values from the Kreischa group, $X_{REV_{KR}}$, of each marker were used.

Similarly, in order to validate the discriminant function developed by the Kreischa group using data collected by the GH-2000 group, the revised values X_{REV_GH} of each marker were used.

Application of the estimated assay conversion factors showed that differences between the assay methodologies had been eliminated. Specifically, for the IGF-I assay methodology following correction it was shown that there was no significance difference between the serum IGF-I concentrations measured in the two studies at baseline or in the placebo treated group ($p = 0.96$). Similarly, for the P-III-P assay methodology following assay correction there was also no significant differences evident in serum P-III-P concentrations between the two studies ($p = 0.19$). For the IGFBP-3 assay methodology, similar conclusions were made ($p = 0.24$).

Application of the assay conversion factors showed that differences between the assay methodologies had been eliminated. Specifically for IGF-I there was no differences between the two studies at baseline or in the placebo treated group ($p=0.96$). Similarly for P-III-P, following assay correction there were also no significant differences between the two studies ($p=0.19$), likewise for IGFBP-3 ($p=0.24$).

2.5.2.2 ASSAY CONVERSIONS TO USE WITH THE ITALIAN LONGITUDINAL STUDY

There was no paired data available from the Italian Longitudinal study, therefore the indirect method for calculating assay adjustments was used as described in the previous section.

An assay adjustment was necessary for both IGF-I and P-III-P measurements measured in the Italian study. The conversion equations calculated were as following:

Adjusted IGF-I value to GH-2000 scale = Italian IGF-I value (ng/ml) x 90.15+25.63

Adjusted P-III-P value to GH-2000 scale (U/ml) = Italian P-III-P value (mg/l) x 0.1214+0.00067 (as calculated in the previous section).

2.5.3 DISCUSSION ON ASSAY ADJUSTMENTS

There is considerable published evidence to show that assay conversion is possible to allow comparison of IGF-I and P-III-P with historical samples. A recent study undertaken by Abellan et al (Abellan et al. 2005) has indicated that the calculation of statistical assay adjustment is possible and can produce reliable results.

In addition, Ivan et al (Ivan et al. 2005) compared the automated Nichols Advantage chemiluminescence IGF-I immunoassay and an enzyme-linked immunosorbent IGF-I DSL-10-2800 assay. The study included 95 serum samples from 88 patients, adults as well as children, with different or no endocrine disorders, and acromegalic as well as growth hormone-deficient patients. IGF-I measurements were performed with both methods. The results have shown a very high correlation between the IGF-1 values obtained with the two assays ($r=0.971$, $p<0.0001$).

Rinaldi and colleagues (Rinaldi et al. 2005) compared IGF-I as measured by an in-house IGF-I immunoassay and commercially available DSL ELISA. Pearson's correlations between IGF-I concentrations in the original study and DSL ELISA were very high [$r = 0.92$; 95% CI, 0.90-0.94].

Massart and Poirier studied the performances of four IGF-I immunoassays in the follow-up of acromegaly it allows the use of assays (Massart et al. 2006). The assays were radioimmunoassays from Immunotech, DiaSorin and Schering Laboratories and the chemiluminescent automated immunoassay (Advantage) from Nichols. Normal age- and gender-matched subjects constituted the reference population for two assays. Deming regression and Bland-Altman analyses showed a high correlation with the four methods tested.

These papers all indicate that it is possible to convert between assays. This is very encouraging, as it allows the use of samples measured with assay methodology other than the one employed by the GH-2000 study when developing the GH-2000 detection score.

There have been suggestions that the assay conversion methodology is not acceptable because there was scatter around regression line. As immunoassays are biological systems, there will always be variation. Indeed even if two identical samples are measured by the same assay there will be variation and therefore this argument is not sustainable unless the use of immunoassays is rejected outright. It has been shown that the variability around the regression line is no more than would be expected from assay variability. As such this can be built into the concept of measurement uncertainty.

2.6 STATISTICAL ANALYSIS TO EXAMINE KEY AIMS OF THIS THESIS

All statistical analysis was performed using the SAS software™ (SAS Institute Inc., Cary, NC, USA) or Minitab Version 15. The analysis was carried out using the log-transformed values of the IGF-I (ng/ml) and P-III-P (U/ml) concentrations.

2.6.1 ASSESSING THE EFFECT OF ETHNICITY

The aim of this study was to examine serum IGF-I and P-III-P concentrations in blood samples obtained from elite athletes of different ethnic groups within 2 hours of competing at a national or international event. These results have been compared with the original results obtained by the GH-2000 group to assess whether the reference ranges obtained for white European athletes are applicable for those of other ethnic backgrounds (Healy et al. 2005).

The model used was: $y = \alpha + \beta x + \varepsilon$ where y represents the logarithm of a marker, and x represents the reciprocal of age and ε represents random error. Prediction intervals (reference ranges) were constructed for each marker by using this model. Under the normal assumption, where the random error is normally distributed with constant variance then be normally distributed with constant variance, regression analysis was used for each combination of marker and gender providing least squares estimates for the model variables; a and b of the intercept α and β slope of the relationship between y and x together with an estimate of the residual standard deviation.

The formula to calculate the prediction intervals, using the variable estimates derived by the regression analysis is:

$$a + bx \pm ts\sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}}}$$

where n is the number of observations, $\bar{x}$ is the mean of the observed values of x, $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$ which is the residual standard deviation and t is the appropriate percentage point of the t distribution.

Due to the large sample size (537 male and 276 female volunteers), under the central limit theorem, t values can be replaced by the equivalent normal value 2.576 for a 99% range. Furthermore, due to the large sample size, the square-root part of this formula, $\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{S_{xx}}}$, can be approximated to the value of 1. This is because $\frac{1}{n}$ is small due to the large sample size, compared to the first term of this part. Similarly, the final term is usually small, as the numerator is a single square, and the denominator is the sum of n such squares, thus making the value of this generally small.

Using these approximations, the approximated lower (L) and upper (U) prediction limits for an observation on y at a given value of x are thus provided as: $L = a + bx - ts$ and $U = a + bx + ts$, where t is the appropriate percentage point of the t distribution. Because the prediction intervals are expressed on a logarithmic scale, the limits for a marker in natural units are e^L and e^U.

In order to assess whether ethnicity affected the serum concentrations of IGF-I and P-III-P, analysis of variance was performed for each gender after adjustment for the effect of age and assay methodology. This method enabled calculation of the estimated mean difference in the marker concentrations for each ethnic group, compared to those observed in white-European subjects. All statistical comparisons were two-tailed. $P < 0.05$ was considered statistically significant.

In order to assess the implication of any variation for the application of the GH-2000 formulae to individuals of non-white European background, the formulae scores were

calculated for each individual. For each formula, IGF-I were assay adjusted whereas P-III-P measured were used as measured.

Graphical displays of these values for each gender were plotted, to illustrate how the measured concentrations of IGF-I and P-III-P in each ethnic group would vary. Results were given with ±1 Standard Error of the Mean (SEM) estimates.

2.6.2 ASSESSING THE EFFECT OF INTRA-INDIVIDUAL VARIABILITY

The statistical analysis performed in this study concerned four main areas.

2.6.2.1 QUANTIFYING INTRA-INDIVIDUAL VARIABILITY ON IGF-I AND P-III-P

Analysis of variance (ANOVA) analysis was performed to estimate the intra-individual variability components of the IGF-I and P-III-P concentrations for each population. Further analysis was carried out using ANOVA methodology to assess whether there were any significant differences between genders. All analysis carried out adjusted for differences in the biomarker concentration due to age differences between volunteers.

2.6.2.2 ASSESSING EFFECT OF INTRA-INDIVIDUAL VARIABILITY ON TESTING SCORES

Similarly, ANOVA methodology was performed to estimate the inter and intra-individual variability components of the formulae scores, calculated using the discriminant functions proposed by the GH-2000 group as means for detecting GH abuse in sports.

Graphical displays of these values for each gender were plotted, to present how the measured concentrations of IGF-I and P-III-P in vary for each individual. Analysis of variance techniques were used, to assess whether there was a significant difference in

the calculated function scores, between the ethnic groups, compared to observations collected from white-European volunteers.

Using the samples collected at resting conditions and immediately after exercise in the GH-2004 Longitudinal study paired t-test analysis was carried out in order to determine whether exercise had a significant effect on the intra-individual variability estimates.

2.6.3 ASSESSING THE EFFECT OF INJURY

The main statistical method used in this study was analysis of variance. The analysis was carried out separately for each of the injury types (bony or soft-tissue injury) and separately for each severity group (mild, moderate and severe). The aim of the analysis was to assess the significance of changes in the concentrations of the two biomarkers compared to baseline values. The analysis examined the distribution of the concentration of each biomarker for each of the six groups. The analysis adjusted for the age and gender of the study participants.

In order to assess whether an injury would significantly affect the score obtained from the detection formulae, and more importantly whether it could lead to the false accusation of an athlete for GH doping, the scores of the GH-2000 detection formulae were calculated for each participant at every time point.

Analysis of variance was undertaken for each of the six injured groups to assess whether these scores changed significantly over time. Furthermore, the intra-individual variability of these scores was quantified over the study period.

2.6.4 VALIDATION OF THE GH-2000 DETECTION METHOD

As discussed the GH-2000 project produced gender specific discriminant formulae involving the measured concentrations of IGF-I and P-III-P.

The Kreischa group recruited only male volunteers, therefore only the male version of the GH-2000 detection was considered for validation.

The Kreischa group developed a similar prediction formula that included IGF-I, P-III-P and IGFBP-3. The resulting discriminant function calculated was:

Kreischa test score = $-13.465 + 0.0272 \times$ IGF-I $+ 0.0398 \times$ IGFBP-3 $+ 1.367 \times$ P-III-P $- 0.00271 \times$ (IGF-I $\times$ P-III-P)

with the markers expressed in the actual measured scale.

The major differences between the two discriminant functions is the use of logarithmically transformed values by the GH-2000 function (to normalise distribution of data) and the use of IGFBP-3 as an additional marker and the presence of a product of IGF-I and P-III-P in the Kreischa formula.

The scores calculated by each research group have been standardised such that the placebo and baseline (pre-treatment) samples have a mean value of 0 and standard deviation 1.

These standardised scores of the two discriminant function equations were found to be normally distributed and therefore a score of 3.7 is equivalent to the 99.99% reference limit. The chance of a normal individual having a score of >3.7 is therefore 1 in 10,000. As a cut-off point has not yet been agreed by the World Anti-Doping Agency, the cut-off limits for the detection of GH usage for this study were arbitrarily set at 3.7 (Sonksen 2001), giving the chances of a false positive being less than 1:10,000.

THE SENSITIVITY OF THE FUNCTIONS ON A GIVEN DAY HAS BEEN DEFINED AS "THE NUMBER OF OBSERVATIONS PLACED OVER THE 3.7 CUT-OFF VALUE, OUT OF THE TOTAL NUMBER OF OBSERVATIONS FOR THE PARTICULAR DAY FROM VOLUNTEERS RECEIVING GH TREATMENT".

CHAPTER 3: AN ASSESMENT OF THE EFFECT OF ETHNICITY ON GH-DEPENDENT MARKERS

3.1 INTRODUCTION

This study involved the collection of post-competition samples from elite athletes competing at national and international level, predominantly from non-Caucasian ethnic backgrounds. The purpose of this study was to assess whether the physiological reference ranges of the GH-sensitive markers IGF-I and P-III-P, derived by the GH-2000 study using predominantly Caucasian volunteers, are valid for non-Caucasian volunteers. If it was proven that these physiological reference ranges were not applicable for non-Caucasian volunteers, suitable adjustments for ethnicity would be needed to ensure that athletes from non-Caucasian background are not unfairly adjudged by the marker method developed by the GH-2000 group.

Age, height and BMI varied widely between the sports, reflecting the different nature of each discipline.

3.2 METHODS

Organizers of 68 International and National events were contacted to obtain permission to recruit subjects at the event. However permission was obtained from only 12 events (see section 2.1.1 for more details).

Samples were collected from each consented individual within 2 hours of post-competition after subjects confirmed in writing that they had not taken performance-enhancing drugs prior to the venepuncture.

Gender, race, height (self-reported), weight (self-reported), age and sporting category were recorded for each study volunteer. Data collected on the ethnic group background from the GH-2004 volunteers were classified in one of five groups; White-European, Afro-Caribbean, Indo-Asian, Oriental and Other. The "Other" classification included volunteers from a mixed ethnic background and Arabic athletes. This was done because there were insufficient samples from each gender available for each combination of mixed ethnic background and included Arabic athletes ($n<5$).

All samples were kept on ice until centrifuged as IGF-I and P-III-P are stable for up to 5 days under these conditions. After centrifugation, serum was transferred into small storage tubes and immediately frozen (World Anti-Doping Agency 2003; World Anti-Doping Agency 2004a). These were placed on dry ice prior to transport to and storage at -80°C at the GH-2004 centre (University of Southampton) and were subsequently coded and transferred on dry ice to the WADA accredited laboratory at the Drug Control Centre, King's College in London for analysis of serum IGF-I and pro-collagen type III. All assay analysis was performed in duplicate.

Serum IGF-I was measured using the IGF-I Nichols immunoradiometric (IRMA) assay. Serum P-III-P was measured by a two-stage sandwich RIA (CIS Biointernational; Oris Industries, Gif-Sur-Yvette Cedex, France).

The methods employed to analyse IGF-I and P-III-P concentrations in the GH-2000 samples have been reported previously (Healy et al. 2005). Serum IGF-I was measured by a Nichols RIA using a monoclonal antibody after acid-ethanol extraction (17), with within-assay CVs of 6.6%, 4.4%, and 2.3%, and between-assay CVs of 9.7%, 7.0%, and 4.6% at 104, 281, and 1324 ng/ml. The P-III-P was measured by the CIS RIA described above.

Statistical Analysis

The results obtained during this study were compared with the previously published normative data by the GH-2000 project from 815 elite athlete volunteers(Healy et al. 2005). The data analysis was performed using the SAS software™ (SAS Institute Inc., Cary, NC, USA).

Adjusting for Assay differences

Due to the differences in the assay methodology used as previously discussed in section 2.2.2 statistical assay adjustments were necessary to allow direct comparisons of the measurements to be made across the GH-2000 and the GH-2004 studies. These

conversions have been given in sections 2.5.1.1 for IGF-I measurements and 2.5.1.2 for P-III-P measurements.

Assessing the effect of Ethnicity

The data from both the GH-2000 and GH-2004 studies were combined to assess the effect of ethnicity on serum IGF-I and P-III-P in men and women which was performed by analysis of variance after adjustment for the effect of age and assay methodology. The estimated mean difference between ethnic groups calculated. All statistical comparisons were two-tailed. $P < 0.05$ was considered statistically significant.

3.3 RESULTS

3.3.1 EFFECT OF ETHNICITY ON IGF-I

IGF-I concentrations in Afro-Caribbean male volunteers were approximately 21.7 ± 2.6% lower than white European men ($p<0.0001$) (Figure 3-1). By contrast, there was no significant difference in serum IGF-I between Indo-Asian ($p=0.65$), Oriental ($p=0.57$) and "Other" ethnic ($p=0.63$) subjects when compared to white European men.

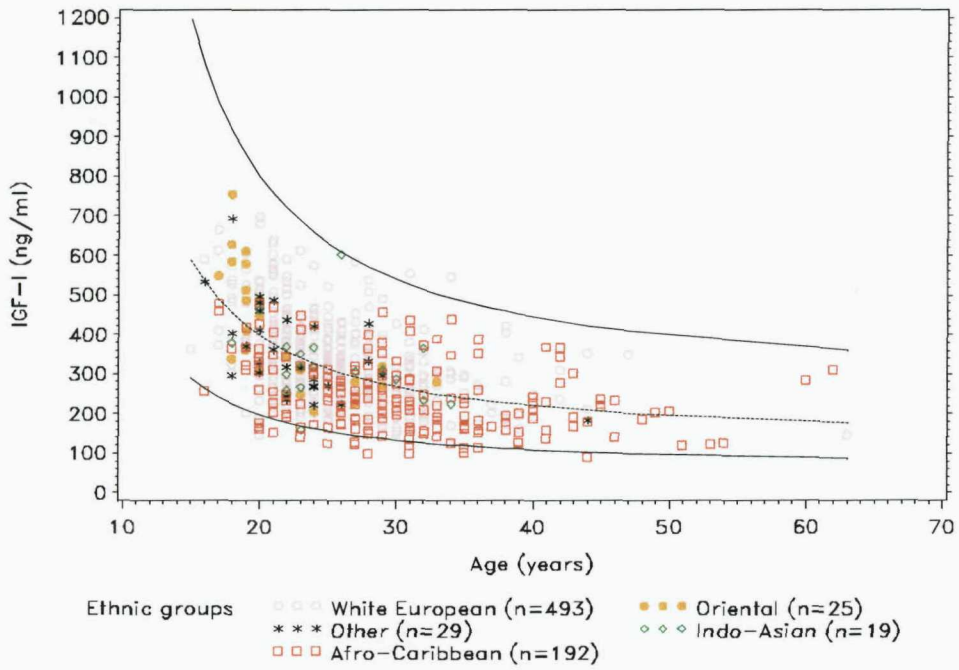

FIGURE 3-1 - AGE DEPENDENT CHANGE OF IGF-I (NG/ML) SERUM CONCENTRATIONS FROM 758 ELITE MALE ATHLETE VOLUNTEERS. THE SOLID AND DOTTED LINES INDICATE THE 99% PREDICTION INTERVALS CALCULATED USING DATA COLLECTED BY THE GH-2000 WHITE EUROPEAN MALE VOLUNTEERS

Similarly, for women IGF-I concentrations were 14.2 ± 5.1% lower in Afro-Caribbean athletes against white European females ($p=0.005$). By contrast, in Oriental women IGF-I concentrations were 15.6 ± 7.0% higher ($p=0.02$). There was, however, no

significant difference in the IGF-I concentrations between Indo-Asian (p=0.66) and of "Other" ethnicity (p=0.58) subjects when compared to white European subjects (Figure 3-2).

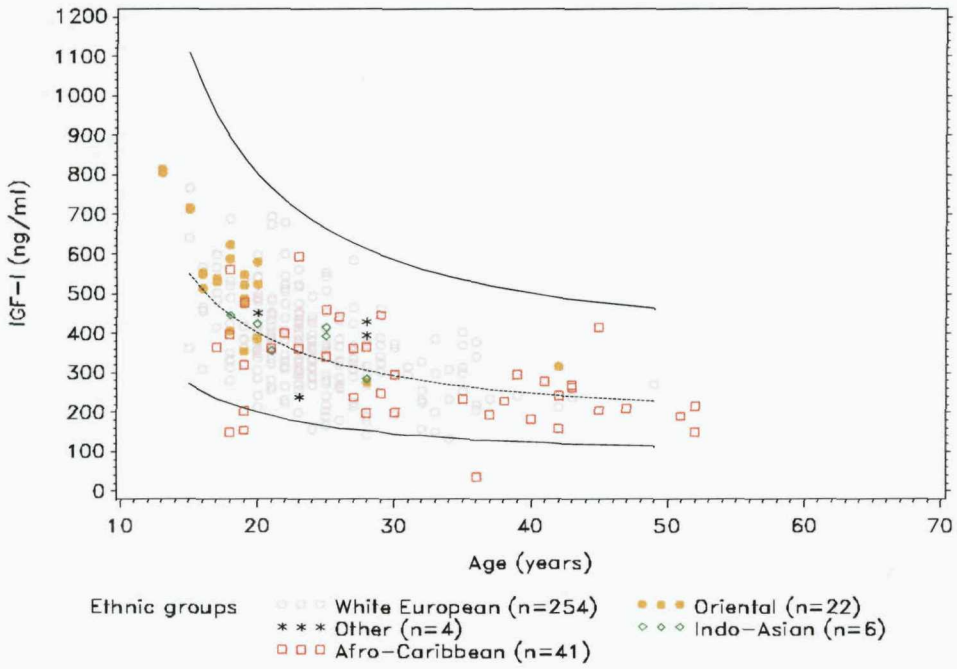

FIGURE 3-2 - AGE DEPENDENT CHANGE OF IGF-I (NG/ML) SERUM CONCENTRATIONS FROM 327 ELITE FEMALE ATHLETE VOLUNTEERS. THE SOLID AND DOTTED LINES INDICATE THE 99% PREDICTION INTERVALS CALCULATED USING DATA COLLECTED BY THE GH-2000 WHITE EUROPEAN FEMALE VOLUNTEERS.

