**Novel markers to detect recombinant human insulin‑like growth factor‑I (rhIGF‑I)/rhIGF binding protein‑3 (rhIGFBP‑3) misuse in athletes**

**Authors:** Richard I.G. Holt1,2, Nishan Guha1,3, Walailuck Böhning1, Christiaan Bartlett4, David A. Cowan4, Peter H. Sönksen1, Dankmar Böhning3

**Institutions:**

1. Human Development and Health Academic Unit, University of Southampton Faculty of Medicine, Southampton, UK
2. University Hospital Southampton NHS Foundation Trust, Southampton, UK
3. Nuffield Division of Clinical Laboratory Sciences, University of Oxford, UK
4. Department of Pharmacy and Forensic Science, Drug Control Centre, King's College London, UK
5. Southampton Statistical Sciences Research Institute, University of Southampton, UK

**Address for Correspondence:** Professor Richard IG Holt, The Institute of Developmental Sciences (IDS Building), MP887, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD, UK

Tel: +44 23 8120 4665, Fax: +44 23 8120 4221

Email address: R.I.G.Holt@soton.ac.uk

**Disclosure information:** All authors have nothing to declare.

**Research Grant:** This project has been carried out with the support of the World Anti‑Doping Agency (WADA) and the Partnership for Clean Competition (PCC).

**Abstract**

Introduction

Insulin‑like growth factor‑I (IGF‑I) is misused by elite athletes for its metabolic and anabolic effects. We have previously shown that it is possible to detect IGF-I misuse by measuring serum IGF-I and procollagen type III amino‑terminal propeptide (P-III-NP) but a pilot study suggested measuring IGF-II, IGF binding protein‑2 (IGFBP-2) and acid‑labile subunit (ALS) may improve the detection of IGF-I administration. The aim of the study was to assess this in a randomized controlled trial.

Methods

26 female and 30 male recreational athletes were randomised to 28 days’ treatment with placebo or recombinant human (rh)IGF‑I/rh IGF binding protein‑3 (IGFBP‑3) complex (30 mg/day or 60 mg/day), followed by 56 days’ washout. IGF-II, IGFBP‑2 and ALS (women only) were measured using commercial immunoassays.

IGFBP‑2 increased and IGF‑II decreased in response to both low and high dose rhIGF‑I/rhIGFBP‑3 in both women and men while ALS decreased in women in response to high dose rhIGF‑I/rhIGFBP‑3. 2 days after discontinuing treatment, significant differences remained between the three treatment groups in IGFBP‑2 and IGF‑II, but not ALS. Thereafter there were no significant differences between the three treatment groups in any of the markers. Combining IGF-I with IGF-II and/or IGFBP-2 improved the performance of the test to detect rhIGF-I/IGFBP-3 administration.

Conclusion

Combining IGFBP‑2 and/or IGF‑II with IGF-I may improve the detection of rhIGF‑I/rhIGFBP‑3 administration in both women and men.

Abstract Word count: 222

**Introduction**

Insulin-like growth factor-I (IGF­I) has potent anabolic and metabolic effects that make it an attractive choice for athletes who wish to misuse performance‑enhancing drugs (1). Two factors appear to have increased the potential for its misuse. First, as GH detection methods have been developed (2), athletes may have turned to IGF‑I as a mediator of the effects of growth hormone. Second, pharmaceutical preparations of rhIGF‑I alone and rhIGF‑I combined with its major binding protein, rhIGF binding protein‑3 (rhIGF‑I/rhIGFBP‑3), have also been developed for clinical use thus increasing the availability of rhIGF-I (1). IGF­I appears on the WADA List of Prohibited Substances because of its potential to improve performance and harm the athlete if misused (3).

As rhIGF-I is structurally identical to endogenous IGF-I, the main challenge in its detection lies in distinguishing exogenous from endogenous IGF-I. The GH-2000 and GH-2004 teams have developed a method to detect GH misuse based on the measurement of serum IGF-I and procollagen type III amino‑terminal propeptide (P‑III‑NP) in combination with sex-specific discriminant functions (the GH-2000 score) that was introduced for the first time at the London 2012 Olympic and Paralympic Games (4;5). We have previously reported that the administration of rhIGF-I/rhIGFBP-3 to healthy recreational athletes led to a substantial increase in IGF-I but a more modest rise in P-III-NP than after rhGH administration (6). Although rhIGF‑I/rhIGFBP‑3 administration could be detected using the GH‑2000 score method, the measurement of serum IGF‑I alone provided better sensitivity.

During the original GH-2000 studies, other markers of GH action were identified that responded to GH administration, including IGF binding protein‑2 (IGFBP­2), IGFBP-3, acid-labile subunit (ALS) and markers of bone and soft tissue turnover (7;8). This led us to hypothesise that the measurement of these proteins may improve the sensitivity to detect rhIGF-I/rhIGFBP-3 administration. Using the samples obtained from the rhIGF-I/rhIGFBP-3 administration study (6), we undertook a pilot study in which we measured IGF-II, IGF binding protein‑2 (IGFBP­2), IGFBP-3, acid-labile subunit (ALS) as well as markers of bone and soft tissue turnover in samples obtained after 21 days administration of 60 mg/day of rhIGF-I/rhIGFBP-3. In women mean IGF‑II decreased by 53% (*P*=0.0028) while mean IGFBP‑2 increased by 119% (*P*=0.0039) and mean ALS decreased by 40% (*P*=0.0022) (9). In men, mean IGF‑II decreased by 51% (*P*<0.0001), mean IGFBP‑2 increased by 125% (*P*=0.0003) but there was no significant change in ALS. By contrast, there were no significant changes in IGFBP‑3 or bone turnover markers in either sex, which may reflect the separate actions of GH and IGF-I (9).

The aim of the current study was to extend the pilot data by measuring analytes that responded in the pilot study to assess whether the measurement of IGF-II, IGFBP-2 or ALS either alone or in combination with either IGF-I or P-III-NP, could improve the sensitivity and specificity of a test to detect IGF-I misuse in athletes.

