Richard IG Holt[1], Ioulietta Erotokritou-Mulligan[1], Cathy M McHugh[1], David Cowan[2], Christiaan Bartlett[2], E Eryl Bassett[3], Peter H Sönksen[1]

The GH-2004 project: the use of growth hormone (GH)-dependent markers in the detection of GH abuse in sport

[1]Endocrinology and Metabolism Sub-Division, Developmental Origins of Adult Health and Disease Division, School of Medicine, University of Southampton, UK, [2]Drug Control Centre, Department of Forensic Science and Drug Monitoring, King's College London, UK, [3]Institute of Mathematics, Statistics and Actuarial Science, University of Kent, UK

Introduction

Growth hormone (GH) is a naturally occurring endogenous peptide hormone produced by the pituitary gland. It has strong anabolic properties regulating body composition and is widely accepted as being a major drug of abuse in sport. Its use is banned by the International Olympic Committee (IOC) and it appears on the World Anti-Doping Agency (WADA) list of prohibited substances.

The detection of exogenously administered GH poses a formidable challenge, as it is almost identical to the GH produced naturally by the pituitary gland [1]. Furthermore the pulsatile secretion of GH leads to wide variations in circulating GH concentrations, not least in the post-competition setting where exercise acts as a potent stimulus for GH secretion.

The methods for detecting the abuse of androgenic anabolic steroids and related substances measured by mass spectroscopy are highly sophisticated but no such methods have been developed for testing for abuse with peptide and glycoprotein hormones such as GH. Immuno-assays and blood sampling are required for the detection of these substances and, because they are rapidly degraded in the body, urine analyses are not an option.

Two approaches for the detection of GH doping have been developed. The first approach is based on the measurement of different GH isoforms. When rhGH is administered as a single 22K

isoform, it inhibits endogenous pituitary production of the multiple GH isoforms by negative feedback regulation, resulting in suppressed concentrations of other GH isoforms and their ratio to 22K GH [2]. This method is able to detect administered rhGH within a short 'window of opportunity' of up to about 36 hours after the last injection but has several limitations: it will not detect any injection of pituitary-derived GH (that is readily available) and its sensitivity and specificity have not been fully assessed. Furthermore this test is unable to detect abuse with GH secretagogues.

The second approach is based on the measurement of GH-dependent protein markers, such as insulin like growth factor-I (IGF-I), IGF binding proteins and bone markers, such as pro-collagen type III (P-III-P) [3]. The administration of rhGH leads to a significant rise in these markers, the magnitude and duration of elevation of which is dependent on the dose of rhGH given, gender and the individual marker [4-7].

As these proteins occur physiologically, detection of GH abuse must rely on detecting levels 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 androgen or oestrogen use [3,8-10]. This makes it more difficult to define cut-off levels beyond which GH abuse could be proven [3]. 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, after considering combinations of up to eight IGF binding proteins and bone markers, proposed a test based on the measurement of IGF-I and P-III-P in conjunction with specific equations, "discriminant functions", derived from the observed changes of these markers during a double-blind placebo-controlled rhGH administration study.

GH-2000 was a research project funded jointly by the European Union (EU) and the IOC, whose aim was to determine a method for the detection of GH abuse. It reported its results to the EU and the IOC in January 1999. These results were reviewed by a panel of international experts (including a representative of the Court of Arbitration in Sport – CAS) at a workshop organized by the IOC in Rome in March 1999. This review produced a number of key issues that needed addressing before the experts felt that there would be a viable test suitable for implementation at an Olympic Games. The two main issues raised concerned possible ethnic effects on the GH

dependent markers proposed (since the large majority of volunteers in GH-2000 were white Europeans) and the effects of injury.

The GH-2004 project

The GH-2004 project was established with funding from the United States Anti-Doping Agency and World Anti-Doping Agency to address these issues. The main aim of the GH-2004 project was to investigate any ethnic differences in IGF-I and P-III-P, both in elite athletes and in response to exogenous rhGH in healthy volunteers. In addition the effects of injury on IGF-I and P-III-P were studied in amateur sportsmen and women who had sustained a bony or soft-tissue injury.