Figure 3-1 and Figure 3-2 present the 99% prediction limits for male and female white European subjects calculated using the previously published GH-2000 data (Healy et al. 2005). The figure shows that all observations from non-white European athletes lie below the upper 99% prediction limits of the white European data.

3.3.2 EFFECT OF ETHNICITY ON TYPE III PROCOLLAGEN

In men, there were small differences in P-III-P concentrations between ethnic groups. P-III-P concentration were approximately 15.2 ± 3.5% lower in Afro-Caribbean men ($p<0.0001$), 26.6 ± 6.6 % lower in Indo-Asian ($p<0.0001$), 19.3 ± 5.8% ($p=0.001$) lower

in Oriental men and 23.6 ± 5.5% (p=0.001) lower in "Other" compared with white European men.

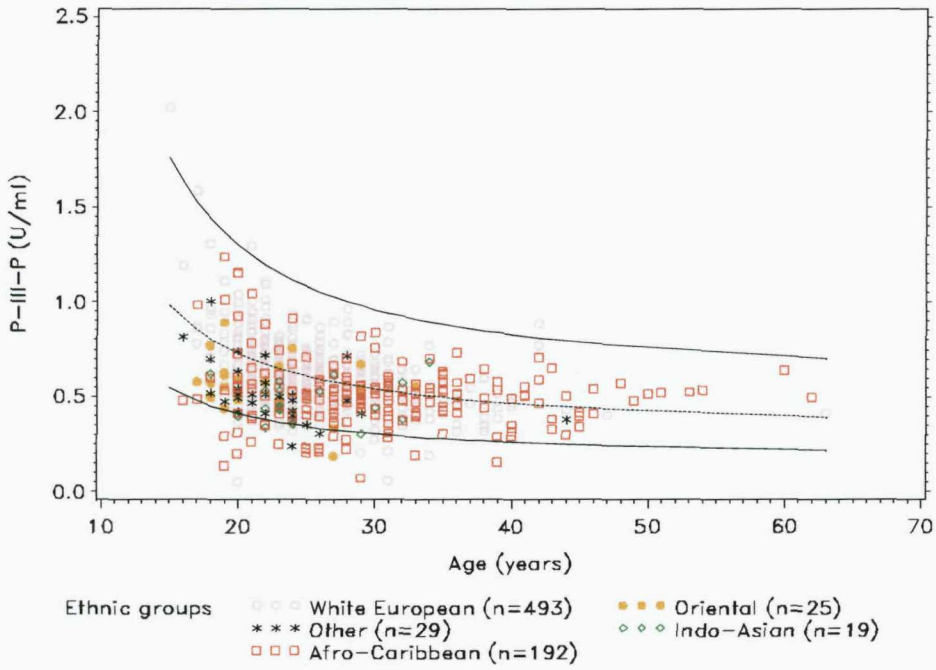

FIGURE 3-3 - AGE DEPENDENT CHANGE OF P-III-P (U/ML) SERUM CONCENTRATIONS FROM 758 ELITE MALE ATHLETE VOLUNTEERS. THE SOLID AND DOTTED LINES INDICATE THE 99% PREDICTION INTERVALS CALCULATED USING DATA COLLECTED BY THE GH-2000 WHITE EUROPEAN MALE VOLUNTEERS

Similarly in females, P-III-P concentrations were 15.7± 4.7% lower in Afro-Caribbean subjects compared to white European women (p=0.0009). There was no difference in P-III-P concentrations between the rest non-white European female ethnic groups; Indo-Asian (p=0.14), Oriental (p=0.85) and "other" (p=0.25).

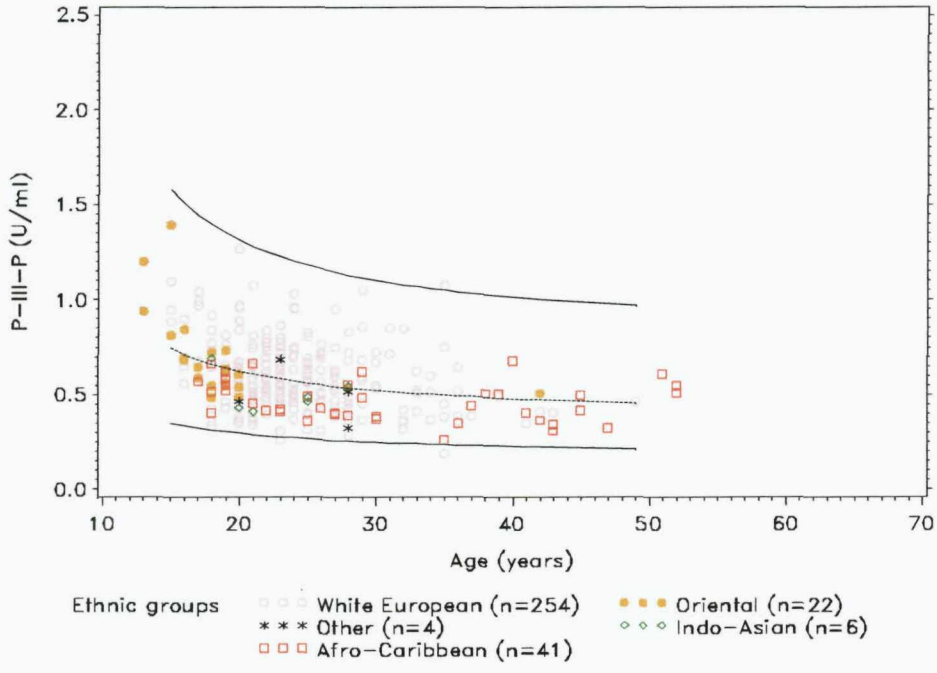

FIGURE 3-4 - AGE DEPENDENT CHANGE OF P-III-P (U/ML) SERUM CONCENTRATIONS FROM 327 ELITE FEMALE ATHLETE VOLUNTEERS. THE SOLID AND DOTTED LINES INDICATE THE 99% PREDICTION INTERVALS CALCULATED USING DATA COLLECTED BY THE GH-2000 WHITE EUROPEAN FEMALE VOLUNTEERS.

The vast majority of observations remain below the upper 99% prediction limit derived from the white European ethnic group in the GH-2000 study, albeit with a number of observations below the lower 99% prediction intervals (Figure 3-3 and Figure 3-4).

3.3.3 AGE DEPENDENT DECLINE IN IGF-I AND P-III-P

The normal age related decline in IGF-I and P-III-P serum concentrations was observed in athletes of all ethnicities and is best represented by a model in which the logarithm of the marker concentration depends linearly on the reciprocal of age. The estimated relationship with age has shown IGF-I serum concentrations decreasing by approximately 35% per decade, whereas P-III-P serum concentrations decreased by 19% per decade (Healy et al. 2005).

3.3.4 EFFECT OF ETHNICITY ON THE GH DETECTION METHOD

The GH-2000 scores could not be calculated for three subjects because their P-III-P concentration was below the lowest limit of quantification (0.01 U/mL). In addition, the GH-2000 score could not be calculated for 1 subject because the IGF-I concentration was the lowest limit of quantification. None of these subjects had a medical condition to explain these unexpected low values.

For men, the age adjusted GH2000 score was calculated for each individual. Calculated scores were plotted against the age of each individual for observations collected from the GH-2000 and GH-2004 study, Figure 3-5. Using the provisional cut-off point of 3.7, it is shown that no observation exceeds this value, indicating that no individual would have been falsely accused of doping, regardless of their ethnic background.

For females, a plot of the GH-2000 scores was constructed including observations from both GH-2000 and GH-2004 studies (Figure 3-6). No individual exceeds 3.7, and thus none would be 'false positives'.

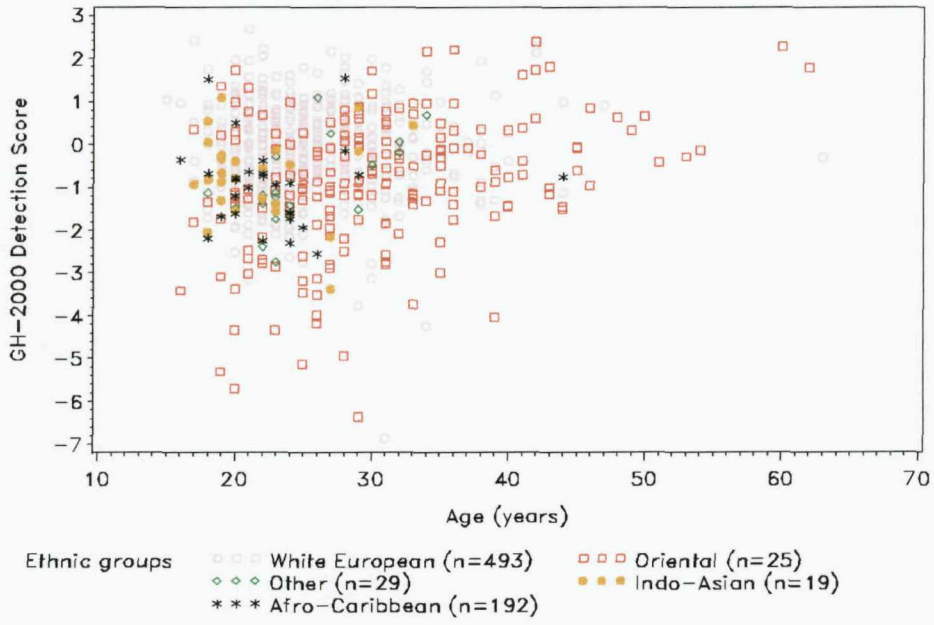

FIGURE 3-5 - CALCULATED GH-2000 SCORES FOR MALE VOLUNTEERS RECRUITED BY THE GH-2000 AND GH-2004 STUDIES, BY ETHNIC GROUP CLASSIFICATION.

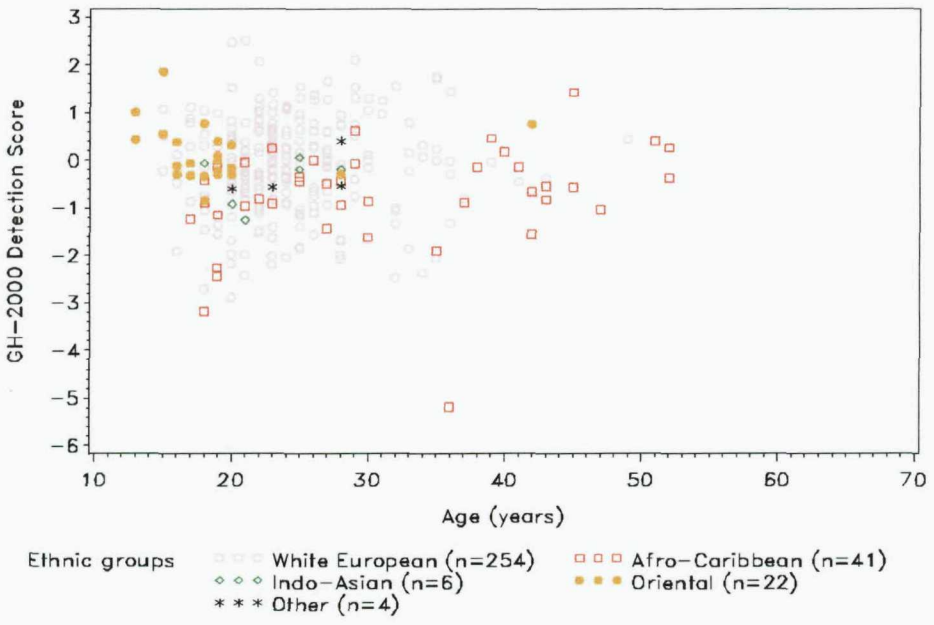

FIGURE 3-6 - CALCULATED GH-2000 SCORES FOR FEMALE VOLUNTEERS RECRUITED BY THE GH-2000 AND GH-2004 STUDIES, BY ETHNIC GROUP CLASSIFICATION.

In white European athletes, the mean GH-2000 score is by definition 0 ± 1. Analysis of variance showed that the female GH-2000 score was significantly lower in Afro-Caribbean women by 0.64 ± 0.17 (p=0.0001) but there was no difference between any other ethnic group and white European women.

In men, analysis of variance showed that the male GH-2000 scores were significantly lower in all ethnic groups compared with Caucasians. In Afro-Caribbean men, the mean score was -0.80 ± 0.10, in Indo-Asian subjects the mean score was -0.85 ± 0.27, for Orientals the mean score was 0.63 ± 0.24 and for athletes in the "Other" ethnic group, the mean score was -0.85 ± 0.22.

3.4 DISCUSSION

This study addressed the concerns raised by the Rome expert working group that the subjects in the original GH-2000 study were predominantly white European and provides evidence that the GH-2000 test is applicable for athletes of other ethnic origins.

Published studies of 'non-elite athlete' healthy volunteers have found little difference in IGF-I concentrations between ethnic groups. Bagg et al found that serum IGF-I concentrations were not different between Samoan, Maori and European populations after adjusting for differences in age, gender, and body composition (Bagg et al. 2006). Furthermore, in a separate study assessing breast cancer risk, IGF-I concentrations were examined in women from four ethnic groups: Latino-American, African-American, Japanese-American and Non-Latino Caucasian. This study showed that following adjustment for age, only the mean age-adjusted IGF-I levels for Latino-American women (116 ng/mL) was significantly lower from the other three groups, with the remaining three groups having similar mean age-adjusted IGF-I values of around 145 ng/mL (Delellis et al. 2003). In a third study involving men, the age adjusted mean IGF-I concentrations were approximately 10% lower in black men than

in white men (McGreevy et al. 2005). The common finding from these studies is that no non-White European ethnic group appears to have significantly higher IGF-I concentrations than the White-European ethnic group.

Recently, an independent research group based in Australia has published results from a large cross-sectional trial involving elite athletes from African, Asian, Oceanian and Caucasian ethnic groups (Nelson et al. 2006). The aim of this research was to assess the effect of ethnicity on the concentrations of a number of GH-dependent biomarkers, including IGF-I and P-III-P. Similarly, to the results obtained by the GH-2000 group, it was concluded that the key determinant of IGF-I and P-III-P concentration was age and (like GH-2000) there was also a small difference between genders. This is in keeping with our findings that the major determinant of IGF-I and P-III-P concentrations in the post-competition setting is age and that the normal age related decline occurs in athletes of all ethnicities in a similar manner to the general population.

In the Australian study, there was no significant difference in IGF-I concentrations between the four ethnic groups but P-III-P was approximately 8.5% higher in Asians compared with Caucasians. Overall, however, the contribution of ethnicity to the variation of IGF-I and P-III-P was less than 2%.

The main difference between the Australian and GH-2004 study is the setting in which the samples were taken. While the Australian samples were predominantly taken out-of-competition, the primary aim of the GH-2004 cross-sectional study was to assess the P-III-P an IGF-I post-competition reference ranges when most dope testing is undertaken. The results shown here are broadly consistent with the studies of the general population and the Australian study indicating that there are no major ethnic differences in IGF-I and P-III-P in the immediate post-competition setting. While there were small differences in the mean values, the vast majority of individuals lie within the 99% prediction limits for the white European athletes.

The second and key finding of this work was that the scores from the proposed detection formulae calculated using non-white European samples were not significantly different to those calculated using white European subjects. This is important to ensure that no athlete would be falsely accused of doping with GH because of his or her ethnicity.

The study has a number of limitations. Despite support from the World Anti-Doping Agency, many event organizers refused permission for us to attend the meetings and coaches often refused permission for us to approach their athletes. Consequently, recruitment was lower than the GH-2000 research project. The number of high profile doping cases may have made athletes and coaches suspicious that our research may have led to a case against the volunteer despite reassurances that all data and samples would be anonymized. Although we provided translated information sheets, recruitment may have been harder because we were targeting non-White European athletes for whom English may not have been their first language and for whom cultural differences may have led to reluctance to provide a blood sample. Although all athletes were competing at a national or international standard, the reluctance to participate may have resulted in our recruiting slightly less highly performing athletes than in GH-2000.

A second criticism of the study is that some athletes might have been included in the study taking GH, which would have extended the upper limit of the reference range. To minimise this risk, the research team required subjects to confirm in writing that they had not taken performance-enhancing drugs. The absence of outliers well above the upper prediction limit supports the assumption that subjects complied with this requirement.

In conclusion, this analysis showed that there are no major differences in the IGF-I and P-III-P concentrations in elite athletes of differing ethnicities in the immediate post-competition period. Furthermore, and of greater importance in the context of a doping test, we have shown that the same holds for the proposed detection formulae.

This study provides further evidence of the validity of a test for GH abuse based on the measurement of GH dependent markers.

CHAPTER 4: THE EFFECT OF SPORTING INJURY ON GH-DEPENDENT MARKERS

4.1 INTRODUCTION

This study was carried out to in order to examine the effects of injuries and fractures on the "markers" of growth hormone abuse in amateur and elite athletes to ensure that an athlete could not use a skeletal injury as a reason for a failed doping test

The proposed GH-2000 detection method is based on the measurement of GH-dependent protein markers, insulin like growth factor-I (IGF-I) and type 3 pro-collagen (P-III-P) (Sonksen 2001).

As these proteins occur physiologically, detection of GH abuse must rely on finding concentrations in excess of those found in an established reference range. There have been fears that skeletal injury may alter the concentrations of both IGF-I and P-III-P. The latter is of particular concern as it is a marker of soft tissue and bone turnover (Haukipuro et al. 1990; Bail et al. 2001). For example, following a tibial shaft fracture, bone turnover increases for at least 24 weeks, as demonstrated by an increase in P-III-P (Veitch et al. 2006).

It is therefore important to examine whether elevations in either of these proteins following a sporting injury could lead to a false accusation of doping with GH. The aim of this study was to assess the effect of a musculo-skeletal injury on serum IGF-I and P-III-P concentrations in amateur and elite sportsmen and women and most importantly assess the effect of injury on the performance of the GH detection method proposed by the GH-2000 group.

4.2 METHODS

127 male (age 29.6 ± 0.9 yrs, range 17 – 68 yrs) and 30 female (32.0 ± 2.1 yrs, 19 – 63) amateur athletes were recruited from the Accident & Emergency outpatient clinic or orthopaedic fracture clinic at the Southampton University Hospitals Trust, UK. 16 male (age 23.8 ± 1.4 yrs) and 10 female (24.7 ± 1.6 yrs) professional athletes were recruited from the British Olympic Medical Centre. Subjects were recruited within 10

days of a bony or soft tissue musculoskeletal injury of sufficient severity to prevent training for at least 3 days. Exclusion criteria included neoplastic disease, diabetes, pregnancy or lactation and any condition likely to affect the GH-IGF axis.

A brief medical and sporting history was obtained including gender, race, age and sporting category. The nature of the injury was classified as either soft-tissue (joint dislocations, ligament or tendon injury, sprain injury or bruises) or bony (fractures, cracks or chips).

Injury severity was classified by a blinded physician as severe (fracture of a major long bone or disruption of a major joint or multiple injuries), moderate (fracture of a cancellous bone, wrist or forefoot fracture, ligament damage) or mild (digit fracture or minor soft tissue injury).

A 10ml venous blood sample was obtained. Follow up blood samples were scheduled on 7 ± 3, 14 ± 3, 21 ± 3, 28 ± 3, 42 ± 7, 84 ± 14 days after the injury.

Analytical Methods

Blood samples were collected and were centrifuged within 12 hours of collection. The serum was stored at -80C until analysis at the WADA accredited laboratory at the Drug Control Centre, King's College London, UK.