**Participants and Methods**

**Participants**

The study design has been described previously (6). In brief, 56 healthy recreational athletes (30 men, 26 women) aged between 18­30 years who engaged in regular physical activity (≥2 sessions per week) were recruited. All participants declared that they had not used performance‑enhancing drugs previously and this was verified by the participants’ primary care physicians. Participants were ineligible if they were competing at elite level. All participants provided written informed consent and the study protocol was approved by the Southampton and South West Hampshire Research Ethics Committee. The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. Baseline demographic data were recorded..

**Study protocol**

The study was a randomised, double‑blind, placebo‑controlled trial involving three treatment arms. Participants were randomly assigned to receive low dose (30 mg/day) rhIGF‑I/rhIGFBP­3 complex, high dose (60 mg/day) rhIGF‑I/rhIGFBP­3 complex or placebo. Insmed Incorporated (Virginia, USA) provided the rhIGF‑I/rhIGFBP­3 complex (Mecasermin rinfabate, iPLEX™ 60 mg/mL) and matching placebo in identical packaging. .

Mecasermin rinfabate is a temperature‑sensitive drug and storage instructions were provided for each participant. Drug vials were stored frozen at ­20˚C until 30 minutes prior to injection when the required dose was allowed to thaw at room temperature. Participants self-administered the drug subcutaneously with their evening meal for 28 consecutive days. The volunteers were asked to complete a treatment diary and return empty drug vials at the end of the treatment period.

**Blood sample collection**

Venous blood samples were collected at baseline, at the end of each week during the treatment period (Days 7, 14 and 21) and during the washout period on Days 28, 30, 33, 42 and 84. These were analysed for IGF-II and IGFBP-2. ALS was also measured in women only because there was no change in ALS in men in response to rhIGF-I/rhIGFBP-3 administration in the pilot study. In contrast in women, there was a 40% decrease in ALS. Baseline and Day 28 samples were collected at 0900, with the participant in the fasting state, while samples on other days were collected when the participant was able to attend. Blood was collected according to WADA guidelines (10). In brief, 15ml of venous blood was collected from the antecubital fossa into 5ml SSTII *Advance***™** Vacutainers (Becton Dickinson, Oxford, UK) and allowed to clot at room temperature for 15 minutes. Samples were centrifuged for 15 minutes at 1300*g*; serum was separated and stored frozen at ‑80**°**C until analysis. All samples were coded and anonymised before analysis.

**Analytical methods**

All markers were measured by commercial immunoassay in the Department of Clinical Biochemistry, Oxford University Hospitals NHS Trust, Oxford, UK according to the manufacturer’s instructions. All samples from one individual were analysed in duplicate within the same assay batch to minimise the effects of inter‑assay variability on results. .

IGF­II was measured using the Mediagnost Human IGF­II ELISA (Mediagnost®, Reutlingen, Germany). In brief, IGF­II was dissociated from IGF binding proteins (IGFBPs) by dilution in an acidic buffer and the addition of excess IGF­I prevented further interference by IGFBPs. Intra‑assay CV was 9.5% and 4.4% at concentrations of 291 and 590 ng/mL respectively. Inter­assay CV was 11.5% and 9.5% at concentrations of 327 and 725 ng/mL respectively.

IGFBP­2 was measured using the Mediagnost IGFBP­2 ELISA (Mediagnost®, Reutlingen, Germany). Intra‑assay CV was 8.7% and 5.0% at concentrations of 221 and 658 ng/mL, respectively. The corresponding inter­assay CVs were 12.6%, and 6.0%.

ALS was measured using the Mediagnost Human ALS ELISA (Mediagnost®, Reutlingen, Germany). Intra‑assay CV was 3.5% and 8.1% at concentrations of 677 and 2230 mU/mL respectively. Inter­assay CV was 7.2% and 11.8% at concentrations of 619 and 2120 mU/mL respectively.

**Statistical analysis**

Statistical analyses were performed using Minitab version 17 software (Minitab Ltd, Coventry, UK). Differences in baseline characteristics between treatment groups were assessed using one‑way analysis of variance where the grouping factor corresponds to the treatment group. As marker concentrations were skewed and their distribution was normalised by log‑transformation, all analyses of marker results were performed on the natural log‑transformed scale.

Logistic regression was used to determine which combinations of markers could be used to predict treatment group (rhIGF‑I/rhIGFBP‑3 or placebo). The results of the previously reported IGF-I and P-III-NP were included in this analysis (6).

In more detail, the parameters were estimated in the logistic regression model



where log is the natural logarithm, is the probability for being a case (under treatment) and is one of the markers of interest with associated parameter . From the estimated score  sensitivity and specificity values for each combination of included markers were calculated for each visit day.

The statistical analyses used to determine the optimal combination of markers to detect rhIGF‑I/rhIGFBP‑3 misuse in this report are different from those that will be required to determine the decision limits for any formula selected to detect rhIGF‑I/rhIGFBP‑3 misuse in elite athletes. These analyses are the first step in developing discriminant functions to detect rhIGF‑I/rhIGFBP‑3 misuse and indicate whether the combination is able to discriminate between those taking rhIGF‑I/rhIGFBP‑3 and those taking placebo. As this is exploratory work, we examined three possible decision limits; an empirical decision limit using the highest value in the placebo group, one based on a false positive rate of 1 in 1,000 calculated using a theoretical construction using a Normal curve and one based on a false positive rate of 1 in 10,000 which is the figure to which the WADA works.

**Results**

**Baseline characteristics**

Table 1 shows the baseline characteristics of the groups. No participant discontinued the study because of adverse effects related to the study medication. Participants in all treatment groups reported local erythema and pain at the site of subcutaneous injections and it is likely that this was a reaction to the solvent used to dissolve the drug and placebo. These symptoms were mild and resolved completely after stopping treatment. Three participants (all in the high dose rhIGF‑I/rhIGFBP­3 group) reported increased appetite during treatment.