Study of Elite Athletes

As the detection of GH abuse relies on detecting levels of IGF-I and P-III-P outside normal physiological variability, it is necessary to construct appropriate normative data against which an individual result may be compared. Where reference ranges exist for these markers, they have been constructed from non-athletic populations. Values in athletes may differ subtly from those in a non-athletic population, reflecting the effects of physical training, acute exercise, genetic background, and the athletic physique and therefore it is important to study elite athletes. The GH-2000 study obtained blood samples from 813 athletes within 2 hours of a major national or international competition [8]. The GH-2004 study obtained a further 304 samples from 9 international sporting events involving 13 disciplines. The ethnic breakdown of the GH-2004 athletes is shown in table 1.

Although the same assays (Nichols IGF-I and CIS P-III-P) were used for both studies, the manufacturers had changed the kits in the interim and therefore adjustments for the changes were needed for both assays.

There was no significant difference in the concentration of IGF-I for male Indo-Asians and Orientals when compared with white Europeans. On the other hand, there was a small but statistically significant difference when white Europeans were compared with Afro-Caribbean athletes ($p=0.0008$), with the sample values for Afro-Caribbean subjects being 8.2% lower on average than white Europeans.

Table 1: Ethnic and gender breakdown of subjects in the GH-2004 study of elite athletes

Ethnic Group	Men	Women	Total
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

There was no significant difference in IGF-I concentration in female Indo-Asian and Afro-Caribbean athletes, when compared with white Europeans. On average IGF-I concentrations were 16.7% higher in the samples of the Oriental volunteers compared to white Europeans although the numbers studied were small ($p=0.02$).

After adjustment for assay run used and the age of the volunteers, there was no significant difference on the levels of P-III-P for male Orientals, Indo-Asians and Others when compared to values collected for Caucasians. On the other hand, P-III-P concentrations were 11% higher in Afro-Carribean men compared with white Europeans. There was no difference in P-III-P concentrations between ethnic groups in women.

Although there were small ethnic variations in both proteins, virtually all individuals lay within the 99% prediction intervals for white European subjects (figure 1).

Double blind rhGH administration Study

If a test based on the measurement of GH-dependent markers is to be applied universally, it is important to assess whether individuals of differing ethnicity respond to rhGH in a similar way, at least in terms of changes in the markers. The GH-2000 project has shown that IGF-I and P-III-P rise in a dose dependent manner following a course of injections of rhGH [6,7]. In order to assess whether there was a similar change in non-white European subjects, the GH-2004