All samples were analysed in duplicate. Serum IGF-I was measured by the DSL-5600 ACTIVE® IGF-I IRMA (Diagnostic Systems Laboratories, Inc., Webster, TX). Serum P-III-P was measured by a two-stage sandwich RIA (CIS Biointernational; Oris Industries, Gif-Sur-Yvette Cedex, France).

Statistical Analysis

Statistical analysis was performed using the SAS software (SAS Institute, Inc., Cary, NC). IGF-I and P-III-P were log-transformed. The changes in the biomarkers and GH-2000 score compared to baseline values was assessed by ANOVA by injury type and severity, after adjustment for age and gender.

Adjusting for IGF-I Assay differences

Due to the differences in the assay methodology used in the GH-2004 Injury and the assay methodology used to derive the GH-2000 detection methodology, statistical assay adjustments were necessary in order to be able to use this methodology. Further details regarding these statistical adjustments are given in section 2.5.1.

In order to assess whether injury affected the GH-2000 score significantly, the score for each participant was calculated at every time point using the measured P-III-P and assay adjusted IGF-I and subject age.

A cut-off point for these GH-detection formulae has not been agreed by the World Anti-Doping Agency. It has been suggested that a possible cut-off point should be at the value of 3.7 equivalent to a false positive rate of approximately 1 in 10,000 tests.

A total of 482 blood samples were taken representing an average of 2.63 ± 0.14 samples per subject.

4.3 RESULTS

4.3.1 EFFECT OF INJURY ON IGF-I

There was no significant difference at baseline IGF-I (ng/ml) concentrations between the two injury types (p=0.91) and the varying levels of severity (p=0.54). The baseline IGF-I concentration was 522.6 ± 21.8 ng/ml. There was no significant change in IGF-I during the study either in the whole population (Figure 4-1) or when analysed separately according to severity or injury type.

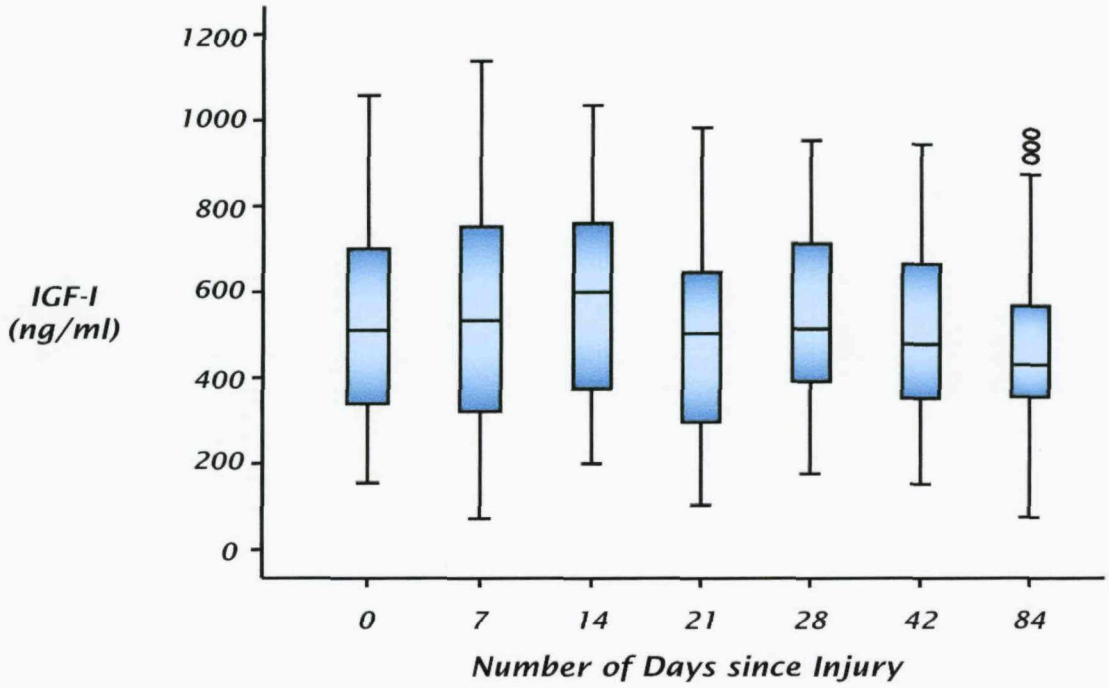

FIGURE 4-1 - BOXPLOT SHOWING CHANGE IN IGF-I OVER 84 DAYS FOLLOWING INJURY. THE MEDIAN AND UPPER AND LOWER 25% CENTILES ARE REPRESENTED BY THE BOX. THE WHISKER REPRESENTS THE UPPER AND LOWER 95% CONFIDENCE INTERVALS. MILD OUTLIERS ARE REPRESENTED BY CIRCLES AND EXTREME OUTLIERS ARE REPRESENTED BY STARS.

The intra-individual variability of the IGF-I (ng/ml) concentrations for this study period was estimated as 24.3 ± 8.5%. Although we do not have the statistical power to assess differences formally between elite and amateur athletes, the pattern seen in the nine elite athletes with multiple samples appeared different from amateur athletes.

4.3.2 EFFECT OF INJURY ON TYPE III PROCOLLAGEN

The baseline P-III-P concentration was 0.42 ± 0.02 U/L. There was no significant difference at baseline P-III-P (U/ml) concentrations between the two injury types (p=0.28) and the varying levels of severity (p=0.49).

In contrast to IGF-I, there was a significant rise in P-III-P following both soft tissue and bony injuries. The pattern of response differed between soft tissue and bony injury and therefore the two types of injury were analysed separately. Following a soft tissue injury, there was a rise in P-III-P that peaked 14 days after the injury, the magnitude of which varied 41.1 ± 16.6% and 44.3 ± 19.2% according to the severity of the injury (Figure 4-2).

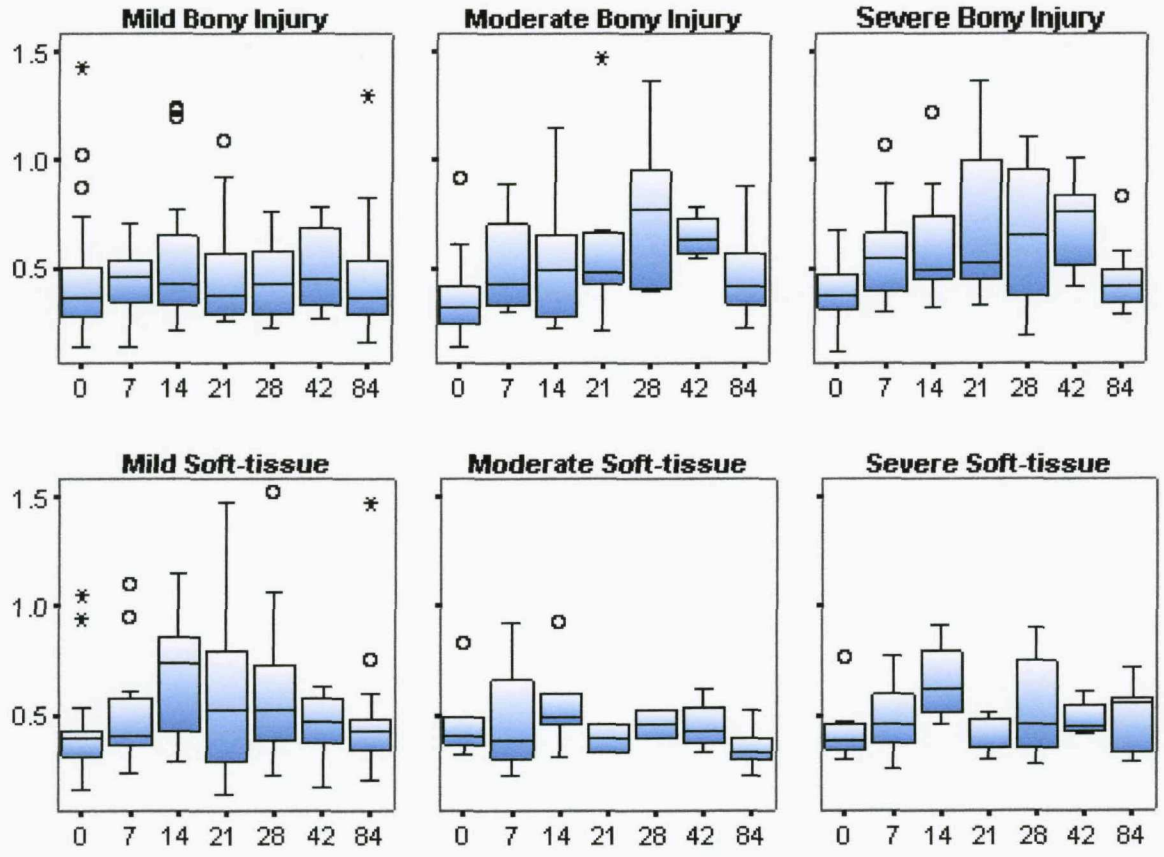

FIGURE 4-2 - BOX PLOTS SHOWING CHANGE IN P-III-P CONCENTRATION OVER 84 DAYS FOLLOWING INJURY ACCORDING TO INJURY TYPE AND SEVERITY. THE MEDIAN AND UPPER AND LOWER 25% CENTILES ARE REPRESENTED BY THE BOX. THE WHISKER REPRESENTS THE UPPER AND LOWER 95% CONFIDENCE INTERVALS. MILD OUTLIERS ARE REPRESENTED BY CIRCLES AND EXTREME OUTLIERS ARE REPRESENTED BY STARS.

Following a bony injury, the peak rise occurred 42 days after injury and remained significantly elevated until day 84. When the bony group was analysed by injury severity, no significant change was seen in the mild injury group. In contrast, the peak rise of P-III-P was 75.1 ± 26% (p= 0.002) in those with a moderate injury and 58.0 ± 17.7% (p=0.002) in those with severe bony injuries (Figure 4-2b).

The intra-individual variability of the P-III-P (U/ml) concentrations for this study period was estimated as 33.1 ± 11.1%.

Although we do not have the statistical power to assess differences formally between elite and amateur athletes, the pattern seen in the nine elite athletes with multiple samples appeared no different from amateur athletes.

4.3.3 EFFECTS OF INJURY ON GH-2000 DETECTION METHOD

The change in male GH-2000 score with time following injury is shown in Figure 4-3. There was no significant change in the calculated GH-2000 score with time following an injury (p=0.78). The intra-individual variation with time following injury was 0.77 ± 0.27. This is up to six times higher than the intra-individual variation seen in non-injured elite and amateur athletes; 0.12 and 0.15 for amateur and elite male athletes respectively (Holt et al. 2006).

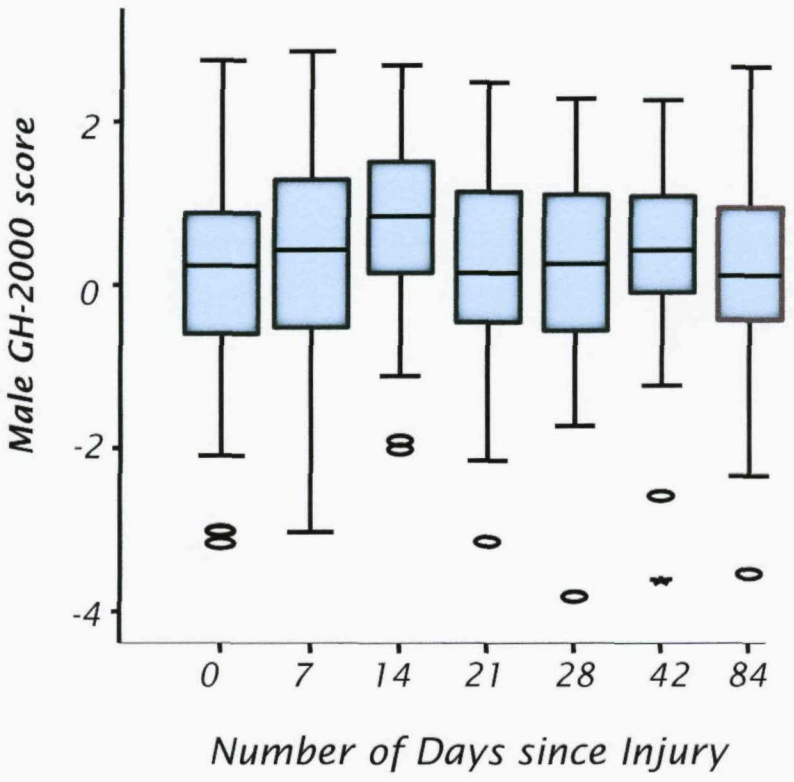

FIGURE 4-3 - BOX PLOT SHOWING CHANGE IN GH-2000 SCORE IN MEN OVER 84 DAYS FOLLOWING INJURY. THE MEDIAN AND UPPER AND LOWER 25% CENTILES ARE REPRESENTED BY THE BOX. THE WHISKER REPRESENTS THE UPPER AND LOWER 95% CONFIDENCE INTERVALS. MILD OUTLIERS ARE REPRESENTED BY CIRCLES AND EXTREME OUTLIERS ARE REPRESENTED BY STARS.

The scores of the GH-2000 age adjusted discriminant function formulae for females is shown in Figure 4-4 and did not change significantly following injury. The intra-individual variation with time following injury was 0.56 ± 0.33%. This is twice as high as the 0.25 intra-individual variation measured in non-injured elite and amateur female athletes (Holt et al. 2006).

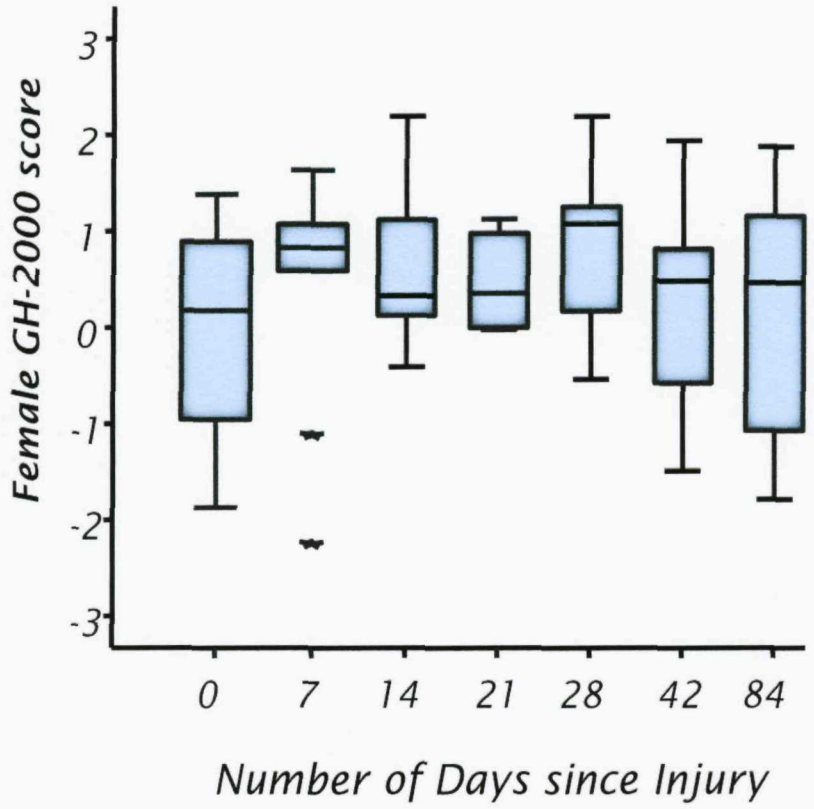

FIGURE 4-4 - BOX PLOT SHOWING CHANGE IN GH-2000 SCORE IN WOMEN OVER 84 DAYS FOLLOWING INJURY. THE MEDIAN AND UPPER AND LOWER 25% CENTILES ARE REPRESENTED BY THE BOX. THE WHISKER REPRESENTS THE UPPER AND LOWER 95% CONFIDENCE INTERVALS. MILD OUTLIERS ARE REPRESENTED BY CIRCLES AND EXTREME OUTLIERS ARE REPRESENTED BY STARS.

All samples from both men and women were below the 3.7 cut-off limit, thus placing all injured individuals, below the cut-off limit. Thus, no individual would have been falsely accused of GH doping.

4.4 DISCUSSION

The purpose of this part of the thesis was to assess the effect of injury on the serum concentrations of IGF-I and P-III-P and in particular whether an identified effect would have an adverse effect on the performance of the discriminant function scores used to identify GH doping in sports.

Injuries are common in sports and it has been suggested that injuries may cause a significant elevation in the serum concentrations of P-III-P (Veitch et al. 2006; Kurdy et al. 1998) and IGF-I (Bail et al. 2001). Therefore, it is vital to ensure that no athlete would be falsely accused of GH doping under the proposed methodology because of an injury. Equally important is the need to eliminate any arguments of a defence lawyer that the accused athlete had been falsely accused of doping because they had sustained an injury, which caused them to have elevated measurements of IGF-I and/or P-III-P and therefore GH-2000 score.

In order to gain additional information on how P-III-P and IGF-I are affected by injury, the type of injury incurred by the recruited volunteers was also considered, as well as the severity of the injury. Although fluctuations were found in IGF-I concentrations during the study period, these were small and no larger than the intra-individual variation among healthy elite and amateur athletes. There were no significant changes in IGF-I due to injury during the study, regardless of the type or severity of injury. This finding was unexpected because it is well known that IGF-I concentrations fall with acute illness (Mesotten et al. 2006). It is possible that the reason why we observed no change was the severity of the injury compared with the illness needed to result in a fall in IGF-I on the intensive care unit.

In contrast to IGF-I, we demonstrated a significant rise in P-III-P following injury, the magnitude and duration of which varied according to the injury type and severity. These findings are in keeping with previous studies of fracture that have shown an increase in P-III-P following fracture (approximately 57±9% at week 8 following injury). In a previous study of individuals with malleolar fracture, P-III-P reached a

peak after 2 weeks of injury, whereas for individuals who had sustained a tibial fracture, P-III-P values were maximal after 12 weeks (Stoffel et al. 2007). Two further studies of subjects with tibial fractures showed maximum P-III-P concentrations after two and eight weeks – agreeing with the finding in this work (Joerring et al. 1994; Veitch et al. 2006). Previous work examining poorly healing fractures demonstrated elevated P-III-P concentrations for up to 10 weeks (Kurdy 2000). There are no previous reports of the change in P-III-P following soft tissue injury.

The observed increase in P-III-P can be explained by considering the process of fracture healing which is divided into 3 distinct phases; inflammation, regeneration, and remodelling. During the initial inflammatory phase, there is an increased production of type III collagen caused by the initial hematoma following the injury. This hematoma is replaced rapidly by moderately dense fibrous tissue containing predominately type III collagen (Stoffel et al. 2007). As part of the healing process this tissue is later replaced by type I collagen. Time for fracture healing and extent of changes in markers of bone metabolism are mainly dependent on fracture size.

Although the rise in P-III-P was statistically significant, the key aspect of this work was to assess whether the rise in P-III-P concentration could lead to the false accusation of an athlete for GH doping. In contrast to the 41-75% increase in P-III-P following injury, the rise in P-III-P following the administration of rhGH was 299% over a 14 day period (Powrie et al. 2007).

No athlete was found to exceed the provisional cut-off point of 3.7, equivalent to a 1:10000 false positive rate. This reinforces the conclusion that an injury would not adversely affect the performance of this detection method.

The study has a number of limitations. The number of female and elite volunteer participants in this study was small and a bigger sample size would have improved the power of the study to detect differences. A second criticism of the study was the need to make an adjustment for the change in the IGF-I assay in order to calculate the GH-2000 score. The withdrawal of the original Nichols RIA used in the GH-2000

study meant that all the samples in the GH-2000 and GH-2004 studies could not be measured using the same assay. By using a direct comparison between assays, we were able to make an assay adjustment with few assumptions (See section 2.4.1). The main assumption was that there was a linear proportional relationship between the two assays and the regression line passed through zero on the x and y-axis. It also assumes that there is no degradation in the freezer with time. This is a reasonable assumption because previous studies have shown little degradation of samples stored at -80C (Evans et al. 2001; Zhang et al. 1998). The weakness of the approach is that by necessity, its uses a relatively small number of samples to derive the regression.