**Change in Markers in Women**

In women, there were no significant differences between treatment groups in baseline concentrations of any of the markers. During and immediately after the administration period (Days 7, 14, 21 and 28), there were significant differences between the three treatment groups in IGF‑II and IGFBP‑2 concentrations (figure 1a-b). IGF‑II was lower in the rhIGF‑I/rhIGFBP‑3 administration groups than in the placebo group (*P* < 0.001 on Days 7, 14, 21 and 28) while IGFBP‑2 was higher (Day 7, *P* = 0.013; Day 14, *P* = 0.009; Day 21, *P* = 0.02; Day 28, *P* = 0.01). There were significant differences between treatment groups in ALS on Days 7 (*P* = 0.024) and 21 (*P* = 0.016) but not on Days 14 (*P* = 0.079) and 28 (*P* = 0.156, figure 1c). On Day 30, significant differences remained between the three treatment groups in IGFBP‑2 (*P* = 0.048) and IGF‑II (*P* = 0.049) but not ALS (*P* = 0.360). Thereafter there were no significant differences between the three treatment groups in any of the markers.

In men, there were no significant differences between treatment groups in baseline concentration of any of the markers. During and immediately after the administration period, there were significant differences between the three treatment groups in IGF-II and IGFBP‑2 concentrations (figure 2a-b). IGF-II was lower in the rhIGF‑I/rhIGFBP‑3 administration groups than the placebo group (*P* < 0.001 on Days 7, 14, 21 and 28) while IGFBP‑2 was higher (*P* < 0.001 on Days 7, 14, 21 and 28). There were no significant differences between the three treatment groups in IGFBP-2 (*P* = 0.084 on Day 30) or IGF-II (*P* = 0.196 on Day 30) after discontinuation of treatment (Day 30 onwards).

**Combinations of markers**

Logistic regression was applied to find the optimal combination of markers. The previously reported changes in IGF-I and P-III-NP were also included in this analysis (6). At this stage, only significant markers (IGF‑I, IGFBP‑2 and IGF‑II) were included in the modelling while non‑significant markers (P‑III‑NP and ALS) were discarded.

We first devised four models to discriminate between those athletes administering rhIGF‑I/rhIGFBP‑3 and those administering placebo. “Score 0” included IGF-I alone as we previously showed that this had better discriminatory power than the GH-2000 score (6). “Score 1” was based on the measurement of IGF‑I and IGFBP‑2 while “Score 2” was based on the measurement of IGF‑I and IGF‑II. “Score 3” was based on the measurement of IGF‑I, IGF‑II and IGFBP-2.

**“Score 0”** = α + β0 log (IGF‑I).

**“Score 1”** = α + β0 log (IGF‑I)+ β1 log (IGFBP‑2).

**“Score 2”** = α + β0 log (IGF‑I) + β2 log (IGF‑II).

**“Score 3”** = α + β0 log (IGF‑I) + β1 log (IGFBP‑2) + β2 log (IGF‑II).

For each equation, α is a non-interpretable constant and β(n) is the coefficient for log of analyte and “log” is the natural logarithm (table 2).

As there was no difference in the baseline scores between men and women, the sexes were combined for the further analysis (supplementary figure 1).

All four scores increased rapidly during the first week of administration, remained elevated throughout the administration period and then declined. Figure 3 shows the data for score 3 presented as (a) mean and 95% confidence intervals, (b) dot plots with the decision limit showing how many athletes would have been detected at the relevant time points and, (c) spaghetti plots to show the individual responses. Similar figures for scores 0-2 are available as supplementary figures 2-4. When the spaghetti graphs are plotted separately for men and women (supplementary figure 5), the similarity of response between men and women becomes apparent

Figure 4 shows the distribution of four scores in recreational athletes during treatment with rhIGF‑I/rhIGFBP‑3 compared with placebo in the form of histograms with the underlying best Normal distribution fits. If the score is greater than zero then the model predicts that the sample is from the “treatment” (rhIGF‑I/rhIGFBP‑3) group while if the score is less than zero then the model predicts that the sample is from the “control” (placebo) group. All scores increased in response to rhIGF‑I/rhIGFBP‑3 administration.

Using empirical data, 4.6% of the data points overlapped for score 0 compared with 2.8% for score 1, 3.2% for score 2 and 2.3% for score 3. A data point was deemed to contribute to the overlap if it fell in the range defined by the difference between the largest value in the placebo group and smallest value in the rhIGF‑I/rhIGFBP‑3 administration group. The smaller the overlap is, the better the separability of the two groups.

As the decision limits for any anti-doping test are dependent on a Normal distribution of the scores, this was tested using the Anderson-Darling method. For both “Score 0” and “Score 1”, the scores were not normally distributed in the rhIGF‑I/rhIGFBP‑3 administration group (Supplementary figure 6). By contrast, “Score 2” and “Score 3” were normally distributed although the Anderson-Darling test approached statistical significance for “Score 2”.

**Estimation of Sensitivity**

We employed three methods to estimate the sensitivity of the four scores. First we employed an empirical approach where the decision limit was defined as the highest placebo value. In the second, the decision limit was defined as the mean of the placebo group plus 3.09 SD which equates to a specificity of 1 in 1,000. In the third, the decision limit was defined as the mean of the placebo group plus 3.72 SD which equates to a specificity of 1 in 10,000. Table 3 shows the sensitivity of the scores for detecting rhIGF‑I/rhIGFBP‑3 administration on each visit day of the study. A more detailed version of table 3 is available as supplementary table 1. The estimated 1:1000 and 1:10,000 sensitivities for men and women separately are given in supplementary tables 2-3.

**Reliability of Scores**

Within the placebo group, there were 9 measurements taken over an 84 day period. These values were used to estimate the reliability of the score measurement in order to assess the extent to which a score value is reproducible if measured again. High values of reliability indicate high reproducibility and vice versa. A score with reliability ≥0.6 is considered acceptable while a score with reliability ≥0.7 is considered good. The reliability measurements ranged from 0.686 for Score 0 to 0.763 for Score 3. The reliability of scores 1 and 2 were 0.719 and 0.742 respectively.