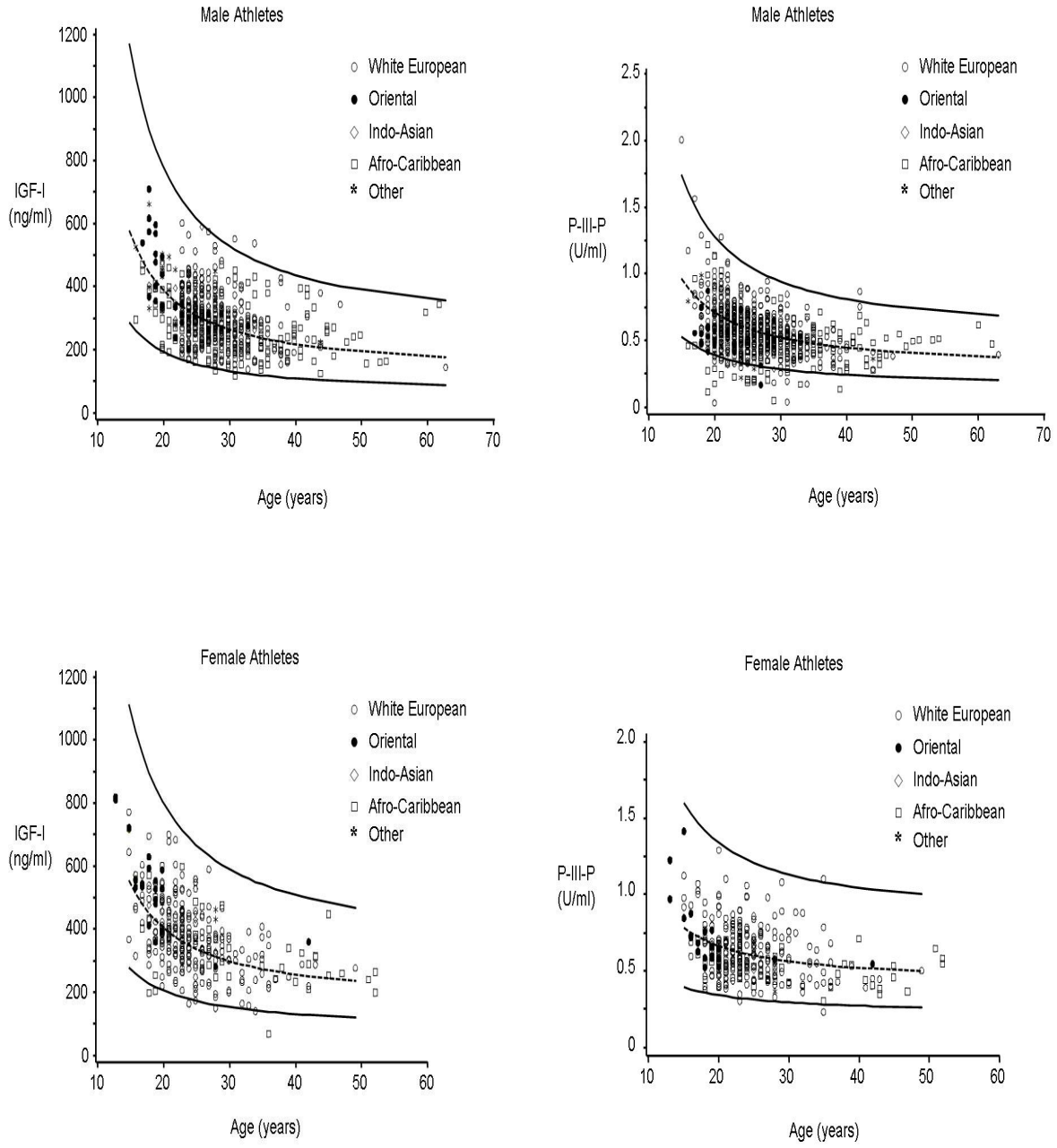

Figure 1: IGF-I and P-III-P in male and female elite athletes in the immediate post-competition setting. The lines represent the 99% prediction intervals for white European athletes.After adjustment for assay run used and the age of the volunteers, there was no significant difference on the levels of P-III-P for male Orientals, Indo-Asians and other ethnic groups when compared to values collected for white Europeans. On the other hand, P-III-P concentrations were 11% higher in Afro-Caribbean men compared with white Europeans. There was no difference in P-III-P concentrations between ethnic groups in women.

project undertook a further double-blind placebo-controlled rhGH administration study in 31 men and 13 women from Afro-Caribbean, Indo-Asian or Oriental backgrounds. rhGH (0.1 or 0.2 units/kg/day) or placebo were self-administered for 28 days by subcutaneous injection. Blood samples were taken before during and up to 56 days after the last injection. The preliminary analysis of the study shows that there are no major ethnic differences in the response to rhGH.

Validation of Discriminant Function

The procedure used in the GH-2000 study to generate the discriminant functions involved splitting the available data into two; a "training" set of data was used to calculate the discriminant function and a "confirmatory" set was used to validate the sensitivity and specificity of the discriminant function and determine whether the function was successful in discriminating between the treatment groups. The confirmatory set was required in order to ensure the model is applicable to the whole population and not just the "training" set. Before a test could be used at an Olympic games, further validation is needed using independent data sets to evaluate whether the GH discriminant function formulae perform reliably in completely different sets of data. Thanks to the work of Astrid Kniess and her colleagues, we have been able to apply the GH-2000 formula to the results of a second independent rhGH administration study undertaken at the Institut für Doping Analytik und Sportbiochemie in Kreischa, in which 0.06 IU/kg body weight /day rhGH or placebo was administered to 15 healthy male, non-competitive athletes for 14 days [11]. When the GH-2000 formula was applied to the data obtained from this study (kindly provided to us by Dr Kniess), there was clear discrimination between those receiving rhGH compared with placebo (figure 2). Furthermore, the GH-2000 formula was able to detect 90% of those receiving rhGH in the Kreischa study correctly, with sensitivity similar to that obtained on the original GH-2000 dataset.