In conclusion, this study shows that following an injury, IGF-I concentration is not significantly affected. In agreement with previously published research, there is a rise in P-III-P concentration following injury. The key finding from this study, however, is the fact that the application of the GH-2000 test for GH abuse would not have resulted in any false positives in athletes with an injury, regardless of its severity or type. This study provides further evidence for the validity of a test for GH abuse based on the measurement of GH dependent markers.

CHAPTER 5: INTRA-INDIVIDUAL VARIABILITY OF GH-DEPENDENT MARKERS

5.1 INTRODUCTION

The aim of this study is to assess the stability of the IGF-I and P-III-P markers, over the period of 1 year. By monitoring the concentrations of these GH-dependent markers in volunteers for a period of 12 months, information can be given on the stability of these markers over time in individuals. Following that, the effect of the intra-individual variability on the examined markers over this period of time was assessed on the discriminant function formulae used to detect GH doping. This is an important part for the development of reliable methodology to detect GH, because if found that there is considerable intra-individual variability, this could lead to the false accusation of an athlete for GH doping. On the other hand, if the intra-individual variation is considerably less than the population variability this raises the possibility of using the athlete as his or her own control and extending the concept of an athlete's biological passport.

5.2 METHODS

The results considered in this article originated from 4 research studies. All four were of a longitudinal design and collected observations over a period of time lasting up to 12 months. These studies were:

GH-2000 Longitudinal Study – Elite Athletes

Up to 4 fasting blood samples were obtained from 175 male and 83 female elite athletes over a period of up to 1 year. The athletes were recruited at national or international sporting events. The vast majority of the athletes were white European with only 4 Afro-Caribbean and 1 Oriental subject. The mean age of the men and women was 25.9 ± 0.4 yrs and 24.7 ± 0.5 yrs respectively.

After the blood sample collection, blood samples were allowed to clot and then centrifuged at 15 minutes at 3000rpm before the serum was separated and frozen at -80C prior to analysis. Serum IGF-I was analysed after hydrochloric acid-ethanol

extraction by radioimmunoassay using authentic IGF-I for labelling (Nichols Institute Diagnostics, San Juan Capisttrano, CA). Serum P-III-P concentration was determined by a RIA (International CIS, Gif-sur Yvette, France).

GH-2000 Placebo controlled double blind study– Amateur Athletes

The GH-2000 project undertook a double blind placebo controlled GH administration study (Dall et al. 2000; Longobardi et al. 2000). In this study, GH was administered for 28 days and then subjects were follow-up for a further 56 days during the washout period. During the 3-month observation period, the subjects had up to 7 blood samples taken. The placebo treated subjects provide us with an opportunity to examine the intra-individual variation of IGF-I and P-III-P in amateur athletes. The study included 18 women (24.3 ± 0.9 years) and 21 men (26.0 ± 0.9 years) who received placebo. All subjects were white European.

Samples in this study were stored and analysed using the same analytical procedures as the GH-2000 Longitudinal study.

GH-2004 Longitudinal Study – Amateur Athletes

41 amateur athletes (20 men and 21 women aged 18-42 yrs) were recruited through the University of Southampton. Most volunteers were members of sports clubs and participated in team or individual sports. Each athlete attended the Southampton Wellcome Trust Clinical Research Facility on four occasions over a four month period for a blood sample. In addition, on one occasion, body composition was measured by anthropometry and bioimpedance. Each subject underwent a maximal exercise test on a treadmill to ascertain the effect of exercise on the GH2000 testing score.

2 subjects dropped out of the study after the first visit; one because of injury and the other for personal reasons. Exercise testing was performed on 35 volunteers.

All samples were analysed in duplicate at the Drug Control Centre, King's College London, UK. Serum IGF-I serum concentration was measured by the DSL-5600

ACTIVE® IGF-I IRMA manufactured by DSL Laboratories (Diagnostic Systems Laboratories, Inc., Webster, TX).

Serum P-III-P was measured by a two-stage sandwich RIA (CIS Biointernational; Oris Industries, Gif-Sur-Yvette Cedex, France).

Italian Longitudinal Study –Elite Athletes

IGF-I and P-III-P had been previously measured in a longitudinal follow-up study of 25 male and 22 female elite Italian athletes from 9 different sporting disciplines (Sartorio et al. 2006; Sartorio et al. 2004b). The mean age of the men and women was 22.6 ± 0.2 yrs and 22.5 ± 0.2 yrs respectively. Four blood samples had been taken over a six-month period. IGF-I concentrations were determined by using a commercial immunoassay kit (Mediagnost Gmbh, Tübingen, Germany). P-III-P levels were determined using the Orion Diagnostica RIA kits (Oy, Espoo, Finland).

Statistical Analysis

The statistical analysis performed in this study concerned four main areas. Analysis was carried out using the SAS software™ (SAS Institute Inc., Cary, NC, USA) and Minitab version 15.1.1.

Calculating Assay Methodology conversion factor

In order to be able to make direct comparisons between the results from each study, due to the different assay methodology used between studies, it was necessary to derive an assay conversion methodology that would transform IGF-I & P-III-P concentrations to the scales used by the GH-2000 assays.

Statistical assay conversion equations were calculated to allow the direct comparison of measurements measured from each study and allow the use of the GH-2000 detection methodology. Further details can be found in section 2.4.1.

Quantifying intra-individual variability on IGF-I and P-III-P

Analysis of variance (ANOVA) analysis was performed to estimate the intra-individual variability components of the IGF-I and P-III-P concentrations for each population. Further analysis was carried out using ANOVA methodology to assess whether there were any significant differences between genders. All analysis carried out adjusted for differences in the biomarker concentration due to age.

Assessing effect of intra-individual variability on testing scores

Similarly, ANOVA methodology was performed to estimate the inter and intra-individual variability components of the GH-2000 discriminant function formulae scores.

The intra-individual variability estimate for IGF-I, P-III-P and the calculated testing scores was calculated using the data collected at resting conditions as well as immediately post-competition. The aim of this was to assess whether intra-individual variability estimates would be adversely affected by exercise. The intra-individual variability was estimated using one way ANOVA.

5.3 RESULTS

5.3.1 QUANTIFYING INTRA-INDIVIDUAL VARIABILITY ON IGF-I, P-III-P AND THE TESTING SCORES

All subjects across the four studies had similar baseline characteristics. Table 5-1 shows the average age, weight and height between males and females for each of the four studies.

		Males	Females
	Age (years)	26.0 ± 0.22	24.7 ± 0.24
GH-2000 Longitudinal Study	**Weight (kg)**	78.3 ± 0.33	65.5 ± 0.68
	Height (cm)	181.4 ± 0.40	171.8 ± 0.58
	Age (years)	22.5 ± 1.35	20.8 ± 0.48
GH-2004 Longitudinal study	**Weight (kg)**	80.0 ± 2.03	67.9 ± 2.53
	Height (cm)	1.83 ± 0.189	1.71 ± 0.168
	Age (years)	26.0 ± 0.93	24.3 ± 0.85
GH-2000 Placebo controlled DB study	**Weight (kg)**	81.2 ± 2.89	57.5 ± 1.28
	Height (cm)	181.6 ± 1.53	168.0 ± 1.46
	Age (years)	22.5 ± 1.35	20.8 ± 0.48
Italian Longitudinal Study	**Weight (kg)**	80.0 ± 2.03	67.9 ± 2.53
	Height (cm)	1.83 ± 0.189	1.71 ± 0.168

TABLE 5-1 - BASELINE CHARACTERISTICS FOR AGE, WEIGHT, HEIGHT OF THE RECRUITED VOLUNTEERS BY GENDER FOR EACH OF THE LONGITUDINAL STUDIES CONSIDERED

The estimated intra-individual variability for IGF-I, P-III-P and the GH-2000 scores as calculated using the testing functions are presented for each study in Table 5-1. The estimated intra-individual variability estimates is similar across the four studies. For this reason, only one of these studies was selected to illustrate each of these results graphically.

	IGF-I (ng/ml)	P-III-P (U/ml)	Male GH-2000 Score	Female GH-2000 Score
GH-2000 Longitudinal Study	14.7±0.62%	15.0±0.62%	0.60±0.03	0.57±0.04
GH-2000 Placebo controlled study	14.0±0.65%	12.2±0.57%	0.48±0.03	0.54±0.04
Italian Longitudinal Study	16.1±0.10%	17.9±0.01%	0.44±0.04	0.38±0.03
GH-2004 Longitudinal Study (only resting samples)	14.5±0.96%	7.21±0.48%	0.41±0.04	0.38±0.04
GH-2004 Longitudinal Data (Resting & Post-exercise samples)	14.1±0.82%	7.20±0.42%	0.38±0.03	0.39±0.03

TABLE 5-2 – ESTIMATES OF THE PHYSIOLOGICAL VARIABILITY FOR IGF-I, P-III-P AND THE MALE AND FEMALE TESTING FUNCTIONS FOR EACH LONGITUDINAL STUDY CONSIDERED. FOR THE GH-2004 STUDY TWO SETS OF ESTIMATES ARE GIVEN TO ILLUSTRATE THE EFFECT OF EXERCISE ON THE MAGNITUDE OF INTRA-INDIVIDUAL VARIABILITY

Figure 5-1 and Figure 5-2 show the IGF-I and P-III-P concentrations respectively for each individual at each visit date for the GH-2004 Longitudinal study.

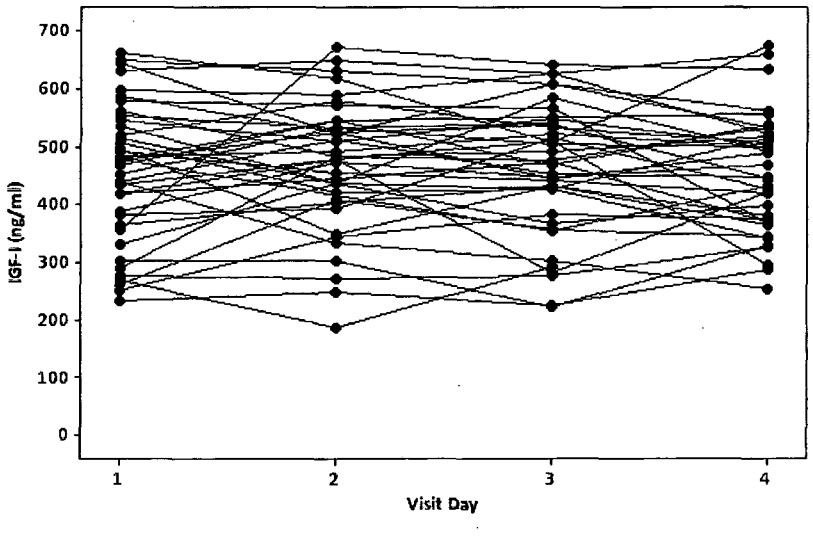

FIGURE 5-1 – IGF-I (NG/ML) CONCENTRATIONS AS MEASURED FOR EACH INDIVIDUAL VOLUNTEER IN THE GH-2004 LONGITUDINAL STUDY. THE DOTS ARE JOINED FOR EACH INDIVIDUAL AND REPRESENT THE MEASURED IGF-I (NG/ML) CONCENTRATIONS.

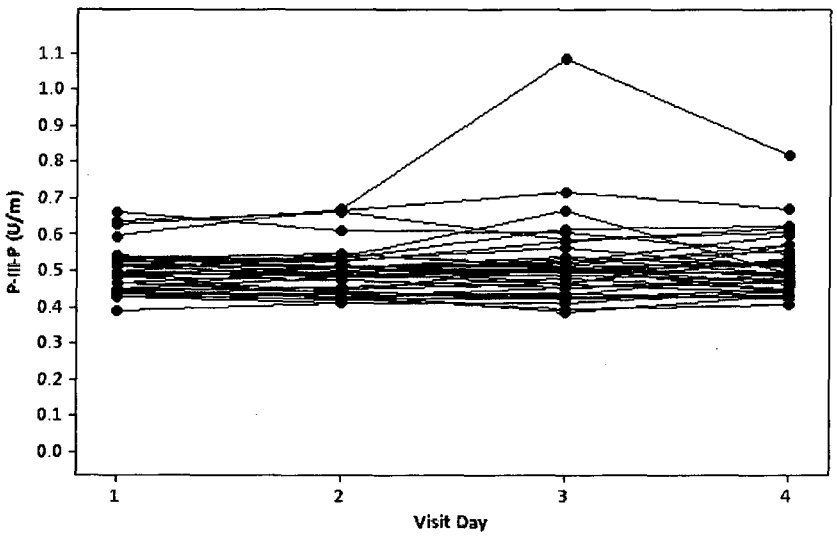

FIGURE 5-2 – P-III-P (U/ML) CONCENTRATIONS AS MEASURED FOR EACH INDIVIDUAL VOLUNTEER IN THE GH-2004 LONGITUDINAL STUDY. THE DOTS ARE JOINED FOR EACH INDIVIDUAL AND REPRESENT THE MEASURED P-III-P (U/ML) CONCENTRATIONS.

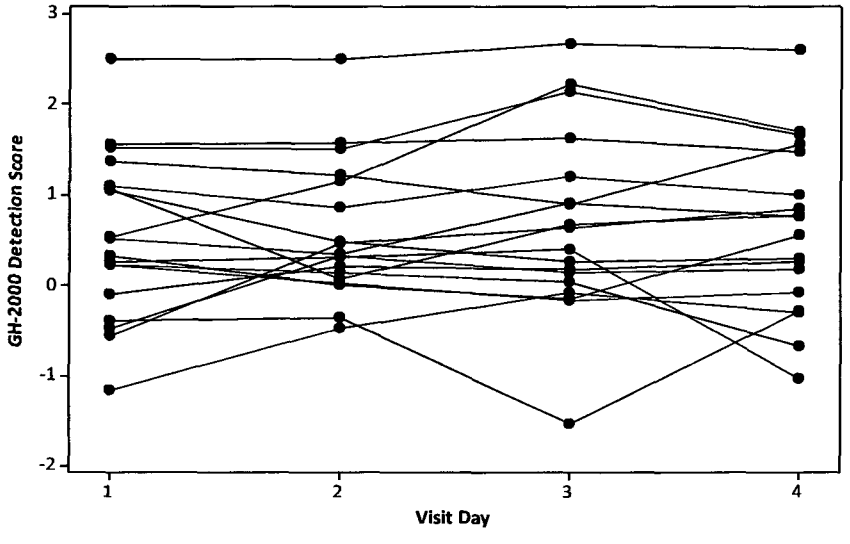

FIGURE 5-3 – CALCULATED GH-2000 SCORES FOR EACH MALE VOLUNTEER IN THE GH-2004 LONGITUDINAL STUDY. THE DOTS ARE JOINED FOR EACH INDIVIDUAL.

Figure 5-3 and Figure 5-4 illustrate the testing scores for the male and female subjects of the GH2004 longitudinal study respectively. All three figures include resting samples only.

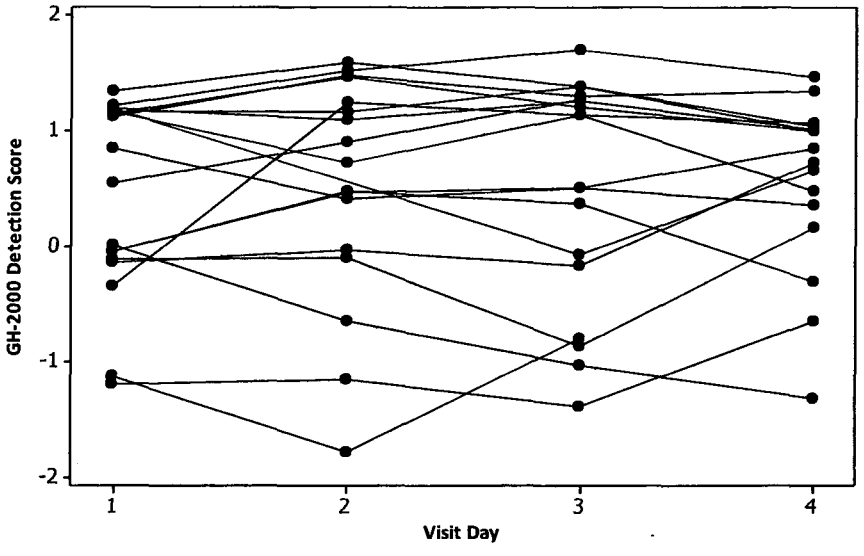

FIGURE 5-4 - CALCULATED GH-2000 SCORES FOR EACH FEMALE VOLUNTEER IN THE GH-2004 LONGITUDINAL STUDY. THE DOTS ARE JOINED FOR EACH INDIVIDUAL.

5.4 DISCUSSION

The results from the four longitudinal studies showed remarkably consistent results, with no apparent difference between amateur and elite athletes. The intra-individual variability for IGF-I varies between 13.9 – 16% while the variability for P-III-P varies from 12 – 19%. These data suggest that the sensitivity of a test for GH based on markers might be improved by the concept of an athlete "passport" or "profiling".

Although the GH-2000 detection method provides good sensitivity and specificity to detect those abusing GH, there are some limitations. Like many endocrine tests, there is a trade-off between sensitivity and specificity. The usual laboratory ranges of mean plus or minor two standard deviations are not applicable in doping analysis because 5% of individuals would lie outside this range by definition and so the risk of a false

positive is unacceptably high. Although the specificity of the test has not been determined, it is thought that a false positive risk of 1 in 10000 is likely to be acceptable. At this level of specificity, the GH-2000 formula will correctly identify 79% of men receiving GH but only 36% of women. The sensitivity improves with dose of GH in women and deteriorates in both sexes after the discontinuation of GH (Powrie et al. 2007).

A further disadvantage of the marker approach is that the formula needs to include an adjustment for age because both GH secretion and markers decrease with age after the 2nd decade of life. Although there are good data about the effect of age on GH markers in elite athletes (Nelson et al. 2006; Healy et al. 2005), the need for the correction could potentially reduce the sensitivity because this assumes an average rate of fall in markers which may vary from one individual to another.

Recently there have been discussions about the possible use of a biological "athlete passport". As it is known that there are smaller differences in intra-individual variation of marker concentrations than inter-individual variation, the knowledge of previous measures of IGF-I and P-III-P may improve the sensitivity of detection. Furthermore, there would be no need for an age correction or assay adjustment if the subsequent samples were taken with 1 year of the first sample and measured by the same assay.

The analysis in this thesis has shown that the intra-individual variation of IGF-I and P-III-P in elite and amateur athletes is small and not much greater than assay variability. These data raise the possibility that an athlete passport or profiling might improve the sensitivity of the GH-2000 test but further work, for example examining the variables that affect the intra-individual variation of IGF-I and P-III-P is needed.

CHAPTER 6: INDEPENDENT VALIDATION OF THE GH-2000 DETECTION METHOD

6.1 INTRODUCTION

The proposed GH-2000 detection method as discussed is based on the measurement of two GH-dependent markers, insulin-like growth factor–I (IGF-I) and type 3 pro-collagen (P-III-P), both of which rise in response to GH administration in a dose dependent manner (Sonksen 2001).

The Institut fur Dopinganalytik und Sportbiochemie in Kreischa undertook a further placebo controlled double blind GH administration study and developed an alternative formula based on IGF-I and P-III-P but also including IGFBP-3 (Kniess et al. 2003).

The statistical procedure used to generate the discriminant functions involved splitting the available data into two; a "training" set (used to calculate the discriminant function) and a "confirmatory" set (used to validate the sensitivity and specificity of the discriminant function)(Strike 1991). The confirmatory set was required to ensure the model is applicable to the general population. Ideally, further validation is needed using a completely independent data set.

The aim of this study was to validate the GH-dependent marker approach by assessing whether the GH-2000 and Kreischa formulae could be used reliably to detect those receiving GH when applied to the alternate data set.

6.2 METHODS

The GH-2000 study

The GH-2000 research group undertook a multi-centre, randomised, double blind, placebo controlled study, which has been reported previously (Longobardi et al. 2000; Dall et al. 2000). As the Kreischa study only involved men, only the male formula and male subjects from the GH-2000 study was included in this analysis (Kniess et al. 2003).