**Discussion**

This study has shown that serum IGFBP‑2 increased and IGF‑II decreased in response to rhIGF‑I/rhIGFBP‑3 administration in both male and female recreational athletes and either marker could be used to discriminate between the treatment and placebo groups during the administration period. While ALS decreased in women in the high dose group, it was less effective than the other two markers at discriminating between the treatment and placebo groups. We therefore developed three new provisional discriminant functions (“Scores”) based upon logistic regression for detecting rhIGF‑I/rhIGFBP‑3 misuse using the combination of serum IGF‑I results with IGFBP‑2 and/or IGF‑II results. These were also compared with the use of a score based on IGF-I alone.

**Changes in markers**

Our findings that rhIGF‑I/rhIGFBP‑3 administration caused a significant decrease in IGF­II concentrations have also been demonstrated previously in people with type 1 diabetes (11;12); it is believed that this decrease is caused by the displacement of IGF­II from IGFBPs and subsequent rapid clearance of IGF­II from the circulation. As the rhIGF‑I and rhIGFBP‑3 are non-covalently bound, these peptides dissociate and re-associate within the circulation leading to an increase in free IGF‑I. This should behave similarly to exogenous rhIGF‑I alone used in the earlier studies and therefore displace IGF‑II from IGF binding proteins, increasing IGF-II clearance. In the study by Carroll *et al*, rhIGF‑I administration to adults with type 1 diabetes led to an increase in serum IGFBP­2 (12), and our study has similar results. We have previously reported that GH secretion was diminished following rhIGF‑I/rhIGFBP‑3 administration (6) and this could explain the rise in IGFBP­2 as IGFBP­2 concentration increases in starvation and GH deficiency (13;14) while levels are lowered with GH treatment (15). The GH suppression may also explain the fall, albeit of smaller magnitude, of ALS in women with IGF-I administration (16).

**Score Developments and Analyses**

The combination of IGF‑I and IGFBP‑2 results was used to devise the “Score 1” method for detecting rhIGF‑I/rhIGFBP‑3 misuse while the combination of IGF‑I and IGF‑II results was used to devise “Score 2”. All three markers were used to devise “Score 3” and these scores were compared with a score based on IGF-I alone “Score 0”. Despite the well-known sexual dimorphism of the GH-IGF axis, there was no difference in the mean and distribution of the scores between men and women allowing the sexes to be combined. From the perspective of anti-doping this will facilitate implementation of a test based on these scores.

All four scores rose rapidly in response to rhIGF‑I/rhIGFBP‑3 administration but fell rapidly after treatment discontinuation. This suggests that each of the scores could be used reliably to detect misuse while the athlete is taking rhIGF‑I/rhIGFBP‑3. Furthermore 11-28% of those receiving rhIGF-I/rhIGFBP-3 were still detected 2 days after discontinuation of treatment. Although this represents a shorter window of opportunity to detect the misuse after treatment is stopped than the GH-2000 test for GH misuse, reflecting the much shorter half-lives of IGF-I, IGF-II and IGFBP-2 compared with P-III-NP, this window is longer than the WADA GH isoform test (2).

For anti‑doping purposes, the formulae derived from these analyses will need to be applied to normative data obtained from elite athletes in order to derive the decision limits and any appropriate age correction, as happened for the previous GH‑2000 and GH‑2004 studies. The adaptations will also need to take account of the WADA required specificity of 1 in 10,000.

During this preliminary work, we examined three possible decision limits; an empirical decision limit, one based on a false positive rate of 1 in 1,000 and one based on a false positive rate of 1 in 10,000 which is the figure to which the WADA works. As the empirical decision limits are constructed using the highest value occurred in the placebo group, for these observed data, we have a 100% specificity. However, in practice as we do not have an infinitely large sample, the use of empirical decision limits is not possible for anti-doping purposes. The 1:10,000 decision limits are much higher than the empirical ones and the effect of this is to reduce the sensitivity. The 1:1000 decision limits are still 1.5 SDs higher than the empirical ones but give a 100% specificity for the observed data in this study. By reducing the specificity, we see an improved sensitivity, which is most marked in the 3 days window after treatment cessation.

The estimated sensitivities were similar between the four scores and therefore any score could be used. All four scores performed better than the previously reported sensitivities with the GH-2000 score, which were between 53-61% in women and 75-80% in men while the athletes were receiving the drug (6). However, the development of decision limits with a specificity of 1 in 10,000 relies on a Normal distribution of the scores (5). As the scores involving IGF-I alone (score 0) and IGF-I and IGFBP-2 (score 1) are not normally distributed, the decision limits from these scores would be less reliable. Score 3 involving all three markers appears to be the most robust but the addition of a third marker will add to the cost of the test.

It is not possible to determine the actual sensitivity of an anti‑doping test from this study because IGF-I doping regimens are unknown and it is possible that athletes are administering doses that are higher than those used in this study or in combination with other drugs. The doses used in this study were proposed by the drug manufacturer based on the doses used in clinical practice and by using safety data from their own clinical trials. Furthermore, athletes may be misusing rhIGF‑I alone rather than the rhIGF‑I/rhIGFBP‑3 complex administered in this study. When the current study was designed, the rhIGF‑I/rhIGFBP‑3 complex was chosen for investigation because it appeared to offer the anabolic effects of IGF‑I with lower risk of significant hypoglycaemia compared with administering rhIGF‑I alone. We hypothesise that rhIGF‑I misuse will have the same (or possibly even greater) effects on serum IGF‑I, IGFBP‑2 and IGF‑II because rhIGF‑I misuse will also suppress endogenous GH secretion and displace IGF‑II from IGF binding proteins, with subsequent rapid clearance of IGF‑II from the circulation. However, the shorter half-life of IGF-I alone may affect the window of opportunity of the detection method.