Effect of Injury

There have been fears that skeletal injury may alter the concentrations of both IGF-I and P-III-P. The latter is a particular concern as it is a marker of soft tissue and bone turnover [12,13]. Elevations in either of these proteins could lead to a false accusation of doping with GH.

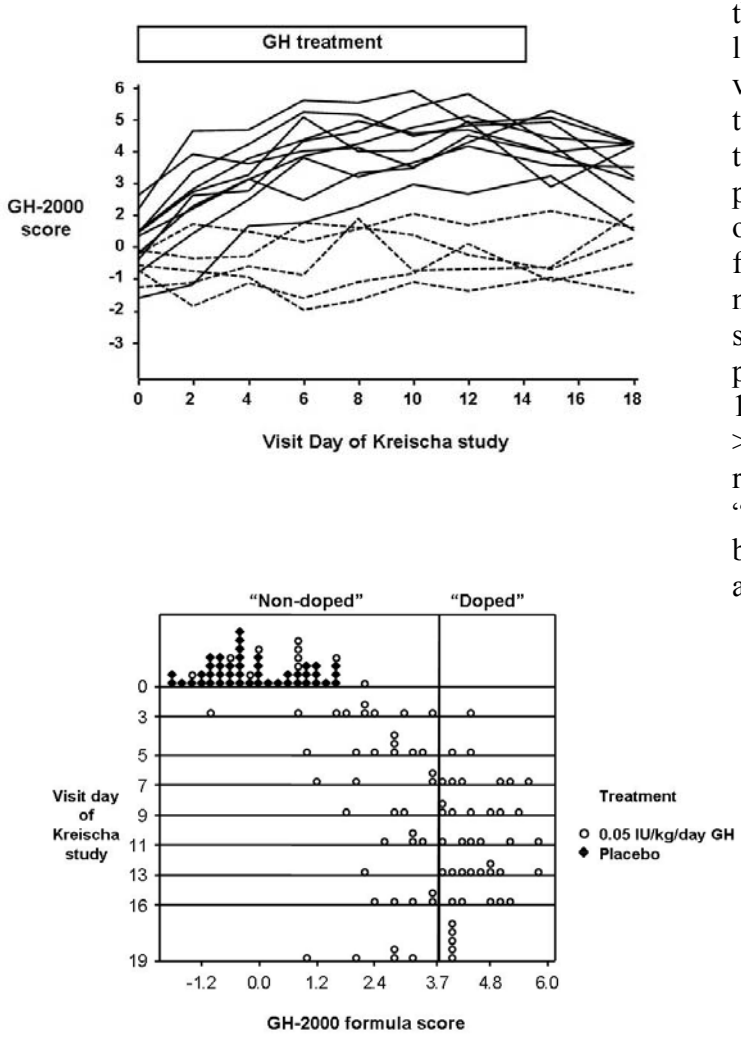

Figure 2: left: the time traces for the GH-2000 score when applied to the Kreischa datasets. The solid line time traces represents volunteers on Growth Hormone treatment. The dotted line time traces represent volunteers on the placebo treatment. Right: Dotplot of the standardised GH-2000 score for each visit day. The mean of this 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 correctly as receiving rhGH (labelled as "doped"). Note that none of the baseline or placebo values are above or even close to 3.7.

The GH-2004 project therefore undertook a study of 137 men (30.0 ± 0.9 yrs) and 34 women (33.3 ± 2.4 yrs) who had sustained either a bony (n=105) or soft tissue (n=66) sporting injury. Blood samples were obtained within a week of the injury and then serial blood tests were taken up to 84 days after the injury. 589 samples were taken, equivalent to 3.4 samples per patient.