Forty-nine healthy male volunteers, (age range 18-35 years, mean 25.6 years, body mass index 19.4–31.7 kg/m^2) were randomly assigned to one of three treatments; placebo (n=21), 0.1 IU/kg body weight/day recombinant GH (rhGH) (n=14), or 0.2 IU/kg/day rhGH (n=14). Treatment was self-administered by subcutaneous injection daily for 28 days. Blood samples were taken fasting at baseline (day 0) and on days 21, 28, 30, 33, 42, and 84 for the measurement of a range of markers including IGF-I and P-III-P(Longobardi et al. 2000; Dall et al. 2000).

The Institut fur Dopinganalytik und Sportbiochemie, Kreischa Study

Fifteen healthy male, non-competitive athletes (age range 21–33 years, mean 24 years, body mass index 23–27 kg/m^2) were assigned randomly to either placebo (n=5) or 0.06 IU/kg body weight /day rhGH (n=10). Treatment was administrated daily by medical staff to the athletes by subcutaneous injection for 14 days. Blood samples were taken for analysis of IGF-I, IGFBP-3 and P-III-P fasting at baseline (day 0) and on days 2,4,6,8,10,12,15 and 18 at 6pm (Kniess et al. 2003).

All subjects gave written informed consent and the studies were conducted in accordance with the ethical principles of the Declaration of Helsinki. The study protocols were approved by the Ethics Committee of the centres involved.

Analytical Methods

In the GH-2000 project, serum IGF-I was analysed after hydrochloric acid-ethanol extraction by radioimmunoassay using authentic IGF-I for labelling (Nichols Institute Diagnostics, San Juan Capisttrano, CA) .

In the Kreischa study, serum IGF-I was analysed after hydrochloric acid-ethanol extraction by competitive fluorescence immunoassay (Kniess et al. 2003).

Discriminant Function Formulae

Each research group used linear discriminant analysis to establish a combination of markers that would distinguish most reliably between placebo and the active treatment groups.

The GH-2000 detection method developed for males was considered in this study. The score was developed from the samples collected from volunteers in the placebo group and those collected at baseline (pre-treatment) from the volunteers in the active treatment group and by definition, had mean of 0 and standard deviation of 1.

The Kreischa group developed a similar prediction formula that included IGF-I, P-III-P and IGFBP-3. The resulting discriminant function calculated was:

$$\textit{Kreischa test score} = -13.465 + 0.0272 \times \textit{IGF-I} + 0.0398 \times \textit{IGFBP-3} + 1.367 \times \textit{P-III-P} - 0.00271 \times (\textit{IGF-I} \times \textit{P-III-P})$$

with the markers expressed in the actual measured scale.

The major differences between the two discriminant functions is the use of logarithmically transformed values by the GH-2000 function (to normalise distribution of data) and the use of IGFBP-3 as an additional marker and the presence of a product of IGF-I and P-III-P in the Kreischa formula.

Statistical Analysis

Adjustment for Assay Differences

Due to the fact that the two research groups used different assay methodology to measure concentrations of IGF-I, P-III-P and IGFBP-3, it was necessary to use statistical adjustment functions which will adjust for the differences in the assay used. This would therefore allow enable the application of the two scoring functions on each dataset. Details of these functions are given in section 2.5.2.

Evaluation of the discriminant functions

The scores calculated by each research group have been standardised such that the placebo and baseline (pre-treatment) samples have a mean value of 0 and standard deviation 1.

These standardised scores of the two discriminant function equations were found to be normally distributed and therefore a score of 3.7 is equivalent to the 99.99% reference limit. The chance of a normal individual having a score of >3.7 is therefore 1 in 10,000. As a cut-off point has not yet been agreed by the World Anti-Doping Agency, the cut-off limits for the detection of GH usage for this study were arbitrarily set at 3.7 (Sonksen 2001), giving the chances of a false positive being less than 1:10,000.

The sensitivity of the functions on a given day has been defined as "the number of observations placed over the 3.7 cut-off value, out of the total number of observations for the particular day from volunteers receiving GH treatment".

6.3 RESULTS

6.3.1 THE GH-2000 DISCRIMINANT FUNCTION

The GH-2000 discriminant function was applied to the data collected by the GH-2000 and Kreischa groups. Table 6-1 presents the sensitivity estimates of this function (and the Kreischa function) when applied to the Kreischa data. A maximum sensitivity of 90% was seen on day 13 of rhGH treatment and 50% of those who had been on rhGH were still identified by the GH-2000 method on Days 16 and 19 (2 and 5 days after the last injection of rhGH). Table 6-2 shows that using the GH-2000 data the GH-2000 function reached a similar sensitivity (86%) on Day 28 while it still identified 54% of those treated with rhGH 2 days after the last injection, 26% after 5 days and 7% after 2 weeks.

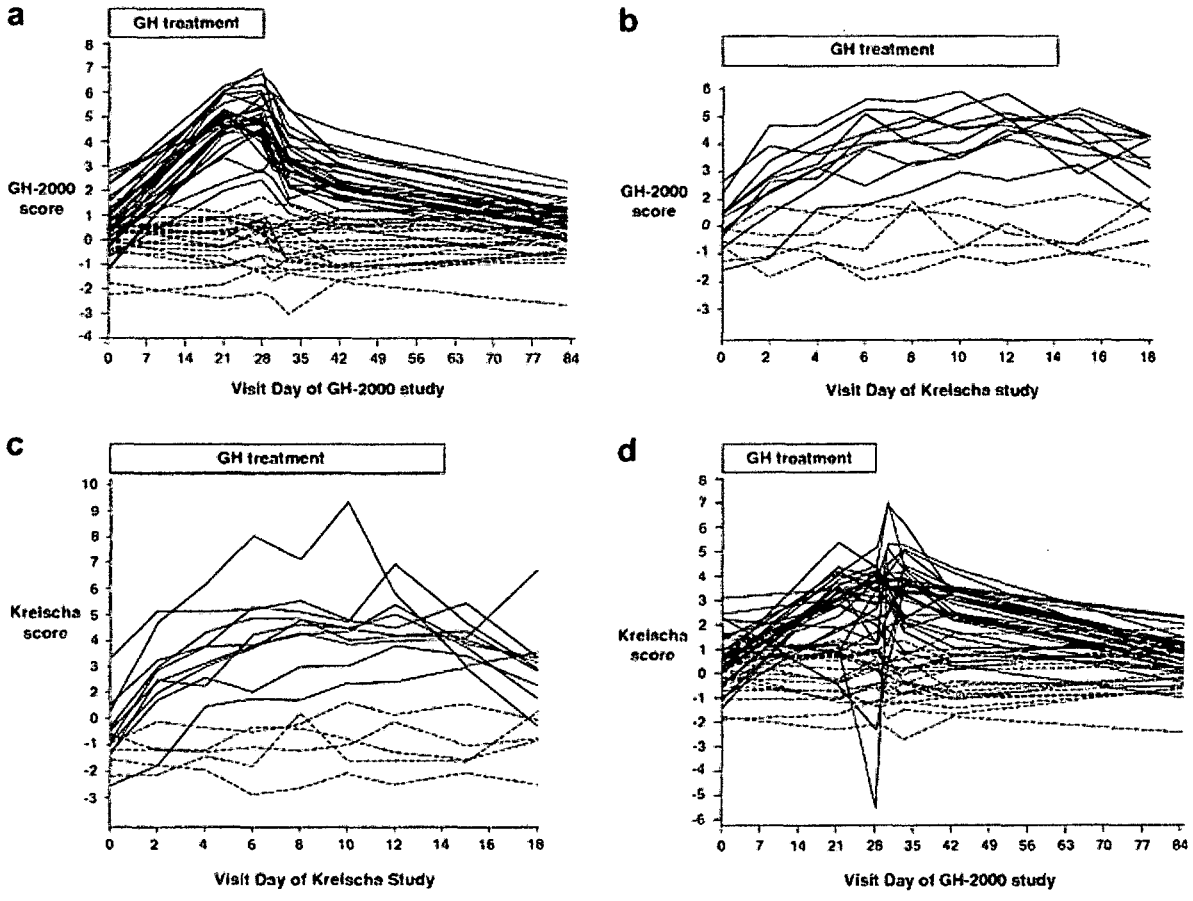

FIGURE 6-1 - TIME TRACES FOR THE TWO DISCRIMINANT FUNCTION FORMULAE WHEN APPLIED TO INDIVIDUAL SUBJECTS FROM THE KREISCHA AND GH-2000 DATASETS. THE SOLID LINE TIME TRACES REPRESENTS VOLUNTEERS ON GROWTH HORMONE TREATMENT. THE DOTTED LINE TIME TRACES REPRESENTS FOR VOLUNTEERS ON THE PLACEBO TREATMENT. (A) THE GH-2000 DISCRIMINANT FUNCTION APPLIED TO THE GH-2000 DATASET. (B) THE GH-2000 DISCRIMINANT FUNCTION APPLIED TO THE KREISCHA DATASET. (C) THE KREISCHA DISCRIMINANT FUNCTION APPLIED TO THE KREISCHA DATASET. (D) THE KREISCHA DISCRIMINANT FUNCTION APPLIED TO THE GH-2000 DATASET.

6.3.2 THE KREISCHA DISCRIMINANT FUNCTION

Table 6-1 presents the sensitivity estimates of the two functions when applied to the Kreischa data. The Kreischa function applied to its own data has a maximum sensitivity of 90% on Day 13, falling to 70% on Day 16 and 10% on Day 19 (2 and 5 days after the last injection of rhGH).

Test Method	Data	Day 3	Day 5	Day 7	Day 9	Day 11	Day 13	Day 16	Day 19
Kreischa	**Kreischa**	2/10 (20%)	5/10 (50%)	8/10 (80%)	8/10 (80%)	8/10 (80%)	9/10 (90%)	7/10 (70%)	1/10 (10%)
GH-2000	**Kreischa**	1/10 (10%)	2/10 (20%)	6/10 (60%)	7/10 (70%)	6/10 (60%)	9/10 (90%)	5/10 (50%)	5/10 (50%)

TABLE 6-1 - SUCCESS RATES AND ESTIMATED SENSITIVITY USING A PRE-DEFINED SPECIFICITY OF 1 IN 10000. SENSITIVITY WAS CALCULATED AS NUMBER OF OBSERVATIONS PLACED OVER THE 3.7 CUT-OFF VALUE OUT OF THE TOTAL NUMBER OF OBSERVATIONS FOR THE PARTICULAR DAY FROM VOLUNTEERS RECEIVING GH TREATMENT. THE TABLE SHOWS THE SENSITIVITY WHEN THE TWO DISCRIMINANT FUNCTION FORMULAE ARE APPLIED TO THE KREISCHA DATASET. SHADED AREA IS TIME OF GH TREATMENT.

Table 6-2 shows that when applied to the GH-2000 data it has a maximum sensitivity of only 36% during rhGH administration (day 28), rising paradoxically to 41% on day 33 (5 days after the last injection of rhGH) and 7% on Day 42 (2 weeks after the last dose of rhGH). Thus, when applied to the GH-2000 data, the Kreischa function has a lower sensitivity than when the function was applied to its own data.

Test Method	Data	Day 21	Day 28	Day 30	Day 33	Day 42	Day 84
Kreischa	**GH-2000**	7/28 (25%)	10/28 (36%)	8/24 (33%)	11/27 (41%)	2/28 (7%)	0/27 (0%)
GH-2000	**GH-2000**	23/28 (82%)	24/28 (86%)	13/24 (54%)	7/27 (26%)	2/28 (7%)	0/27 (0%)

TABLE 6-2 - SUCCESS RATES AND ESTIMATED SENSITIVITY USING A PRE-DEFINED SPECIFICITY OF 1 IN 10000. SENSITIVITY WAS CALCULATED AS NUMBER OF OBSERVATIONS PLACED OVER THE 3.7 CUT-OFF VALUE OUT OF THE TOTAL NUMBER OF OBSERVATIONS FOR THE PARTICULAR DAY FROM VOLUNTEERS RECEIVING GH TREATMENT. THE TABLE SHOWS THE SENSITIVITY WHEN THE TWO FORMULAE ARE APPLIED TO THE GH-2000 DATASET. SHADED AREA IS TIME OF RHGH TREATMENT.

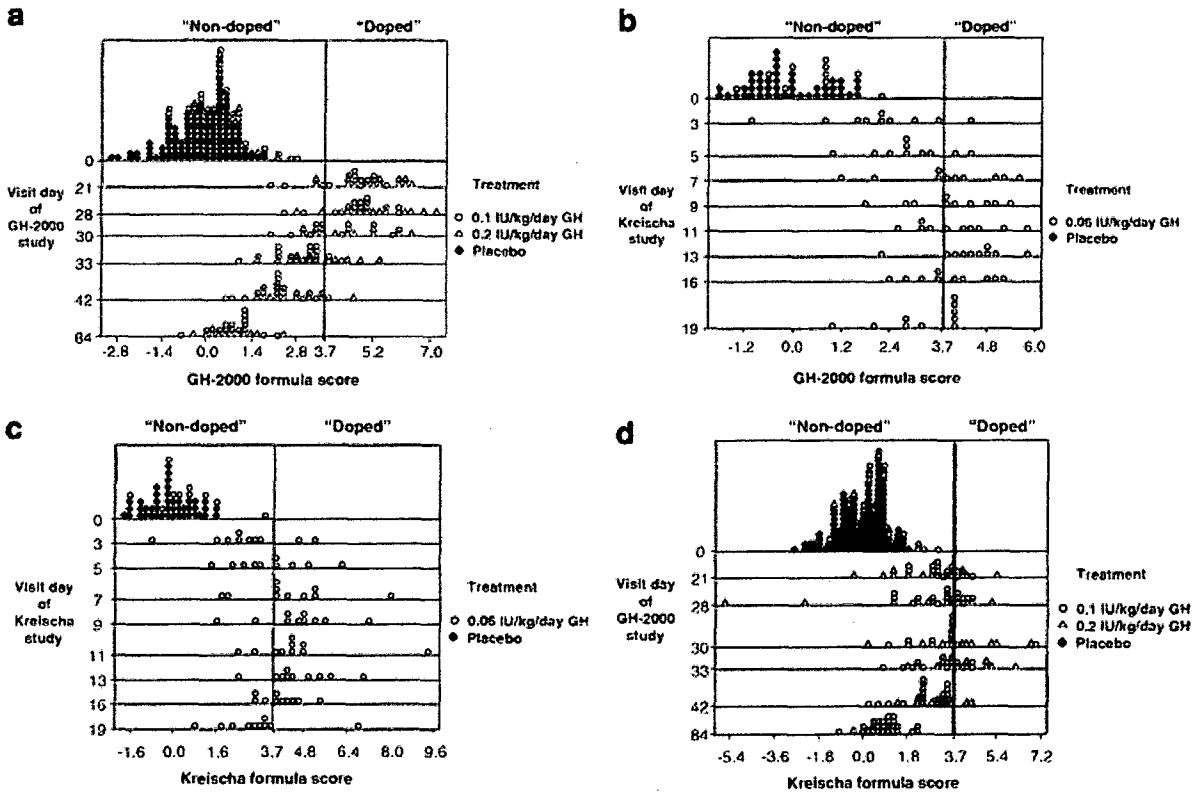

FIGURE 6-2 - DOTPLOT OF THE STANDARDISED SCORES FOR EACH VISIT DAY OF THE STUDIES BY GROUP AND DATASET. THE MEAN OF A NORMAL POPULATION IS 0 AND THE STANDARD DEVIATION IS 1. USING A PRE-DEFINED SENSITIVITY OF 1 IN 10 000, SAMPLES WITH A SCORE OF >3.7 ARE IDENTIFIED AS RECEIVING GH (LABELLED AS "DOPED"). NOTE NONE OF THE BASELINE OR PLACEBO VALUES ARE ABOVE 3.7. (A) THE STANDARDISED GH-2000 SCORE APPLIED TO THE GH-2000 DATASET. (B) THE STANDARDISED GH-2000 SCORE APPLIED TO THE KREISCHA DATASET. (C) THE STANDARDISED KREISCHA SCORE APPLIED TO THE KREISCHA DATASET. (D) THE STANDARDISED KREISCHA SCORE APPLIED TO THE GH-2000 DATASET.

6.4 DISCUSSION

Article 2 of The WADA Anti-Doping Code states:

"Anti-Doping Rule Violations – The following constitute anti-doping rule violations:

2.1 The presence of a Prohibited Substance or its Metabolites or Markers in an Athlete's bodily Specimen." .... constitutes an offence.

As already discussed in considerable detail in previous sections of this thesis, the detection of exogenously administered GH poses a formidable challenge, as it is almost indistinguishable from GH produced naturally from the pituitary gland. Furthermore, the pulsatile secretion of GH leads to wide variations in circulating GH concentrations. Exercise is a potent stimulus for GH secretion thus elevating levels in a post-competition setting (Ehrnborg et al. 2003; Wallace et al. 1999).

The GH-2000 project showed that IGF-I and P-III-P were the simplest and best combination of GH-sensitive markers to distinguish between those receiving GH and those receiving placebo. Both markers increase in a dose dependent manner in response to rhGH administration and are relatively refractory to the acute effects of exercise. The combination of IGF-I and P-III-P is attractive as it combines a hepatic marker of GH action with a collagen marker and therefore reduces the chance of confounding variables affecting the score. The use of P-III-P has the particular advantage of having a much longer half-life than GH thereby lengthening the 'window of opportunity' for detection (Wallace et al. 2000). The Institut fur Dopinganalytik und Sportbiochemie in Kreischa created a second formula based on IGF-I and P-III-P but additionally included IGFBP-3.

Although each formula was devised and validated within their respective datasets, the current study provides further important validation that this marker approach is capable of detecting a proportion of athletes abusing GH during 2-5 days after withdrawal of GH. For the first time we have shown that, the formulae can detect

individuals receiving GH in independent studies with useful sensitivity at a pre-defined specificity of 1 in 10,000.

The formula derived by the GH-2000 study group performs as well, at least in men, in terms of sensitivity when applied to the independent Kreischa dataset as it did on its own data, despite the lower GH doses used by the volunteers in this study. As expected, the score increased over the treatment period, giving good separation from placebo values and remained positive in some cases for as long as 14 days.

The Kreischa formula also performed well when applied to the data obtained from individuals receiving low dose GH in the GH-2000. The Kreischa formula performed less well, however, when applied to those receiving high dose GH and those who had a large increase in markers. This occurred because as P-III-P concentrations rose above a certain figure, the negative product of the term (IGF-I × P-III-P) leads to a fall in score to below the cut-off point and thus false negative results. This is important because it has been suggested that athletes may be currently receiving up to 10 times higher doses than the amount of GH administered in either the GH-2000 or Kreischa studies (Kraemer W. et al. 2002; Kniess et al. 2003). Consequently, the Kreischa function may not be able to detect athletes administering higher GH doses. A further potential disadvantage of the Kreischa formula is the inclusion of IGFBP-3, as anabolic steroids, which may be commonly used with GH, will reverse the GH-induced increase in IGFBP-3 (Karila et al. 1998).

The important conclusion of these studies is that both formulae are able to detect the use of GH in the other's dataset despite the differences in the protocols. Furthermore, reassuringly, neither formula resulted in a false positive when applied to the other dataset. This is of paramount importance to prevent the false conviction and disqualification of any innocent athlete.

A major challenge for the marker approach is to ensure harmonisation between the different assays used to measure P-III-P and IGF-I. Currently there are no established methodologies to adjust the measured P-III-P and IGF-I from one assay to

another. This problem is not insoluble, as a similar problem has arisen for many assays including glycosylated haemoglobin (Goodall 2005). The establishment of international reference material and quality control schemes has led to harmonisation of assays within the clinical arena. As this methodology was not available, adjustment factors were calculated using the values measured for the untreated subjects (the placebo groups and the pre-treatment values from the active treatment groups) by the two research groups. By assuming that there was no difference in the normal ranges for the subjects in Kreischa and the GH-2000 study (because the subjects were Europeans of similar age, body composition, fitness and ethnicity), we were able to use existing statistical methodology to adjust for the P-III-P, IGF-I and IGFBP-3 assays used in the differing studies. It is hoped that similar methodology could be applied in future to any assays used to measure concentrations, thus preventing these formulae from becoming assay-dependent. Any assumptions made in the future using this assay adjustment methodology should ideally be confirmed by direct comparison of the assays where possible, as not all peptide assays would be amenable to this approach.