Athletes may also misuse rhGH in combination with rhIGF‑I or rhIGF‑I/rhIGFBP‑3. In that scenario it is likely that serum IGF‑II will show similar changes to the current study because IGF‑II is not regulated by GH concentrations. “Score 2” should not therefore be affected. It is possible, however, that co‑administration of rhGH with rhIGF‑I may prevent the significant increase in IGFBP‑2 observed in this study and thus decrease the sensitivity of “Score 1” and “Score 3”. However, given the likely additive effects on IGF-I, it is likely that co-administration of IGF-I and GH would be detected by the GH-2000 method for detecting GH misuse.

This study is only the first step in developing a test that could be used for anti-doping purposes. Our experience with testing for GH has indicated that further validation will be required before the test could be implemented. WADA rules specify that all analytes measured by immunoassay must be measured by two separate assays recognizing different epitopes. Since this study was performed, a group of laboratories has developed a mass spectrometry method to measure IGF-I (17) and it is possible that a similar method could be adapted to measure IGF‑II (18). This would allow the rapid implementation of the test as duplicate testing is not required for mass spectrometry methods. However, a second immunoassay to measure IGFBP-2 is required. It will be important to investigate pre-analytical variability on the performance of the new assays. Finally it will be essential to develop decision limits based on samples from elite athletes in order to develop a test with a pre-specified specificity of 99.99%, equivalent to a false positive rate of 1 in 10,000. The latter work is on-going. Unlike the GH-2000 test which extended the detection window for GH substantially beyond the last injection, the IGF-I test proposed by this work has a relatively short window-of-opportunity. Further work is needed to discover new markers or methods to detect IGF-I with a longer detection window.

In conclusion, we have developed a further method to detect IGF-I misuse using discriminant functions based upon logistic regression including IGF-I and either or both of IGF-II or IGFBP-2. These formulae are likely to improve the sensitivity to detect IGF-I misuse compared with the GH-2000 score that was originally designed to detect GH misuse.

**Acknowledgements**

This project has been carried out with the support of the Partnership for Clean Competition Research Collaborative and the World Anti‑Doping Agency. The content of this manuscript does not necessarily reflect the views or policies of the Research Collaborative. We would like to thank all the volunteers for their participation in the study. We would like to thank Simon Nevitt, Michael Francis, John Woodland and the nurses of the Wellcome Trust Clinical Research Facility at Southampton General Hospital for their assistance with the study. We thank the University of Southampton Faculty of Medicine and sports societies for allowing us to approach potential participants on campus. We thank Insmed Incorporated for providing the rhIGF‑I/rhIGFBP‑3 complex and placebo and for organising the randomisation for the study. We thank the UHS NHS Trust Pharmacy staff for their assistance with dispensing the study medication.

Conflict of Interest: None

Table 1. Baseline characteristics of 56 recreational athletes. BMI = body mass index, IGF‑I = rhIGF‑I/rhIGFBP‑3 administration. \*Significant difference (*P*<0.05) between placebo, low dose and high dose groups.

|  |  |  |
| --- | --- | --- |
| **Variable** | **Women** | **Men** |
|  | **Placebo (n=8)** | **Low dose IGF‑I (n=9)** | **High dose IGF‑I (n=9)** | **Placebo (n=10)** | **Low dose IGF‑I (n=10)** | **High dose IGF‑I (n=10)** |
| **Mean age (SD), years** | 21.9 (2.2) | 21.7 (3.4) | 21.4 (1.7) | 22.0 (2.8) | 21.9 (2.7) | 23.2 (2.7) |
| **Mean height (SD), cm** | 167.5 (7.7) | 165.2 (2.3) | 169.0 (6.6) | 185.0 (5.8) | 179.2 (10.2) | 181.3 (6.2) |
| **Mean weight (SD), kg** | 61.7 (7.0) | 60.2 (4.9) | 60.5 (7.4) | 92.4 (16.2)\* | 76.9 (12.0)\* | 80.7 (12.9)\* |
| **Mean BMI (SD), kg/m2** | 22.0 (1.6) | 22.0 (1.8) | 21.2 (2.4) | 27.0 (4.3) | 23.8 (2.5) | 24.6 (3.9) |

Table 2: Coefficients employed in Scores 0-3 in men and women combined and their standard errors (SE)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Constant | Log(Immunotech IGF-I) | Log(IGFBP-2) | Log(IGF-II) |
|  | Coefficient | SE | Coefficient | SE | Coefficient | SE | Coefficient | SE |
| Score 0 | -32.96 | 4.68 | 5.49 | 0.79 |  |  |  |  |
| Score 1 | -47.15 | 7.92 | 4.77 | 0.75 | 3.25 | 0.95 |  |  |
| Score 2 | 13.23 | 9.94 | 3.83 | 0.83 |  |  | -5.96 | 1.52 |
| Score 3 | -6.6 | 13.4 | 3.92 | 0.85 | 2.49 | 1.15 | -5.08 | 1.65 |

Log is the natural logarithm

Table 3. Estimated Sensitivity the four scores detecting rhIGF‑I/rhIGFBP­3 administration in 26 female and 30 male recreational athletes. rhIGF‑I/rhIGFBP‑3 was administered between Days 0 and 28. “Score 0” incorporates IGF‑I alone while “Score 1” incorporates IGF‑I and IGFBP-2 results. “Score 2” incorporates IGF‑I and IGF‑II results while “Score 3” incorporates IGF‑I, IGF-II and IGFBP‑2 results. The empirical decision limit was defined as the highest placebo value. The 1 in 1,000 specificity decision limit was defined as the mean of the placebo group plus 3.09 SD. The 1 in 10,000 specificity decision limit was defined as the mean of the placebo group plus 3.72 SD.