There was no significant change in IGF-I in the post-injury period up to 84 days (figure 3). Previous studies have suggested that IGF-I concentrations fall with severe illness and it is possible that the lack of any change reflected the severity of injury. Overall there was no statistically significant change in P-III-P over the 84 days. Some individuals, however, appeared

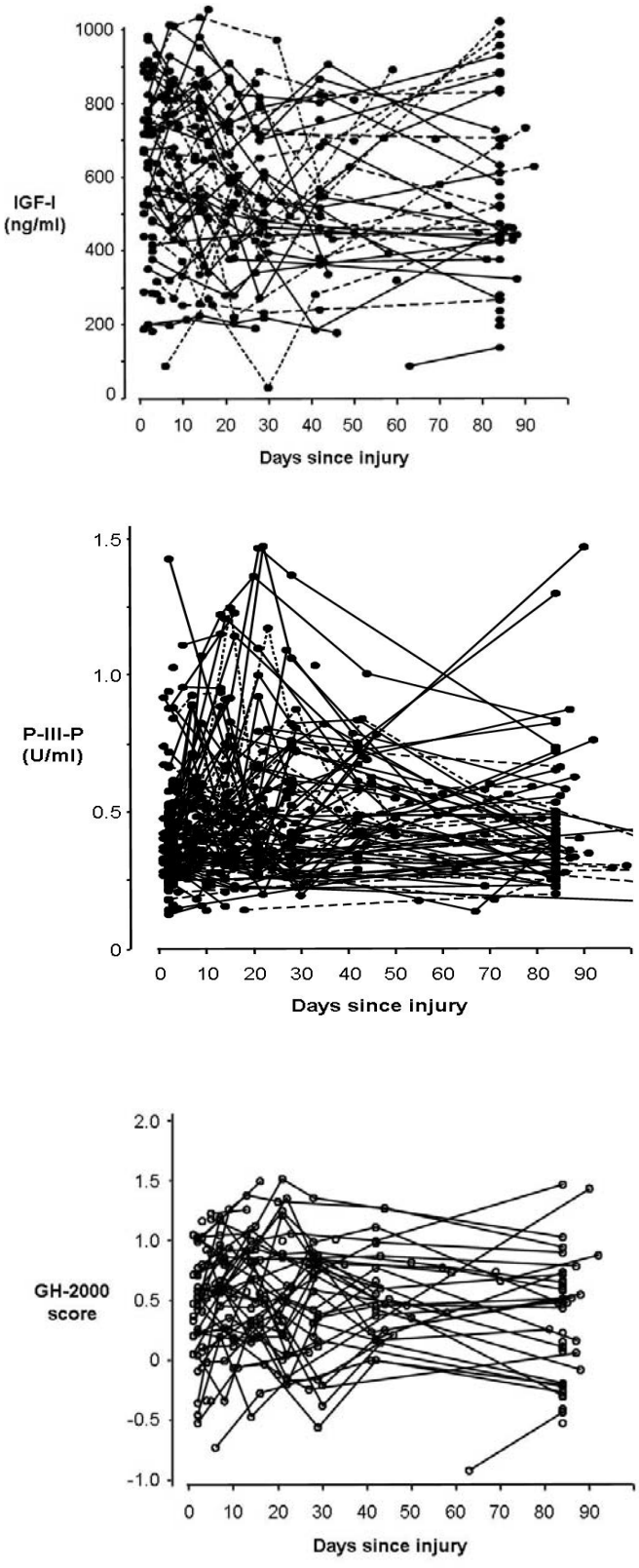

Figure 3: Top left: Time trace of the change in IGF-I following a sporting injury. Top right: Time trace of the change in P-III-P following a sporting injury. Note how the P-III-P rises to a peak after 2-3 weeks in some individuals. Solid line: bony injury. Dotted line: soft tissue injury. Bottom right: Time trace of the change in GH-2000 score following a sporting injury. Despite the rise in P-III-P in some individuals, no subject approached the cut off limit of 3.7 (equivalent to a specificity of 1 in 10,000).

to have a rise in P-III-P after 3-4 weeks. This rise may have occurred in individuals with the most severe injuries and further analysis of this is planned. The GH-2000 formula was applied to data from all subjects in the post-injury period and no subject developed a score greater than 1.6 indicating that none would have been falsely accused because of their injury.