There are two main limitations to the study. The first is that the Kreischa study included only male subjects. It is known that women are more resistant to the actions of GH and the GH-2000 study has shown that it is more difficult to detect GH abuse in women because of a smaller rise in markers for any given dose of GH (Dall et al. 2000). This may be offset by a need for women to receive higher doses to obtain a performance enhancing benefit. The second limitation is the relatively young age of the subjects in both studies. It is known that GH secretion falls by 14% every decade after a peak in the 20s which is accompanied by a fall in the GH-dependent markers (Juul et al. 1997). The GH-2000 group has proposed an age-adjusted formula to account for the fall in markers in older athletes. It was not possible to validate this adjustment in the current study because of the narrow age range of the subjects in both studies. A further limitation is that this study is unable to show how well these formulae would perform in "real life" where the patterns and doses of GH abused by athletes are unknown.

In summary, the study has provided further validation that the test proposed by GH-2000, based on the measurement of IGF-I and P-III-P can reliably be used to detect subjects receiving exogenous GH.

CHAPTER 7: DISCUSSION

Doping is a serious threat to the integrity of sport and the health of athletes. It is important the sporting regulatory authorities undertake all necessary measures to ensure that the sports arena is clean of any performance enhancing substances.

Developing reliable detection methodology is a lengthy and a very expensive process. However the regulatory authorities have a responsibility that any adopted methodology is reliable and robust and does not falsely accuse any athletes of doping. So despite the fact that a lot of investment goes into developing this detection methodology, it would be inappropriate to put a methodology in place which could not produce a verdict which could be relied on.

A doping offense carries serious consequences for an athlete. At best if an athlete is accused of doping, an athlete's reputation is at risk. This in turn could mean that the athlete might lose the funding they would need to train. At worst, an athlete could be banned from competing for a significant length of time – or even indefinitely.

So it is crucial, that any accusation made for an athlete is reliable and robust, and it is always beyond reasonable doubt. On the other hand however, it is essential that athletes have confidence in the doping system - that it is actually in place and that it does exist. A limit of reliability needs to be determined which when reached, athletes can be declared guilty with a given level of certainty. This result can never be 100% accurate in the case of GH testing just like many other performance enhancing substances because they occur naturally in the body system already – such as in the case of gene doping for example. So to be able to detect these substances, it is necessary to accept that a level of uncertainty will need to be involved when testing for these substances.

To ensure that any testing methodology is robust and reliable it is necessary that all measurement uncertainty is taken into account. This will in turn reduce the number of false positives. In order to be able to do this, it is essential that all sources of error are controlled for thus ensuring consistent performance of the test. Uncertainty is an unavoidable part of any measurement and this will never be reduced to zero levels. However, uncertainty becomes significant when a test is near to the cut-off limit for declaring a positive result.

When the results of the GH-2000 study were presented on the proposed GH detection methodology at the IOC organised Workshop in Rome in April 1999, a number of concerns were raised regarding the reliability of this methodology. There were two key suggested areas for further investigation: the effect of sporting injury and the effect of ethnicity on the reliability of this proposed methodology. The GH-2004 study was then set up to undertake further research to examine the effect of each of these, as well as the effect of physiological intra-individual variability on the reliability of this test.

This thesis has examined these areas, and has shown that the effect of these variables is not significant on the performance of the GH-2000 detection method. The key message from the analysis and the results described in this thesis is that the proposed GH-2000 detection methodology is a robust and reliable methodology for detecting GH abuse in sports with minimal risk of false positive results but also with good levels of sensitivity at the same time.

However, there are some limitations to this study, which will need to be addressed, before any results can be used to prosecute an athlete with confidence in the court of law.

The current limitation of this thesis is the need to carry out further assay validation work before this test can be officially implemented at future Olympic Games. In the event where an immunoassay is used for the screening testing process, then the immunoassay to be used for confirmation is required to be use a different antibody

that will recognise a different epitope of the peptide/protein than the assay used for screening.

The future validation work will need to cover three key areas:

1. Determine which pair of immunoassays for each of the two biomarkers; IGF-I and P-III-P will be used in a testing situation. This work is necessary to be able to provide WADA with recommendations for the best immunoassays to be used, which will be able to measure IGF-I and P-III-P measurements reliably with high reproducibility estimates.
2. Perform validation work for each pair of immunoassays to determine assay conversion equations, to enable direct comparison with each other, but also comparison with the GH-2000 assays, since this is the methodology used by the GH-2000 discriminant functions. Furthermore, the error introduced through the use of assay conversion equations needs to be quantified and be taken into account when determining the cut-off point of the GH-2000 detection methodology.
3. Perform inter-laboratory validation to ensure that any inter-laboratory error due to different equipment and operators would adversely affect the outcome of the detection method.

This laboratory analysis is important for two reasons. Firstly, this assay validation work needs to be undertaken to ensure that the results derived from any assay methodology are robust and reliable and therefore would not bias the outcome of any GH test. Secondly, it is known that for each of these three areas mentioned, there will be a measurement uncertainty associated with it. This is an important step in ensuring that all sources of uncertainty are accounted for.

This can be done either at the stage of determining the cut-off limit the GH-2000 detection method or when measuring the test samples. Every lab performing the assays must determine their measurement uncertainty, which must be within the values of the expected measurement uncertainty. The expected measurement

uncertainty can either be used to determine a universal cut-off limit, or for each lab to have their own cut-off limit.

This stage of the laboratory analysis will enable the quantification of all measurement uncertainty and will therefore allow us to take this into account when devising the cut-off point for the GH-2000 detection method in a testing. Having done this, a new calculated cut-off point can be calculated which will now take into account the effect of all measurement uncertainty.

This laboratory validation work is essential before proceeding with the official implementation of the test at a sporting event. However, there are further areas for further enhancement of the GH-2000 detection test. These areas mainly concern further validation of the GH-2000 detection methodology.

It is known during adolescence, healthy individuals have higher IGF-I values than adults, however the relationship with age in these younger ages is different compared to adults (Juul et al. 1994; Juul 2001; Rogol 2003). From birth, IGF-I concentrations increase linearly with age until they reach a peak at approximately the age of 18 for males and 16 for females (Juul et al. 1994). Furthermore, it has been shown by Healy et al (Healy et al. 2005) that elite athletes have significantly higher values than healthy normal individuals. Therefore, it is likely that for adolescent athletes the IGF-I physiological reference ranges established by the GH-2000 group and validated in this study will not be applicable to this age group of athletes. This is because the GH-2000 study did not recruit sufficient numbers of adolescent elite athletes. In addition, there is currently no adequate research available for P-III-P concentrations for adolescent subjects. Before being able to use this detection methodology in adolescent athletes it will be necessary to ensure that they are at no increased risk for a false positive conclusion. However, it is just as important that the methodology will be sensitive enough to be able to detect young individuals abusing GH.

A third area for further research would be to validate the female discriminant function in an independently collected sample from a GH administration study. This

validation has already been undertaken for the male discriminant function and has been discussed in this thesis.

A recent GH administration study undertaken recently by K. Ho et al involving male and female young healthy volunteers can provide the opportunity for further validation of the GH-2000 detection method in both genders (Blackman et al. 2007). In addition, this study also involved an additional treatment arm, which included testosterone administration as well as GH. Validation of the GH-2000 detection method using the data collected from this study will be very important, as it is anecdotally known that athletes will often abuse GH in combination with testosterone.

Despite the fact that the discriminant functions derived have been validated in a previously unused part of the sample collected by the GH-2000 GH administration study, there is still a positive bias in these validation results. This is because the samples would have been collected from the study where exactly the same study protocol and design was followed.

This work is important and it would confirm even further the validity of the GH-2000 detection method. This would be of great value particularly if the study to be considered had different study protocol and design to the one used in the GH-2000 study. For example if different assay methodology was used, and/or different GH dose, this would provide further an additional testing environment under which the testing method can be tested.

A fourth area for further research is the need to validate the proposed GH detection methodology in samples where GH was co-administered using other performance enhancing substances. It is anecdotally well known that in real life situations, athletes who take GH will often so in combination with other performance enhancing substances. A common performance enhancing substance used concomitantly with GH is testosterone. Unpublished results (Nelson AE et al. 2008) has shown that the effect on IGF-I and P-III-P is further enhanced with concomitant use of testosterone and

GH. It would be valuable to perform a further validation of the testing method in this environment, to illustrate how this would affect the sensitivity of the method as well as the risk of a false positive result.

Despite these four issues discussed in this section, the GH detection methodology has already been rigorously validated in a variety of settings and it has proven to be a reliable and robust methodology. Any further work on this is only expected to enhance the strength of these conclusions.

This thesis set out to assess whether the GH detection methodology proposed by the GH-2000 group back in 1999 is a reliable and robust methodology to be used in sports to detect GH abuse in sport. It has been shown that indeed this methodology is a powerful method and should help to cut down the abuse of GH, which carries so many health risks and its use is unfair to competitor athletes.

Before this method is officially implemented there is important laboratory analysis which needs to be undertaken, which will enable us to determine reliable cut-off limits which will both maximise the sensitivity of the test but also reduce the rate of false positive results.

Nevertheless, apart from these logistical issues, the proposed GH detection methodology has been shown to be reliable and a promising method of reducing GH abuse in sports.

REFERENCES

Abellan, R., Ventura, R., Pichini, S., Pascual, J. A., Pacifici, R., Di, C. S., Bacosi, A., Segura, J., & Zuccaro, P. 2005, "Evaluation of immunoassays for the measurement of insulin-like growth factor-I and procollagen type III peptide, indirect biomarkers of recombinant human growth hormone misuse in sport", *Clin.Chem.Lab Med.*, vol. 43, no. 1, pp. 75-85.

Adidas. Adidas official website. www.adidas.com . 26-3-2006. 26-3-2006.
Ref Type: Electronic Citation

Allen, N. E., Roddam, A. W., Allen, D. S., Fentiman, I. S., Dos, S. S., I, Peto, J., Holly, J. M., & Key, T. J. 2005, "A prospective study of serum insulin-like growth factor-I (IGF-I), IGF-II, IGF-binding protein-3 and breast cancer risk", *Br.J.Cancer*, vol. 92, no. 7, pp. 1283-1287.

Armanini, D., Faggian, D., Scaroni, C., & Plebani, M. 2002, "Growth hormone and insulin-like growth factor I in a Sydney Olympic gold medallist", *Br.J.Sports Med.*, vol. 36, no. 2, pp. 148-149.

Bagg, W., Aoina, J., Cross, P. A., Whalley, G. A., Gamble, G. D., Doughty, R. N., & Holdaway, I. M. 2006, "Serum IGF-I levels are similar in Samoan, Maori and European populations despite differences in body composition", *Growth Horm.IGF.Res.*, vol. 16, no. 1, pp. 57-60.

Bail, H. J., Kolbeck, S., Lindner, T., Dahne, M., Weiler, A., Windhagen, H. J., Raun, K., Skjaerbaek, C., Flyvbjerg, A., Orskov, H., Haas, N. P., & Raschke, M. J. 2001, "The effect of growth hormone on insulin-like growth factor I and bone metabolism in distraction osteogenesis", *Growth Horm.IGF.Res.*, vol. 11, no. 5, pp. 314-323.

Banfi, G., Marinelli, M., Roi, G. S., Colombini, A., Pontillo, M., Giacometti, M., & Wade, S. 1994, "Growth hormone and insulin-like growth factor I in athletes performing a marathon at 4000 m of altitude", *Growth Regul.*, vol. 4, no. 2, pp. 82-86.

Barkan, A. L. 2004, "Biochemical markers of acromegaly: GH vs. IGF-I", *Growth Horm.IGF Res.*, vol. 14 Suppl A, pp. S97-100.

Barroso, O., Mazzoni, I., & Rabin, O. 2008, "Hormone abuse in sports: the antidoping perspective", *Asian J Androl*, vol. 10, no. 3, pp. 391-402.

Berggren, A., Ehrnborg, C., Rosen, T., Ellegard, L., Bengtsson, B. A., & Caidahl, K. 2005, "Short-term administration of supraphysiological recombinant human growth hormone (GH) does not increase maximum endurance exercise capacity in healthy,

active young men and women with normal GH-insulin-like growth factor I axes", *J.Clin.Endocrinol.Metab*, vol. 90, no. 6, pp. 3268-3273.

Bidlingmaier, M., Wu, Z., & Strasburger, C. J. 2000, "Test method: GH", *Baillieres Best.Pract.Res.Clin.Endocrinol.Metab*, vol. 14, no. 1, pp. 99-109.

Bidlingmaier, M., Wu, Z., & Strasburger, C. J. 2003, "Problems with GH doping in sports", *J.Endocrinol.Invest*, vol. 26, no. 9, pp. 924-931.

Blackman, M., Ho, K. K., & Nelson, A. E. "Testosterone and Growth Hormone Replacement: Impact on Health and Fitness".

Bollerslev, J., Hallen, J., Fougner, K. J., Jorgensen, A. P., Kristo, C., Fagertun, H., Gudmundsen, O., Burman, P., & Schreiner, T. 2005, "Low-dose GH improves exercise capacity in adults with GH deficiency: effects of a 22-month placebo-controlled, crossover trial", *Eur J Endocrinol*, vol. 153, no. 3, pp. 379-387.

Bollerslev, J., Moller, J., Thomas, S., Djoseland, O., & Christiansen, J. S. 1996, "Dose-dependent effects of recombinant human growth hormone on biochemical markers of bone and collagen metabolism in adult growth hormone deficiency", *Eur.J.Endocrinol.*, vol. 135, no. 6, pp. 666-671.

Bristow A.F. 1998, *Standardisation of protein hormone immunoassays: current controversies, EQAS'98 Participants Meeting.*

Brown, P., Brandel, J. P., Preece, M., & Sato, T. 2006, "Iatrogenic Creutzfeldt-Jakob disease: the waning of an era", *Neurology*, vol. 67, no. 3, pp. 389-393.

Brown-Sequard C.E. 1889, "The effects produced in man by subcutaneous injections of a liquid obtained from the testicles of animals", *Lancet*, vol. 2, pp. 105-107.

Bryn Palmer. The dirtiest Games ever? www.bbc.co.uk . 30-8-2004.
Ref Type: Electronic Citation

Butler, A. A. & Le, R. D. 2001, "Control of growth by the somatropic axis: growth hormone and the insulin-like growth factors have related and independent roles", *Annu.Rev.Physiol*, vol. 63, pp. 141-164.

Buzzini, S. R. 2007, "Abuse of growth hormone among young athletes", *Pediatr.Clin North Am.*, vol. 54, no. 4, pp. 823-43, xiii.

Chung, L., Clifford, D., Buckley, M., & Baxter, R. C. 2006, "Novel biomarkers of human growth hormone action from serum proteomic profiling using protein chip mass spectrometry", *J.Clin.Endocrinol.Metab*, vol. 91, no. 2, pp. 671-677.

CIS BioInternational. RIA-gnost-PIIIP: Protocol and results.
http://www.cisbiointernational.fr/ria/ressources/produits/AG_OCFK07_PIIIP_V3.pdf .

1-10-2007.
Ref Type: Electronic Citation

Colao, A., Ferone, D., Marzullo, P., & Lombardi, G. 2004, "Systemic complications of acromegaly: epidemiology, pathogenesis, and management", *Endocr.Rev.*, vol. 25, no. 1, pp. 102-152.

Crist, D. M., Peake, G. T., Egan, P. A., & Waters, D. L. 1988, "Body composition response to exogenous GH during training in highly conditioned adults", *J.Appl.Physiol*, vol. 65, no. 2, pp. 579-584.

Csaky, T. Z. 1972, "Doping", *J Sports Med Phys.Fitness*, vol. 12, no. 2, pp. 117-123.

Cuneo, R. C., Salomon, F., McGauley, G. A., & Sonksen, P. H. 1992, "The growth hormone deficiency syndrome in adults", *Clin Endocrinol (Oxf)*, vol. 37, no. 5, pp. 387-397.

Cuneo, R. C., Salomon, F., Wiles, C. M., Hesp, R., & Sonksen, P. H. 1991, "Growth hormone treatment in growth hormone-deficient adults. I. Effects on muscle mass and strength", *Journal of Applied Physiology*, vol. 70, no. 2, pp. 688-694.

Cuneo, R. C., Salomon, F., Wiles, C. M., & Sonksen, P. H. 1990, "Skeletal muscle performance in adults with growth hormone deficiency", *Horm.Res.*, vol. 33 Suppl 4, pp. 55-60.

Cushing, H. 1909, "III. Partial Hypophysectomy for Acromegaly: With Remarks on the Function of the Hypophysis", *Ann.Surg.*, vol. 50, no. 6, pp. 1002-1017.

Dall, R., Longobardi, S., Ehrnborg, C., Keay, N., Rosen, T., Jorgensen, J. O., Cuneo, R. C., Boroujerdi, M. A., Cittadini, A., Napoli, R., Christiansen, J. S., Bengtsson, B. A., Sacca, L., Baxter, R. C., Basset, E. E., & Sonksen, P. H. 2000, "The effect of four weeks of supraphysiological growth hormone administration on the insulin-like growth factor axis in women and men. GH-2000 Study Group", *J.Clin.Endocrinol.Metab*, vol. 85, no. 11, pp. 4193-4200.

De Palo, E. F., Gatti, R., Lancerin, F., Cappellin, E., & Spinella, P. 2001, "Correlations of growth hormone (GH) and insulin-like growth factor I (IGF-I): effects of exercise and abuse by athletes", *Clin.Chim.Acta*, vol. 305, no. 1-2, pp. 1-17.

de, B. H., Blok, G. J., & Van, d., V 1995, "Clinical aspects of growth hormone deficiency in adults", *Endocr.Rev.*, vol. 16, no. 1, pp. 63-86.

Dean, H. 2002, "Does exogenous growth hormone improve athletic performance?", *Clin.J.Sport Med.*, vol. 12, no. 4, pp. 250-253.

Defrancesco, L. 2004, "The faking of champions", *Nat.Biotechnol.*, vol. 22, no. 9, pp. 1069-1071.

Dekkers, O. M., Biermasz, N. R., Pereira, A. M., Romijn, J. A., & Vandenbroucke, J. P. 2008, "Mortality in acromegaly: a metaanalysis", *J Clin Endocrinol Metab*, vol. 93, no. 1, pp. 61-67.

Delellis, K., Ingles, S., Kolonel, L., Kean-Cowdin, R., Henderson, B., Stanczyk, F., & Probst-Hensch, N. M. 2003, "IGF1 genotype, mean plasma level and breast cancer risk in the Hawaii/Los Angeles multiethnic cohort", *Br.J.Cancer*, vol. 88, no. 2, pp. 277-282.

Deyssig, R., Frisch, H., Blum, W. F., & Waldhor, T. 1993, "Effect of growth hormone treatment on hormonal parameters, body composition and strength in athletes", *Acta Endocrinol.(Copenh)*, vol. 128, no. 4, pp. 313-318.

Di Pasquale, M. G. 1990, "Growth Hormone," in *Beyond Anabolic Steroids*, MGD press, pp. 6-16.

Dickerman, R. D., Douglas, J. A., & East, J. W. 2000, "Bilateral median neuropathy and growth hormone use: a case report", *Arch.Phys.Med.Rehabil.*, vol. 81, no. 12, pp. 1594-1595.

Doessing, S. & Kjaer, M. 2005, "Growth hormone and connective tissue in exercise", *Scand.J.Med.Sci.Sports*, vol. 15, no. 4, pp. 202-210.