|  |  |  |  |
| --- | --- | --- | --- |
| **Day** | **Empirical**  | **Specificity Defined as 1 in 1,000** | **Specificity Defined as 1 in 10,000** |
| **Score 0** | **Score 1** | **Score 2** | **Score 3** | **Score 0** | **Score 1** | **Score 2** | **Score 3** | **Score 0** | **Score 1** | **Score 2** | **Score 3** |
| **Decision Limit** | -0.181 | 0.236 | 1.246 | 1.230 | 1.588 | 2.197 | 2.490 | 2.680 | 2.368 | 3.231 | 3.624 | 3.929 |
| **0** | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| **7** | 92% | 92% | 92% | 95% | 92% | 89% | 92% | 89% | 92% | 82% | 89% | 86% |
| **14** | 92% | 95% | 95% | 92% | 92% | 92% | 92% | 89% | 87% | 84% | 84% | 81% |
| **21** | 97% | 100% | 94% | 94% | 97% | 97% | 94% | 94% | 94% | 89% | 80% | 86% |
| **28** | 97% | 97% | 95% | 95% | 89% | 95% | 95% | 92% | 87% | 76% | 84% | 87% |
| **30** | 67% | 61% | 39% | 44% | 36% | 39% | 31% | 36% | 28% | 17% | 11% | 17% |
| **33** | 5% | 9% | 3% | 3% | 6% | 6% | 3% | 0% | 6% | 3% | 0% | 0% |
| **42** | 8% | 5% | 0% | 0% | 3% | 3% | 0% | 0% | 0% | 0% | 0% | 0% |
| **84** | 10% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |

§Strictly speaking, sensitivity cannot be defined for day 0, 30, 33, 42 and 84 since at these days no treatment has been given. However, since we know the labels of the athletes receiving treatment within the treatment period at days 7, 14, 21, and 28 we use this group and estimate the proportion classified positive.

**Figure legends**

Figure 1: Change in IGF-II (figure 1a), IGFBP-2 (figure 1b) and ALS (figure 1c) in recreational female athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. Data are mean and SEM.

Figure 2: Change in IGF-II (figure 2a) and IGFBP-2 (figure 2b) in recreational male athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. Data are mean and SEM.

Figure 3: a) Interval plot of IGF score 3 in 56 male and female recreational athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. The rhIGF‑I/rhIGFBP‑3 is shown in red and placebo group in blue. The data are mean and 95% confidence intervals. Individual standard deviations were used to calculate the intervals; b) Scatterplot of IGF score 3 for the same individuals with the decision limit to indicate how many would have had adverse analytical findings; c) Spaghetti plot of IGF score 3 for the same individuals to show the individual response over time.

Figure 4: Histogram of IGF scores in 56 male and female recreational athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or or placebo including the best Normal distribution fits.

Supplementary Material:

Supplementary Figure 1: Individual value plot of the four IGF-I scores for both 56 male and female recreational athletes.

Supplementary Figure 2: a) Interval plot of IGF score 0 in 56 male and female recreational athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. The rhIGF‑I/rhIGFBP‑3 is shown in red and placebo group in blue. The data are mean and 95% confidence intervals of the mean. Individual standard deviations were used to calculate the intervals; b) Scatterplot of IGF score 0 for the same individuals with the decision limit to indicate how many would have had adverse analytical findings; c) Spaghetti plot of IGF score 0 for the same individuals to show the individual response over time.

Supplementary Figure 3: a) Interval plot of IGF score 1 in 56 male and female recreational athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. The rhIGF‑I/rhIGFBP‑3 is shown in red and placebo group in blue. The data are mean and 95% confidence intervals of the mean. Individual standard deviations were used to calculate the intervals; b) Scatterplot of IGF score 1 for the same individuals with the decision limit to indicate how many would have had adverse analytical findings; c) Spaghetti plot of IGF score 1 for the same individuals to show the individual response over time.

Supplementary Figure 4: a) Interval plot of IGF score 2 in 56 male and female recreational athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. The rhIGF‑I/rhIGFBP‑3 is shown in red and placebo group in blue. The data are mean and 95% confidence intervals of the mean. Individual standard deviations were used to calculate the intervals; b) Scatterplot of IGF score 2 for the same individuals with the decision limit to indicate how many would have had adverse analytical findings; c) Spaghetti plot of IGF score 2 for the same individuals to show the individual response over time.

Supplementary Figure 5: Spaghetti plot of IGF scores plotted separately for 30 male and 26 recreational athletes who self-administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. The rhIGF‑I/rhIGFBP‑3 is shown in red and placebo group in blue. a) Score 0; b) Score 1; c) Score 2 and d) Score 3.

Supplementary Figure 6: Probability plot of IGF scores in 56 male and female recreational athletes who administered rhIGF‑I/rhIGFBP‑3 60 mg/day for 28 days (high dose), 30 mg/day (low dose) or placebo. Treated individuals are shown in blue and placebo treated individuals are shown in red

Supplementary Table 1. Estimated Sensitivity the four scores detecting rhIGF‑I/rhIGFBP­3 administration in 26 female and 30 male recreational athletes. rhIGF‑I/rhIGFBP‑3 was administered between Days 0 and 28. “Score 0” incorporates IGF‑I alone while “Score 1” incorporates IGF‑I and IGFBP-2 results. “Score 2” incorporates IGF‑I and IGF‑II results while “Score 3” incorporates IGF‑I, IGF-II and IGFBP‑2 results. The empirical decision limit was defined as the highest placebo value. The 1 in 1,000 specificity decision limit was defined as the mean of the placebo group plus 3.09 SD. The 1 in 10,000 specificity decision limit was defined as the mean of the placebo group plus 3.72 SD.