Outstanding issues

There has now been extensive validation of the GH-dependent marker test and the research suggests that the test is feasible and sensitive. A major challenge for the marker approach, however, 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 [14,15]. The establishment of international reference material and quality control schemes has led to harmonisation of assays within the clinical arena and could be applied in anti-doping science. However as this methodology is not currently available, adjustment factors can be and were calculated using the values measured in “undoped” subjects. By assuming that there is no difference in the normal ranges, existing statistical methodology to adjust for the assays can be used in the differing studies. This would prevent these formulae from becoming assay-dependent.

Conclusion

The GH-2004 study has provided further validation that a test based on the measurement of IGF-I and P-III-P can and should be used to detect subjects receiving exogenous GH. Although there are minor ethnic differences in the mean values of IGF-I and P-III-P in elite athletes, most athletes lie within the 99% prediction intervals for white European athletes and the differences are insufficient to invalidate the discriminant function score proposed by the GH-2000 group. IGF-I and P-III-P increase in a dose-dependent manner following the administration of rhGH to subjects from all ethnic groups and therefore the test will have universal applicability. The GH-2000 formula has been shown to work as effectively on an independent data set as it did on its own data, at least in men. Injury will not invalidate the performance of the test.

Acknowledgements

The GH-2004 project is funded by grants from the US Anti-Doping Agency and World Anti-Doping Agency. We would like to acknowledge the support and personal interest of Larry Bowers (USADA) and Olivier Rabin (WADA). The double blind GH administration study was undertaken in the Southampton Wellcome Trust Clinical Research Facility. We are grateful to Astrid Kniess for sharing the data from the Kreischa study with us.

References

[1] McHugh CM, Park RT, Sonksen PH, Holt RI. Challenges in detecting the abuse of growth hormone in sport. Clin Chem 2005; 51(9): 1587-93.

[2] Wu Z, Bidlingmaier M, Dall R, Strasburger CJ. Detection of doping with human growth hormone. Lancet 1999; 353(9156): 895.

[3] Sonksen PH. Insulin, growth hormone and sport. J Endocrinol 2001; 170(1): 13-25.

[4] Wallace JD, Cuneo RC, Baxter R, *et al.* 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 1999; 84(10): 3591-601.

[5] Wallace JD, Cuneo RC, Lundberg PA, *et al.* 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 2000; 85(1): 124-33.

[6] Dall R, Longobardi S, Ehrnborg C, *et al.* 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 2000; 85(11): 4193-200.

[7] Longobardi S, Keay N, Ehrnborg C, *et al.* 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 2000; 85(4): 1505-1512.

[8] Healy ML, Dall R, Gibney J, *et al.* 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 2005; 90(2): 641-649.

[9] Kam GY, Leung KC, Baxter RC, Ho KK. 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 2000; 85(5): 1918-1922.

[10] Juul A. IGF-I measurements. Epidemiological and clinical aspects. Front Horm Res 2005; 33: 45-67.

[11] Kniess A, Ziegler E, Kratzsch J, *et al.* Potential parameters for the detection of hGH doping. Anal Bioanal Chem 2003; 376(5): 696-700.

[12] Haukipuro K, Risteli L, Kairaluoma M, Risteli J. Aminoterminal propeptide of type III pro-collagen in serum during wound healing in human beings. Surgery 1990; 107(4): 381-388.

[13] Bail HJ, Kolbeck S, Lindner T, *et al.* The effect of growth hormone on insulin-like growth factor I and bone metabolism in distraction osteogenesis. Growth Horm IGF Res 2001; 11(5): 314-323.

[14] Hoelzel W, Weykamp C, Jeppsson JO, *et al.* IFCC reference system for measurement of hemoglobin A1c in human blood and the national standardization schemes in the United States, Japan, and Sweden: a method-comparison study. Clin Chem 2004; 50(1): 166-74.

[15] Jeppsson J, Kobold U, Barr J, Finke A, Hoelzel W, *et al.* Approved IFCC reference method for the measurement of HbA1c in human blood. Clin Chem Lab Med 2002; 40(1): 78-89.