Donahue, L. R., Hunter, S. J., Sherblom, A. P., & Rosen, C. 1990, "Age-related changes in serum insulin-like growth factor-binding proteins in women", *J Clin Endocrinol Metab*, vol. 71, no. 3, pp. 575-579.

Donohoe, T. 1986, *Foul play : drug abuse in sports* Blackwell, Oxford.

Dubin C.J. Commision of inquiry into the use of drugs and banned practices intended to increase athletic performance. 1990. Ottawa, ON, Canadian Publishing Center. Ref Type: Serial (Book,Monograph)

Duchaine D. 1982, *Underground Steroid Handbook* OEM, Santa Monica CA.

Ducommum, P. & DUCOMMUM, S. 1953, "[The effect of pituitary somatotropin in the prevention of infections abetted by overdosage of cortisone.]", *Ann.Endocrinol.(Paris)*, vol. 14, no. 5, pp. 765-771.

Dunn G. & Everitt, B. S. 2001, *Applied Multivariate Data Analysis*, 2 edn.

Dunn, G. A. 2004, *Statistical evaluation of measurement errors : design and analysis of reliability studies* Arnold, London.

Ehrnborg, C., Lange, K. H., Dall, R., Christiansen, J. S., Lundberg, P. A., Baxter, R. C., Boroujerdi, M. A., Bengtsson, B. A., Healey, M. L., Pentecost, C., Longobardi, S., Napoli, R., & Rosen, T. 2003, "The growth hormone/insulin-like growth factor-I axis

hormones and bone markers in elite athletes in response to a maximum exercise test", *J.Clin.Endocrinol.Metab*, vol. 88, no. 1, pp. 394-401.

Eliakim, A., Brasel, J. A., Barstow, T. J., Mohan, S., & Cooper, D. M. 1998, "Peak oxygen uptake, muscle volume, and the growth hormone-insulin-like growth factor-I axis in adolescent males", *Med.Sci.Sports Exerc.*, vol. 30, no. 4, pp. 512-517.

Erotokritou-Mulligan, I., Bassett, E. E., Kniess, A., Sonksen, P. H., & Holt, R. I. 2007, "Validation of the growth hormone (GH)-dependent marker method of detecting GH abuse in sport through the use of independent data sets", *Growth Horm.IGF Res.*, vol. 17, no. 5, pp. 416-423.

Evans, M. J., Livesey, J. H., Ellis, M. J., & Yandle, T. G. 2001, "Effect of anticoagulants and storage temperatures on stability of plasma and serum hormones", *Clin Biochem.*, vol. 34, no. 2, pp. 107-112.

Evans, N. A. 2004, "Current concepts in anabolic-androgenic steroids", *Am.J.Sports Med.*, vol. 32, no. 2, pp. 534-542.

Faria, A. C., Bekenstein, L. W., Booth, R. A., Jr., Vaccaro, V. A., Asplin, C. M., Veldhuis, J. D., Thorner, M. O., & Evans, W. S. 1992, "Pulsatile growth hormone release in normal women during the menstrual cycle", *Clin Endocrinol (Oxf)*, vol. 36, no. 6, pp. 591-596.

Felsing, N. E., Brasel, J. A., & Cooper, D. M. 1992, "Effect of low and high intensity exercise on circulating growth hormone in men", *J.Clin.Endocrinol.Metab*, vol. 75, no. 1, pp. 157-162.

Franke, W. W. & Berendonk, B. 1997, "Hormonal doping and androgenization of athletes: a secret program of the German Democratic Republic government", *Clin Chem.*, vol. 43, no. 7, pp. 1262-1279.

Fryburg, D. A. & Barrett, E. J. 1993, "Growth hormone acutely stimulates skeletal muscle but not whole-body protein synthesis in humans", *Metabolism*, vol. 42, no. 9, pp. 1223-1227.

Fryburg, D. A., Gelfand, R. A., & Barrett, E. J. 1991, "Growth hormone acutely stimulates forearm muscle protein synthesis in normal humans", *Am.J.Physiol*, vol. 260, no. 3 Pt 1, p. E499-E504.

Galbo, H. 1982, "Endocrinology and metabolism in exercise", *Curr.Probl.Clin.Biochem.*, vol. 11, pp. 26-44.

GH-2000 group. Final Report to International Olympic Committee and European Union. www.gh2004.soton.ac.uk . 1999. 29-9-2006.
Ref Type: Electronic Citation

Giannoulis, M. G., Sonksen, P. H., Umpleby, M., Breen, L., Pentecost, C., Whyte, M., McMillan, C. V., Bradley, C., & Martin, F. C. 2006, "The effects of growth hormone and/or testosterone in healthy elderly men: a randomized controlled trial", *J.Clin.Endocrinol.Metab*, vol. 91, no. 2, pp. 477-484.

Gibney, J., Healy, M. L., & Sonksen, P. H. 2007, "The growth hormone/insulin-like growth factor-I axis in exercise and sport", *Endocr.Rev.*, vol. 28, no. 6, pp. 603-624.

Goodall, I. 2005, "HbA1c standardisation destination--global IFCC Standardisation. How, why, where and when--a tortuous pathway from kit manufacturers, via inter-laboratory lyophilized and whole blood comparisons to designated national comparison schemes", *Clin Biochem.Rev.*, vol. 26, no. 1, pp. 5-19.

Graham M.R., Baker J.S., Evans P., Kicman, A. T., Cowan, D. A., & Davies, B. 2008, "Physical effects of short term GH administration in abstinent Steroid dependency", *Horm.Res.*, vol. In Press.

Haisma, H. J., De Hon, O., Sollie, P., & Vorstenbosch, J. 2004, *Gene Doping - Netherlands Centre for Doping Affairs*.

Hanif, M., Patel, A., & Dunning, J. 2005, "Might gene therapy offer symptomatic relief for patients with 'no option' angina?", *Interact.Cardiovasc.Thorac.Surg.*, vol. 4, no. 6, pp. 627-632.

Hartman, M. L., Veldhuis, J. D., Johnson, M. L., Lee, M. M., Alberti, K. G., Samojlik, E., & Thorner, M. O. 1992, "Augmented growth hormone (GH) secretory burst frequency and amplitude mediate enhanced GH secretion during a two-day fast in normal men", *J Clin Endocrinol Metab*, vol. 74, no. 4, pp. 757-765.

Harvey, S. C. 1995, *Growth hormone Scanes, William H. Daughaday* CRC, Boca Raton ; London.

Hashimoto, Y., Kamioka, T., Hosaka, M., Mabuchi, K., Mizuchi, A., Shimazaki, Y., Tsunoo, M., & Tanaka, T. 2000, "Exogenous 20K growth hormone (GH) suppresses endogenous 22K GH secretion in normal men", *J.Clin.Endocrinol.Metab*, vol. 85, no. 2, pp. 601-606.

Haukipuro, K., Risteli, L., Kairaluoma, M. I., & Risteli, J. 1990, "Aminoterminal propeptide of type III procollagen in serum during wound healing in human beings", *Surgery*, vol. 107, no. 4, pp. 381-388.

Haycock, B. Growing Beyond What Nature Intended - Growth Factors Take us to the Next Level of Sophistication in Bodybuilding (Part III: GH, IGF-1, Insulin, and Thryoid to Enhance Anabolic Effects of Androgens). http://www.mesomorphosis.com/articles/haycock/growth-factors-03.htm . 20-3-2006. Ref Type: Electronic Citation

Healy, M. L., Dall, R., Gibney, J., Bassett, E., Ehrnborg, C., Pentecost, C., Rosen, T., Cittadini, A., Baxter, R. C., & Sonksen, P. H. 2005, "Toward the development of a test for growth hormone (GH) abuse: a study of extreme physiological ranges of GH-dependent markers in 813 elite athletes in the postcompetition setting", *J.Clin.Endocrinol.Metab*, vol. 90, no. 2, pp. 641-649.

Healy, M. L., Gibney, J., Russell-Jones, D. L., Pentecost, C., Croos, P., Sonksen, P. H., & Umpleby, A. M. 2003, "High dose growth hormone exerts an anabolic effect at rest and during exercise in endurance-trained athletes", *J.Clin.Endocrinol.Metab*, vol. 88, no. 11, pp. 5221-5226.

Healy, M. L. & Russell-Jones, D. 1997, "Growth hormone and sport: abuse, potential benefits, and difficulties in detection", *Br.J.Sports Med.*, vol. 31, no. 4, pp. 267-268.

Heinemann, U., Krasnianski, A., Meissner, B., Varges, D., Kallenberg, K., Schulz-Schaeffer, W. J., Steinhoff, B. J., Grasbon-Frodl, E. M., Kretzschmar, H. A., & Zerr, I. 2007, "Creutzfeldt-Jakob disease in Germany: a prospective 12-year surveillance", *Brain*, vol. 130, no. Pt 5, pp. 1350-1359.

Hertel, N. T., Stoltenberg, M., Juul, A., Main, K. M., Muller, J., Nielsen, C. T., Lorenzen, I., & Skakkebaek, N. E. 1993, "Serum concentrations of type I and III procollagen propeptides in healthy children and girls with central precocious puberty during treatment with gonadotropin-releasing hormone analog and cyproterone acetate", *J.Clin.Endocrinol.Metab*, vol. 76, no. 4, pp. 924-927.

Hirst, C. 2005, "Iatrogenic Creutzfeldt-Jakob disease presenting 24 years after human growth hormone administration", *Br J Hosp.Med (Lond)*, vol. 66, no. 10, pp. 592-593.

Hoberman, J. M. 1992, *Mortal engines : the science of performance and the dehumanization of sport Hoberman* Free Press ; Toronto : Maxwell Macmillan Canada ; New York : Maxwell Macmillan International, New York.

Holt, R. I., Erotokritou-Mulligan, I., McHugh, C. M., Cowan, D. A., Bartlett C, Basset, E. E., & Sonksen, P. H. 2006, "The GH-2004 project: the use of growth hormone (GH)-dependent markers in the detection of GH abuse in sport. In: Schänzer, W., Geyer, H., Gotzmann, A., Mareck, U. (eds.) Recent Advances in Doping Analysis", Schänzer W Geyer H Gotzmann A Mareck, ed..

Holt, R. I. & Sonksen, P. H. 2008, "Growth hormone, IGF-I and insulin and their abuse in sport", *Br J Pharmacol.*

Holt, R. I. G. 2002, "Fetal programming of the growth hormone-insulin-like growth factor axis", *Trends in Endocrinology and Metabolism*, vol. 13, no. 9, pp. 392-397.

Iranmanesh, A., Lizarralde, G., & Veldhuis, J. D. 1991, "Age and relative adiposity are specific negative determinants of the frequency and amplitude of growth hormone

(GH) secretory bursts and the half-life of endogenous GH in healthy men", *J Clin Endocrinol Metab*, vol. 73, no. 5, pp. 1081-1088.

Ivan, D., Brabant, G., & Kann, P. H. 2005, "Applicability of recently established reference values for serum insulin-like growth factor 1: A comparison of two assays--an (automated) chemiluminescence immunoassay and an enzyme-linked immunosorbent assay", *Clin Lab*, vol. 51, no. 7-8, pp. 381-387.

Janssen, J. A., Stolk, R. P., Pols, H. A., Grobbee, D. E., de Jong, F. H., & Lamberts, S. W. 1998, "Serum free IGF-I, total IGF-I, IGFBP-1 and IGFBP-3 levels in an elderly population: relation to age and sex steroid levels", *Clin Endocrinol (Oxf)*, vol. 48, no. 4, pp. 471-478.

Janssen, Y. J., Doornbos, J., & Roelfsema, F. 1999, "Changes in muscle volume, strength, and bioenergetics during recombinant human growth hormone (GH) therapy in adults with GH deficiency", *J.Clin.Endocrinol.Metab*, vol. 84, no. 1, pp. 279-284.

Jenkins, P. J. 1999, "Growth hormone and exercise", *Clin.Endocrinol.(Oxf)*, vol. 50, no. 6, pp. 683-689.

Jenkins, P. J. 2001, "Growth hormone and exercise: physiology, use and abuse", *Growth Horm.IGF.Res.*, vol. 11 Suppl A, p. S71-S77.

Joerring, S., Krogsgaard, M., Wilbek, H., & Jensen, L. T. 1994, "Collagen turnover after tibial fractures", *Arch.Orthop.Trauma Surg.*, vol. 113, no. 6, pp. 334-336.

Jokl, E. 1968, "Notes on doping," in *Exercise and Altitude*, Karger, Basel.

Jorgensen, J. O., Moller, J., Wolthers, T., Vahl, N., Juul, A., Skakkebaek, N. E., & Christiansen, J. S. 1994a, "Growth hormone (GH)-deficiency in adults: clinical features and effects of GH substitution", *J Pediatr.Endocrinol*, vol. 7, no. 4, pp. 283-293.

Jorgensen, J. O., Muller, J., Moller, J., Wolthers, T., Vahl, N., Juul, A., Skakkebaek, N. E., & Christiansen, J. S. 1994b, "Adult growth hormone deficiency", *Horm.Res.*, vol. 42, no. 4-5, pp. 235-241.

Jorgensen, J. O., Pedersen, S. A., Thuesen, L., Jorgensen, J., Ingemann-Hansen, T., Skakkebaek, N. E., & Christiansen, J. S. 1989, "Beneficial effects of growth hormone treatment in GH-deficient adults", *Lancet*, vol. 1, no. 8649, pp. 1221-1225.

Juul, A. 2001, "Determination of insulin-like growth factor I in children: normal values and clinical use", *Horm.Res.*, vol. 55 Suppl 2, pp. 94-99.

Juul, A. 2005, "IGF-I measurements. Epidemiological and clinical aspects", *Front Horm.Res.*, vol. 33, pp. 45-67.

Juul, A., Bang, P., Hertel, N. T., Main, K., Dalgaard, P., Jorgensen, K., Muller, J., Hall, K., & Skakkebaek, N. E. 1994, "Serum insulin-like growth factor-I in 1030 healthy children, adolescents, and adults: relation to age, sex, stage of puberty, testicular size, and body mass index", *J.Clin.Endocrinol.Metab*, vol. 78, no. 3, pp. 744-752.

Juul, A., Scheike, T., Pedersen, A. T., Main, K. M., Andersson, A. M., Pedersen, L. M., & Skakkebaek, N. E. 1997, "Changes in serum concentrations of growth hormone, insulin, insulin-like growth factor and insulin-like growth factor-binding proteins 1 and 3 and urinary growth hormone excretion during the menstrual cycle", *Hum.Reprod.*, vol. 12, no. 10, pp. 2123-2128.

Kam, G. Y., Leung, K. C., Baxter, R. C., & Ho, K. K. 2000, "Estrogens exert route- and dose-dependent effects on insulin-like growth factor (IGF)-binding protein-3 and the acid-labile subunit of the IGF ternary complex", *J.Clin.Endocrinol.Metab*, vol. 85, no. 5, pp. 1918-1922.

Kanaley, J. A., Weltman, J. Y., Veldhuis, J. D., Rogol, A. D., Hartman, M. L., & Weltman, A. 1997, "Human growth hormone response to repeated bouts of aerobic exercise", *J.Appl.Physiol*, vol. 83, no. 5, pp. 1756-1761.

Karila, T., Koistinen, H., Seppala, M., Koistinen, R., & Seppala, T. 1998, "Growth hormone induced increase in serum IGFBP-3 level is reversed by anabolic steroids in substance abusing power athletes", *Clin.Endocrinol.(Oxf)*, vol. 49, no. 4, pp. 459-463.

Kay, M. A., Manno, C. S., Ragni, M. V., Larson, P. J., Couto, L. B., McClelland, A., Glader, B., Chew, A. J., Tai, S. J., Herzog, R. W., Arruda, V., Johnson, F., Scallan, C., Skarsgard, E., Flake, A. W., & High, K. A. 2000, "Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector", *Nat.Genet.*, vol. 24, no. 3, pp. 257-261.

Kearns, B., Harkness, R., Hobson, V., & Smith, A. 1942, "Testosterone pellet implantation in the gelding", *Journal of the American Veterinary Medicine Association*, vol. C/780, pp. 197-201.

Kent M. Food and Fitness: A Dictionary of Diet and Exercise. Oxford Reference Online . 1997. Oxford University Press. 26-3-2006.
Ref Type: Electronic Citation

Khosravi, J., Diamandi, A., Bodani, U., Khaja, N., & Krishna, R. G. 2005, "Pitfalls of immunoassay and sample for IGF-I: comparison of different assay methodologies using various fresh and stored serum samples", *Clin.Biochem.*, vol. 38, no. 7, pp. 659-666.

Kicman, A. T. & Cowan, D. A. 1992, "Peptide hormones and sport: misuse and detection", *Br.Med.Bull.*, vol. 48, no. 3, pp. 496-517.

Kicman, A. T., Miell, J. P., Teale, J. D., Powrie, J., Wood, P. J., Laidler, P., Milligan, P. J., & Cowan, D. A. 1997, "Serum IGF-I and IGF binding proteins 2 and 3 as potential markers of doping with human GH", *Clin.Endocrinol.(Oxf)*, vol. 47, no. 1, pp. 43-50.

Kniess, A., Ziegler, E., Kratzsch, J., Thieme, D., & Muller, R. K. 2003, "Potential parameters for the detection of hGH doping", *Anal.Bioanal.Chem.*, vol. 376, no. 5, pp. 696-700.

Kraemer W., Nindl B.C., & Rubin M.R. 2002, "Growth Hormone: Physiological Effects of Exogenous Administration," in *Performance Enhancing Substances in Sport and Exercise*, Bahrke M.S. & Yesalis C.E., eds., Human Kinetics, Champaign IL, pp. 65-78.

Kraemer, W. J., Aguilera, B. A., Terada, M., Newton, R. U., Lynch, J. M., Rosendaal, G., McBride, J. M., Gordon, S. E., & Hakkinen, K. 1995, "Responses of IGF-I to endogenous increases in growth hormone after heavy-resistance exercise", *J.Appl.Physiol*, vol. 79, no. 4, pp. 1310-1315.

Krogsgaard, T. M., Friis, C., Sehested, H. B., Johansen, P., Eschen, C., Nowak, J., & Poulsen, K. 1994, "Studies on the renal kinetics of growth hormone (GH) and on the GH receptor and related effects in animals", *J.Pediatr.Endocrinol.*, vol. 7, no. 2, pp. 93-105.

Kurdy, N. M. 2000, "Serology of abnormal fracture healing: the role of PIIINP, PICP, and BsALP", *J Orthop.Trauma*, vol. 14, no. 1, pp. 48-53.

Kurdy, N. M., Bowles, S., Marsh, D. R., Davies, A., & France, M. 1998, "Serology of collagen types I and III in normal healing of tibial shaft fractures", *J.Orthop.Trauma*, vol. 12, no. 2, pp. 122-126.

Lange, K. H., Isaksson, F., Rasmussen, M. H., Juul, A., Bulow, J., & Kjaer, M. 2001, "GH administration and discontinuation in healthy elderly men: effects on body composition, GH-related serum markers, resting heart rate and resting oxygen uptake", *Clin.Endocrinol.(Oxf)*, vol. 55, no. 1, pp. 77-86.

Lassarre, C., Girard, F., Durand, J., & Raynaud, J. 1974, "Kinetics of human growth hormone during submaximal exercise", *J.Appl.Physiol*, vol. 37, no. 6, pp. 826-830.

Le, R. D., Bondy, C., Yakar, S., Liu, J. L., & Butler, A. 2001, "The somatomedin hypothesis: 2001", *Endocr.Rev.*, vol. 22, no. 1, pp. 53-74.

Le, R. D. & Sonksen, P. H. 1998, "Research on growth hormone (GH) and the insulin-like growth factors (IGFs)", *Growth Horm.IGF.Res.*, vol. 8, no. 1, pp. 1-2.

Leger, J., Reverchon, C., Porquet, D., Noel, M., & Czernichow, P. 1995, "The wide variation in urinary excretion of human growth hormone in normal growing and growth hormone-deficient children limits its clinical usefulness", *Horm.Res.*, vol. 44, no. 2, pp. 57-63.