|  |  |  |  |
| --- | --- | --- | --- |
| **Day** | **Empirical** | **Specificity Defined as 1 in 1,000** | **Specificity Defined as 1 in 10,000** |
| **Score 0** | **Score 1** | **Score 2** | **Score 3** | **Score 0** | **Score 1** | **Score 2** | **Score 3** | **Score 0** | **Score 1** | **Score 2** | **Score 3** |
| **Decision Limit** | -0.181 | 0.236 | 1.246 | 1.230 | 1.588 | 2.197 | 2.490 | 2.680 | 2.368 | 3.231 | 3.624 | 3.929 |
| **0** | 5% (2/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) |
| **7** | 92% (35/38) | 92% (35/38) | 92% (35/38) | 95% (36/38) | 92% (35/38) | 89% (34/38) | 92% (35/38) | 89% (34/38) | 92% (35/38) | 82% (31/38) | 89% (34/38) | 87% (33/38) |
| **14** | 92% (34/37) | 95% (35/37) | 95% (35/37) | 92% (34/37) | 92% (34/37) | 92% (34/37) | 92% (34/37) | 89% (33/37) | 87% (32/37) | 84% (31/37) | 84% (31/37) | 81% (30/37) |
| **21** | 97% (34/35) | 100% (35/35) | 94% (33/35) | 94% (33/35) | 97% (34/35) | 97% (34/35) | 94% (33/35) | 94% (33/35) | 94% (33/35) | 89% (31/35) | 80% (28/35) | 86% (30/35) |
| **28** | 97% (37/38) | 97% (37/38) | 95% (36/38) | 95% (36/38) | 89% (34/38) | 95% (36/38) | 95% (36/38) | 92% (35/38) | 87% (33/38) | 76% (29/38) | 84% (32/38) | 87% (33/38) |
| **30** | 67% (24/36) | 61% (22/36) | 39% (14/36) | 44% (16/36) | 36% (13/36) | 39% (14/36) | 31% (11/36) | 36% (13/36) | 28% (10/36) | 17% (6/36) | 11% (4/36) | 17% (6/36) |
| **33** | 15% (5/34) | 9% (3/34) | 3% (1/34) | 3% (1/34) | 6% (2/34) | 6% (2/34) | 3% (1/34) | 0% (0/34) | 6% (2/34) | 3% (1/34) | 0% (0/34) | 0% (0/34) |
| **42** | 8% (3/37) | 5% (2/37) | 0% (0/37) | 0% (0/37) | 3% (1/37) | 3% (1/37) | 0% (0/37) | 0% (0/37) | 0% (0/37) | 0% (0/37) | 0% (0/37) | 0% (0/37) |
| **84** | 10% (4/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) | 0% (0/38) |

Supplementary Table 2. Estimated Sensitivity the four scores detecting rhIGF‑I/rhIGFBP­3 administration in 26 female and 30 male recreational athletes by sex. rhIGF‑I/rhIGFBP‑3 was administered between Days 0 and 28. “Score 0” incorporates IGF‑I alone while “Score 1” incorporates IGF‑I and IGFBP-2 results. “Score 2” incorporates IGF‑I and IGF‑II results while “Score 3” incorporates IGF‑I, IGF-II and IGFBP‑2 results. The 1 in 1,000 specificity decision limit was defined as the mean of the placebo group plus 3.09 SD.

|  |  |
| --- | --- |
| **Day** | **Specificity Defined as 1 in 1,000** |
| **women** | **men** |
| **Score 0** | **Score 1** | **Score 2** | **Score 3** | **Score 0** | **Score 1** | **Score 2** | **Score 3** |
| **Decision Limit** | 1.588 | 2.197 | 2.490 | 2.680 | 1.588 | 2.197 | 2.490 | 2.680 |
| **0** | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/20) | 0% (0/20) | 0% (0/20) | 0% (0/20) |
| **7** | 89% (16/18) | 83% (15/18) | 89% (16/18) | 83% (15/18) | 95% (19/20) | 95% (19/20) | 95% (19/20) | 95% (19/20) |
| **14** | 88% (15/17) | 88% (15/17) | 88% (15/17) | 82% (14/17) | 95% (19/20) | 95% (19/20) | 95% (19/20) | 95% (19/20) |
| **21** | 100% (17/17) | 100% (17/17) | 100% (17/17) | 100% (17/17) | 94% (17/18) | 94% (17/18) | 89% (16/18) | 89% (16/18) |
| **28** | 89% (16/18) | 94% (17/18) | 94% (17/18) | 89% (16/18) | 90% (18/20) | 95% (19/20) | 95% (19/20) | 95% (19/20) |
| **30** | 50% (8/16) | 50% (8/16) | 31% (5/16) | 31% (5/16) | 25% (5/20) | 25% (5/20) | 30% (6/20) | 40% (8/20) |
| **33** | 6% (1/16) | 6% (1/16) | 6% (1/16) | 0% (0/16) | 6% (1/18) | 6% (1/18) | 6% (1/18) | 6% (1/18) |
| **42** | 6% (1/18) | 6% (1/18) | 0% (0/18) | 0% (0/18) | 0% (0/19) | 0% (0/19) | 0% (0/19) | 0% (0/19) |
| **84** | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/20) | 0% (0/20) | 0% (0/20) | 0% (0/20) |

Supplementary Table 3. Estimated Sensitivity the four scores detecting rhIGF‑I/rhIGFBP­3 administration in 26 female and 30 male recreational athletes by sex. rhIGF‑I/rhIGFBP‑3 was administered between Days 0 and 28. “Score 0” incorporates IGF‑I alone while “Score 1” incorporates IGF‑I and IGFBP-2 results. “Score 2” incorporates IGF‑I and IGF‑II results while “Score 3” incorporates IGF‑I, IGF-II and IGFBP‑2 results. The 1 in 10,000 specificity decision limit was defined as the mean of the placebo group plus 3.09 SD.

|  |  |
| --- | --- |
| **Day** | **Specificity Defined as 1 in 10,000** |
| **women** | **men** |
| **Score 0** | **Score 1** | **Score 2** | **Score 3** | **Score 0** | **Score 1** | **Score 2** | **Score 3** |
| **Decision Limit** | 2.368 | 3.231 | 3.624 | 3.929 | 2.368 | 3.231 | 3.624 | 3.929 |
| **0** | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/20) | 0% (0/20) | 0% (0/20) | 0% (0/20) |
| **7** | 89% (16/18) | 83% (15/18) | 89% (16/18) | 83% (15/18) | 95% (19/20) | 80% (16/20) | 90% (18/20) | 90% (18/20) |
| **14** | 88% (15/17) | 76% (13/17) | 82% (14/17) | 71% (12/17) | 85% (17/20) | 90% (18/20) | 85% (17/20) | 90% (18/20) |
| **21** | 94% (16/17) | 100% (17/17) | 82% (14/17) | 88% (15/17) | 94% (17/18) | 78% (14/18) | 78% (14/18) | 83% (15/18) |
| **28** | 89% (16/18) | 72% (13/18) | 89% (16/18) | 83% (15/18) | 85% (17/20) | 85% (179/20) | 80% (16/20) | 90% (18/20) |
| **30** | 38% (6/16) | 19% (3/16) | 6% (1/16) | 13% (2/16) | 20% (4/20) | 15% (3/20) | 15% (3/20) | 20% (4/20) |
| **33** | 6% (1/16) | 6% (1/16) | 0% (0/16) | 0% (0/16) | 6% (1/18) | 0% (0/18) | 0% (0/18) | 0% (0/18) |
| **42** | 6% (1/18) | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/19) | 0% (0/19) | 0% (0/19) | 0% (0/19) |
| **84** | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/18) | 0% (0/20) | 0% (0/20) | 0% (0/20) | 0% (0/20) |

Reference List

 (1) Guha N, Cowan DA, Sonksen PH, Holt RI. Insulin-like growth factor-I (IGF-I) misuse in athletes and potential methods for detection. Anal Bioanal Chem 2013 Dec;405(30):9669-83.

 (2) Holt RI. Detecting growth hormone abuse in athletes. Anal Bioanal Chem 2011 Aug;401(2):449-62.

 (3) World Anti-Doping Agency. The World Anti-Doping Code: The 2015 Prohibited List International Standard. https://wada-main-prod s3 amazonaws com/resources/files/wada-2015-prohibited-list-en pdf 2015Available from: URL: https://wada-main-prod.s3.amazonaws.com/resources/files/wada-2015-prohibited-list-en.pdf

 (4) Powrie JK, Bassett EE, Rosen T, Jorgensen JO, Napoli R, Sacca L, et al. Detection of growth hormone abuse in sport. Growth Horm IGF Res 2007 Jun;17(3):220-6.

 (5) Holt RI, Bohning W, Guha N, Bartlett C, Cowan DA, Giraud S, et al. The development of decision limits for the GH-2000 detection methodology using additional insulin-like growth factor-I and amino-terminal pro-peptide of type III collagen assays. Drug Test Anal 2015 Jan 21.

 (6) Guha N, Erotokritou-Mulligan I, Bartlett C, Nevitt SP, Francis M, Bassett EE, et al. Biochemical markers of insulin-like growth factor-I misuse in athletes: the response of serum IGF-I, procollagen type III amino-terminal propeptide, and the GH-2000 score to the administration of rhIGF-I/rhIGF binding protein-3 complex. J Clin Endocrinol Metab 2014 Jun;99(6):2259-68.

 (7) Dall R, Longobardi S, Ehrnborg C, Keay N, Rosen T, Jorgensen JO, 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 Nov;85(11):4193-200.

 (8) Longobardi S, Keay N, Ehrnborg C, Cittadini A, Rosen T, Dall R, 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 Apr;85(4):1505-12.

 (9) Guha N, Erotokritou-Mulligan I, Nevitt SP, Francis M, Bartlett C, Cowan DA, et al. Biochemical markers of recombinant human insulin-like growth factor-I (rhIGF-I)/rhIGF binding protein-3 (rhIGFBP-3) misuse in athletes. Drug Test Anal 2013 Nov;5(11-12):843-9.

 (10) World Antidoping Agency. International Standard Testing and Investigations. https://wada-main-prod.s3.amazonaws.com/resources/files/WADA\_IST\_2012\_EN.pdf . 2015.

 (11) Cheetham TD, Holly JM, Clayton K, Cwyfan-Hughes S, Dunger DB. The effects of repeated daily recombinant human insulin-like growth factor I administration in adolescents with type 1 diabetes. Diabet Med 1995 Oct;12(10):885-92.

 (12) Carroll PV, Umpleby M, Alexander EL, Egel VA, Callison KV, Sonksen PH, et al. Recombinant human insulin-like growth factor-I (rhIGF-I) therapy in adults with type 1 diabetes mellitus: effects on IGFs, IGF-binding proteins, glucose levels and insulin treatment. Clin Endocrinol (Oxf) 1998 Dec;49(6):739-46.

 (13) Underwood LE, Thissen JP, Lemozy S, Ketelslegers JM, Clemmons DR. Hormonal and nutritional regulation of IGF-I and its binding proteins. Horm Res 1994;42(4-5):145-51.

 (14) Smith WJ, Nam TJ, Underwood LE, Busby WH, Celnicker A, Clemmons DR. Use of insulin-like growth factor-binding protein-2 (IGFBP-2), IGFBP-3, and IGF-I for assessing growth hormone status in short children. J Clin Endocrinol Metab 1993 Nov;77(5):1294-9.

 (15) Munzer T, Rosen CJ, Harman SM, Pabst KM, St CC, Sorkin JD, et al. Effects of GH and/or sex steroids on circulating IGF-I and IGFBPs in healthy, aged women and men. Am J Physiol Endocrinol Metab 2006 May;290(5):E1006-E1013.

 (16) Clemmons DR, Sleevi M, Allan G, Sommer A. Effects of combined recombinant insulin-like growth factor (IGF)-I and IGF binding protein-3 in type 2 diabetic patients on glycemic control and distribution of IGF-I and IGF-II among serum binding protein complexes. J Clin Endocrinol Metab 2007 Jul;92(7):2652-8.

 (17) Cox HD, Lopes F, Woldemariam GA, Becker JO, Parkin MC, Thomas A, et al. Interlaboratory agreement of insulin-like growth factor 1 concentrations measured by mass spectrometry. Clin Chem 2014 Mar;60(3):541-8.

 (18) Thomas A, Kohler M, Schanzer W, Delahaut P, Thevis M. Determination of IGF-1 and IGF-2, their degradation products and synthetic analogues in urine by LC-MS/MS. Analyst 2011 Mar 7;136(5):1003-12.