Longobardi, S., Keay, N., Ehrnborg, C., Cittadini, A., Rosen, T., Dall, R., Boroujerdi, M. A., Bassett, E. E., Healy, M. L., Pentecost, C., Wallace, J. D., Powrie, J., Jorgensen, J. O., & Sacca, L. 2000, "Growth hormone (GH) effects on bone and collagen turnover in healthy adults and its potential as a marker of GH abuse in sports: a double blind, placebo-controlled study. The GH-2000 Study Group", *J.Clin.Endocrinol.Metab*, vol. 85, no. 4, pp. 1505-1512.

MacGillivray, M. H., Frohman, L. A., & Doe, J. 1970, "Metabolic clearance and production rates of human growth hormone in subjects with normal and abnormal growth", *J.Clin.Endocrinol.Metab*, vol. 30, no. 5, pp. 632-638.

Malcovati, L., Pascutto, C., & Cazzola, M. 2003, "Hematologic passport for athletes competing in endurance sports: a feasibility study", *Haematologica*, vol. 88, no. 5, pp. 570-581.

Massart, C. & Poirier, J. Y. 2006, "Serum insulin-like growth factor-I measurement in the follow-up of treated acromegaly: comparison of four immunoassays", *Clin Chim.Acta*, vol. 373, no. 1-2, pp. 176-179.

Maughan, R. 2004, "Contamination of supplements: an interview with professor Ron Maughan by Louise M. Burke", *Int.J Sport Nutr.Exerc.Metab*, vol. 14, no. 4, p. 493.

Maughan, R. J. 2005, "Contamination of dietary supplements and positive drug tests in sport", *J Sports Sci.*, vol. 23, no. 9, pp. 883-889.

McGreevy, K., Hoel, B., Lipsitz, S., Bissada, N., & Hoel, D. 2005, "Racial and anthropometric differences in plasma levels of insulin-like growth factor I and insulin-like growth factor binding protein-3", *Urology*, vol. 66, no. 3, pp. 587-592.

McHugh, C. M., Park, R. T., Sonksen, P. H., & Holt, R. I. 2005, "Challenges in detecting the abuse of growth hormone in sport", *Clin.Chem.*, vol. 51, no. 9, pp. 1587-1593.

Mesotten, D. & Van den, B. G. 2006, "Changes within the GH/IGF-I/IGFBP axis in critical illness", *Crit Care Clin*, vol. 22, no. 1, pp. 17-28, v.

Moor, J. W. & Khan, M. I. 2005, "Growth hormone abuse and bodybuilding as aetiological factors in the development of bilateral internal laryngocoeles. A case report", *Eur.Arch.Otorhinolaryngol.*, vol. 262, no. 7, pp. 570-572.

Muller, E. E., Locatelli, V., & Cocchi, D. 1999, "Neuroendocrine control of growth hormone secretion", *Physiol Rev.*, vol. 79, no. 2, pp. 511-607.

Nabarro, J. D. 1987, "Acromegaly", *Clin.Endocrinol.(Oxf)*, vol. 26, no. 4, pp. 481-512.

Narayan, S. K. & Dutta, J. K. 2005, "Creutzfeldt-Jakob disease", *J Assoc.Physicians India*, vol. 53, pp. 791-795.

National Center on Addiction and Substance Abuse 2000, *Winning at Any Cost: Doping in Olympic Sport*, National Center on Addiction and Substance Abuse, New York.

Nelson AE, Meinhardt U, Hansen JL, Walker IH, Stone G, Howe CJ, Leung K-C1, Baxter RC, Handelsman DJ, Kazlauskas R, & Ho, K. K. 2008, "Pharmacodynamics of GH abuse biomarkers and the influence of gender andtestosterone: a randomized double-blind placebo-controlled study in young recreational athletes.", *Journal of Clinical Endocrinology Metabolism (in print).*

Nelson, A. E., Howe, C. J., Nguyen, T. V., Leung, K. C., Trout, G. J., Seibel, M. J., Baxter, R. C., Handelsman, D. J., Kazlauskas, R., & Ho, K. K. 2006, "Influence of demographic factors and sport type on growth hormone-responsive markers in elite athletes", *J Clin.Endocrinol.Metab*, vol. 91, no. 11, pp. 4424-4432.

Nike. Nike official website. Nike official website . 26-7-2005. 20-3-2006.
Ref Type: Electronic Citation

Osler, T. & Dodd, E. 1979, *Ultra-Marathoning: The Next Challenge* World, Mountain View, CA.

Parker, M. L., Utiger, R. D., & Ughaday, W. H. 1962, "Studies on human growth hormone. II. The physiological disposition and metabolic fate of human growth hormone in man", *J.Clin.Invest*, vol. 41, pp. 262-268.

Powrie, J., Weissberger, A., & Sonksen, P. 1995, "Growth hormone replacement therapy for growth hormone-deficient adults", *Drugs*, vol. 49, no. 5, pp. 656-663.

Powrie, J. K., Bassett, E. E., Rosen, T., Jorgensen, J. O., Napoli, R., Sacca, L., Christiansen, J. S., Bengtsson, B. A., & Sonksen, P. H. 2007, "Detection of growth hormone abuse in sport", *Growth Horm.IGF Res.*, vol. 17, no. 3, pp. 220-226.

Prokop, L. 1970, "The struggle against doping and its history", *J.Sports Med.Phys.Fitness*, vol. 10, no. 1, pp. 45-48.

Quarmby, V., Quan, C., Ling, V., Compton, P., & Canova-Davis, E. 1998, "How Much Insulin-Like Growth Factor I (IGF-I) Circulates? Impact of Standardization on IGF-I Assay Accuracy", *Journal of Clinical Endocrinology Metabolism*, vol. 83, no. 4, pp. 1211-1216.

Raben, M. S. 1958, "Treatment of a pituitary dwarf with human growth hormone", *J.Clin.Endocrinol.Metab*, vol. 18, no. 8, pp. 901-903.

Ranke, M. B., Osterziel, K. J., Schweizer, R., Schuett, B., Weber, K., Robbel, P., Vornwald, A., Blumenstock, G., & Elmlinger, M. W. 2003, "Reference levels of insulin-like growth factor I in the serum of healthy adults: comparison of four immunoassays", *Clin Chem.Lab Med*, vol. 41, no. 10, pp. 1329-1334.

Refetoff, S. & Sonksen, P. H. 1970, "Disappearance rate of endogenous and exogenous human growth hormone in man", *J.Clin.Endocrinol.Metab*, vol. 30, no. 3, pp. 386-392.

Rennie, M. J. 2003, "Claims for the anabolic effects of growth hormone: a case of the emperor's new clothes?", *Br.J.Sports Med.*, vol. 37, no. 2, pp. 100-105.

Rigamonti, A. E., Cella, S. G., Marazzi, N., Di, L. L., Sartorio, A., & Muller, E. E. 2005, "Growth hormone abuse: methods of detection", *Trends Endocrinol.Metab*, vol. 16, no. 4, pp. 160-166.

Rinaldi, S., Kaaks, R., Zeleniuch-Jacquotte, A., Arslan, A. A., Shore, R. E., Koenig, K. L., Dossus, L., Riboli, E., Stattin, P., Lukanova, A., & Toniolo, P. 2005, "Insulin-like growth factor-I, IGF binding protein-3, and breast cancer in young women: a comparison of risk estimates using different peptide assays", *Cancer Epidemiol.Biomarkers Prev.*, vol. 14, no. 1, pp. 48-52.

Rogol, A. D. 2003, "Growth, body composition and hormonal axes in children and adolescents", *J Endocrinol Invest*, vol. 26, no. 9, pp. 855-860.

Sartorio, A., Agosti, F., Marazzi, N., Maffiuletti, N. A., Cella, S. G., Rigamonti, A. E., Guidetti, L., Di, L. L., & Muller, E. E. 2004a, "Combined evaluation of resting IGF-I, N-terminal propeptide of type III procollagen (PIIINP) and C-terminal cross-linked telopeptide of type I collagen (ICTP) levels might be useful for detecting inappropriate GH administration in athletes: a preliminary report", *Clin.Endocrinol.(Oxf)*, vol. 61, no. 4, pp. 487-493.

Sartorio, A., Jubeau, M., Agosti, F., Marazzi, N., Rigamonti, A., Muller, E. E., & Maffiuletti, N. A. 2006, "A follow-up of GH-dependent biomarkers during a 6-month period of the sporting season of male and female athletes", *J.Endocrinol.Invest*, vol. 29, no. 3, pp. 237-243.

Sartorio, A., Marazzi, N., Agosti, F., Faglia, G., Corradini, C., De, P. E., Cella, S., Rigamonti, A., & Muller, E. E. 2004b, "Elite volunteer athletes of different sport disciplines may have elevated baseline GH levels divorced from unaltered levels of both IGF-I and GH-dependent bone and collagen markers: a study on-the-field", *J.Endocrinol.Invest*, vol. 27, no. 5, pp. 410-415.

Saugy, M., Cardis, C., Schweizer, C., Veuthey, J. L., & Rivier, L. 1996, "Detection of human growth hormone doping in urine: out of competition tests are necessary", *J.Chromatogr.B Biomed.Appl.*, vol. 687, no. 1, pp. 201-211.

Saugy, M., Robinson, N., Saudan, C., Baume, N., Avois, L., & Mangin, P. 2006, "Human growth hormone doping in sport", *British Journal of Sports Medicine*, vol. 40, no. suppl_1, p. i35-i39.

Sonksen, P. H. 2001, "Insulin, growth hormone and sport", *J.Endocrinol.*, vol. 170, no. 1, pp. 13-25.

Stoffel, K., Engler, H., Kuster, M., & Riesen, W. 2007, "Changes in biochemical markers after lower limb fractures", *Clin.Chem.*, vol. 53, no. 1, pp. 131-134.

Striegel, H., Vollkommer, G., Horstmann, T., & Niess, A. M. 2005, "Contaminated nutritional supplements--legal protection for elite athletes who tested positive: a case report from Germany", *J Sports Sci.*, vol. 23, no. 7, pp. 723-726.

Strike, P. W. 1991, *Statistical methods in laboratory medicine* Butterworth-Heinemann.

Sweeney, H. L. 2004, "Gene doping", *Sci.Am.*, vol. 291, no. 1, pp. 62-69.

Tanner, J. M., Hughes, P. C., & Whitehouse, R. H. 1977, "Comparative rapidity of response of height, limb muscle and limb fat to treatment with human growth hormone in patients with and without growth hormone deficiency", *Acta Endocrinol (Copenh)*, vol. 84, no. 4, pp. 681-696.

The World Anti-Doping Agency. The World Anti-Doping Code - The 2006 Prohibited List International Standard. http://www.wada-ama.org/rtecontent/document/2006_LIST.pdf . 2006.
Ref Type: Electronic Citation

U.S.Food and Drug Administration - Center for Food Safety and Applied Nutrition. Overview of Dietary Supplements. http://www.cfsan.fda.gov/~dms/ds-oview.html#what . 4-1-2002. 14-7-2006.
Ref Type: Electronic Citation

Unal, M. & Ozer, U. D. 2004, "Gene doping in sports", *Sports Med.*, vol. 34, no. 6, pp. 357-362.

VanHelder, W. P., Goode, R. C., & Radomski, M. W. 1984, "Effect of anaerobic and aerobic exercise of equal duration and work expenditure on plasma growth hormone levels", *Eur.J.Appl.Physiol Occup.Physiol*, vol. 52, no. 3, pp. 255-257.

Veitch, S. W., Findlay, S. C., Hamer, A. J., Blumsohn, A., Eastell, R., & Ingle, B. M. 2006, "Changes in bone mass and bone turnover following tibial shaft fracture", *Osteoporos.Int.*, vol. 17, no. 3, pp. 364-372.

Veldhuis, J. D. 1998, "Neuroendocrine control of pulsatile growth hormone release in the human: relationship with gender", *Growth Horm.IGF.Res.*, vol. 8 Suppl B, pp. 49-59.

Veldhuis, J. D., Bidlingmaier, M., Anderson, S. M., Evans, W. S., Wu, Z., & Strasburger, C. J. 2002, "Impact of experimental blockade of peripheral growth hormone (GH) receptors on the kinetics of endogenous and exogenous GH removal in healthy women and men", *J.Clin.Endocrinol.Metab*, vol. 87, no. 12, pp. 5737-5745.

Vogel, G. 2004, "A race to the starting line", *Science*, vol. 305, no. 5684, pp. 632-635.

Voy, R. O. 1991, *Drugs, sport, and politics Deeter* Leisure Press, Champaign, Ill.

Wallace, J. D., Cuneo, R. C., Baxter, R., Orskov, H., Keay, N., Pentecost, C., Dall, R., Rosen, T., Jorgensen, J. O., Cittadini, A., Longobardi, S., Sacca, L., Christiansen, J. S., Bengtsson, B. A., & Sonksen, P. H. 1999, "Responses of the growth hormone (GH) and insulin-like growth factor axis to exercise, GH administration, and GH withdrawal in trained adult males: a potential test for GH abuse in sport", *J.Clin.Endocrinol.Metab*, vol. 84, no. 10, pp. 3591-3601.

Wallace, J. D., Cuneo, R. C., Bidlingmaier, M., Lundberg, P. A., Carlsson, L., Boguszewski, C. L., Hay, J., Boroujerdi, M., Cittadini, A., Dall, R., Rosen, T., & Strasburger, C. J. 2001, "Changes in non-22-kilodalton (kDa) isoforms of growth hormone (GH) after administration of 22-kDa recombinant human GH in trained adult males", *J.Clin.Endocrinol.Metab*, vol. 86, no. 4, pp. 1731-1737.

Wallace, J. D., Cuneo, R. C., Lundberg, P. A., Rosen, T., Jorgensen, J. O., Longobardi, S., Keay, N., Sacca, L., Christiansen, J. S., Bengtsson, B. A., & Sonksen, P. H. 2000, "Responses of markers of bone and collagen turnover to exercise, growth hormone (GH) administration, and GH withdrawal in trained adult males", *J.Clin.Endocrinol.Metab*, vol. 85, no. 1, pp. 124-133.

Welle, S. 1998, "Growth hormone and insulin-like growth factor-I as anabolic agents", *Curr.Opin.Clin.Nutr.Metab Care*, vol. 1, no. 3, pp. 257-262.

Welle, S., Thornton, C., Statt, M., & McHenry, B. 1996, "Growth hormone increases muscle mass and strength but does not rejuvenate myofibrillar protein synthesis in healthy subjects over 60 years old", *J.Clin.Endocrinol.Metab*, vol. 81, no. 9, pp. 3239-3243.

Weltman, A., Weltman, J. Y., Hartman, M. L., Abbott, R. D., Rogol, A. D., Evans, W. S., & Veldhuis, J. D. 1994, "Relationship between age, percentage body fat, fitness,

and 24-hour growth hormone release in healthy young adults: effects of gender", *J Clin Endocrinol Metab*, vol. 78, no. 3, pp. 543-548.

Williams L. & Fainaru-Wada M. 2006, *Game of Shadows: Barry Bonds, Balco, and the Steroids Scandal That Rocked Professional Sports* Gotham Books.

Woitge, H. W., Scheidt-Nave, C., Kissling, C., Leidig-Bruckner, G., Meyer, K., Grauer, A., Scharla, S. H., Ziegler, R., & Seibel, M. J. 1998, "Seasonal variation of biochemical indexes of bone turnover: results of a population-based study", *J Clin Endocrinol Metab*, vol. 83, no. 1, pp. 68-75.

Woodhouse, L. J., Mukherjee, A., Shalet, S. M., & Ezzat, S. 2006, "The influence of growth hormone status on physical impairments, functional limitations, and health-related quality of life in adults", *Endocr.Rev.*, vol. 27, no. 3, pp. 287-317.

World Anti-Doping Agency 2003, *International Standard for Testing - Version 3.0* Montréal, Canada.

World Anti-Doping Agency 2004a, *Addendum to the International Standard for Laboratories.*

World Anti-Doping Agency 2004b, *International Standard for Laboratories.*

Wu, Z., Bidlingmaier, M., Dall, R., & Strasburger, C. J. 1999, "Detection of doping with human growth hormone", *Lancet*, vol. 353, no. 9156, p. 895.

Yarasheski, K. E. 1994, "Growth hormone effects on metabolism, body composition, muscle mass, and strength", *Exerc.Sport Sci.Rev.*, vol. 22, pp. 285-312.

Yarasheski, K. E., Campbell, J. A., Smith, K., Rennie, M. J., Holloszy, J. O., & Bier, D. M. 1992, "Effect of growth hormone and resistance exercise on muscle growth in young men", *Am.J.Physiol*, vol. 262, no. 3 Pt 1, p. E261-E267.

Yarasheski, K. E., Zachweija, J. J., Angelopoulos, T. J., & Bier, D. M. 1993, "Short-term growth hormone treatment does not increase muscle protein synthesis in experienced weight lifters", *J.Appl.Physiol*, vol. 74, no. 6, pp. 3073-3076.

Yarasheski, K. E., Zachwieja, J. J., Campbell, J. A., & Bier, D. M. 1995, "Effect of growth hormone and resistance exercise on muscle growth and strength in older men", *Am.J.Physiol*, vol. 268, no. 2 Pt 1, p. E268-E276.

Yesalis C.E. & Bahrke M.S. 2002, "History of Doping in Sport," in *Performance-Enhancing Substances in Sport and Exercise*, 1 edn, Yesalis C.E. & Bahrke M.S., eds., Human Kinetics, Champaign IL, pp. 1-20.

Zachwieja, J. J. & Yarasheski, K. E. 1999, "Does growth hormone therapy in conjunction with resistance exercise increase muscle force production and muscle mass in men and women aged 60 years or older?", *Phys.Ther.*, vol. 79, no. 1, pp. 76-82.

Zhang, D. J., Elswick, R. K., Miller, W. G., & Bailey, J. L. 1998, "Effect of serum-clot contact time on clinical chemistry laboratory results", *Clin Chem.*, vol. 44, no. 6 Pt 1, pp. 1325-1333.

Zorpette, G. 2000, "All doped up--and going for the gold", *Sci.Am.*, vol. 282, no. 5, pp. 20, 22.

PUBLISHED WORK FROM THIS THESIS

The majority of the work presented in this thesis has recently been published in peer reviewed academic journals. Copies of these articles have been included in this thesis.

Results from Chapter 3, have been published at:

Erotokritou-Mulligan, I., Bassett, E. E., Bartlett, C., Cowan, D., McHugh, C., Sonksen, P. H., & Holt, R. I. 2008, "The influence of ethnicity on insulin like growth factor-I and procollagen III peptide in elite athletes and its effect on the ability to detect GH abuse", Clin Endocrinol.

Results from Chapter 4, have been published at:

Erotokritou-Mulligan, I., Bassett, E. E., Bartlett, C., Cowan, D., McHugh, C., Seah, R., Curtis, B., Wells, V., Harrison, K., Sonksen, P. H., & Holt, R. I. 2008, "The effect of sports injury on insulin like growth factor-I and procollagen III peptide: implications for detection of growth hormone abuse in athletes", J Clin Endocrinol Metab.

Results from Chapter 6, have been published at:

Erotokritou-Mulligan, I., Bassett, E. E., Kniess, A., Sonksen, P. H., & Holt, R. I. 2007, "Validation of the growth hormone (GH)-dependent marker method of detecting GH abuse in sport through the use of independent data sets", Growth Horm.IGF Res., vol. 17, no. 5, pp. 416-423.

Results from Chapter 5 are currently submitted pending approval at the Journal of Clinical Endocrinology & Metabolism.

In addition, the research undertaken by the GH-2004 project has also been discussed at:

Holt, R. I., Erotokritou-Mulligan, I., McHugh, C. M., Cowan, D. A., Bartlett C, Bassett, E. E., & Sonksen, P. H. 2006, "The GH-2004 project: the use of growth hormone (GH)-dependent markers in the detection of GH abuse in sport. In: Schänzer,

W., Geyer, H., Gotzmann, A., Mareck, U. (eds.) Recent Advances in Doping Analysis", Schänzer W Geyer H Gotzmann A Mareck, ed.