Challenges in Detecting the Abuse of Growth Hormone in Sports

CATHY M. McHUGH,* RODERICK T. PARK, PETER H. SÖNKENSEN, and RICHARD I.G. HOLT

Background: Growth hormone (GH) is reputed to be in widespread use in the sporting arena as a performance-enhancing agent and is on the list of banned substances published by the World Anti-Doping Agency. The detection of GH abuse poses many challenges. Unlike many substances of abuse, such as synthetic anabolic steroids, GH is a naturally occurring substance; therefore, demonstration of exogenous administration must rely on detecting concentrations in excess of an established reference interval. The purpose of this review is to discuss the methodologies being developed to detect GH abuse.

Methods: We undertook a comprehensive search using multiple electronic databases and hand searches of reference lists of articles. The data for this review reflect our academic interests and experience through work on the GH-2000 and GH-2004 projects.

Results: Two approaches have been taken to detect GH abuse. The first is based on assessment of the effect of exogenous GH on pituitary GH isoforms, and the second is based on measurement of markers of GH action. The advantages of each approach and the difficulties encountered with each technique, as well as future concepts in detection are discussed.

Conclusion: Although there are substantial challenges for the detection of GH, methodologies now exist to detect GH abuse with reasonable sensitivity and specificity.

The human competitive nature is not only innate but essential in evolutionary and survival terms. In the sporting arena, this manifests as an all-consuming drive to win.

Doping is defined in the World Anti-Doping Agency (WADA) code as “the administration of or use by an competing athlete of any substance foreign to the body or any physiological substance taken in abnormal quantity, or taken by an abnormal route of entry into the body with the sole purpose of increasing in an artificial and unfair manner their performance in competition.” The use of artificial means to enhance performance dates back to Ancient Greco-Roman times when figs were used to assist performance. The term “doping” comes from the word “dop”, which is a substance made from grape skins used by Zulu warriors to enhance battle prowess.

The first recorded drug-related fatality occurred in

1 Nonstandard abbreviations: GH, growth hormone; WADA, World Anti-Doping Agency; GC, gas chromatography; rhGH, recombinant human growth hormone; IGF-I, insulin-like growth factor-I; IGFBP, insulin-like growth factor–binding protein; FA, immunofunctional assay; GHBP, growth hormone–binding protein; P-III-P, procollagen III terminal peptide; PICP, C-terminal propeptide of collagen type I; and ICTP, C-terminal cross-linked telopeptide of type I collagen.
1886 when Andrew Linton died on the Bordeaux-Paris cycle race, allegedly from an overdose of strychnine, heroin, and a compound known as “trimethyl”. In 1967, the International Olympic Committee established a Medical Commission and formulated an official list of prohibited substances.

The first systematic testing began at the 1972 Olympic Games in Munich with the analysis of more than 2000 urine samples by gas chromatography (GC) with nitrogen-selective detection for stimulants. Systematic urinary screening was introduced in 1983 at the Pan American Games, and blood testing was used for the first time in 1994 in Lillehammer XVII Olympic Winter Games in an attempt to detect blood doping (4).

WADA was established in 1999 and has produced a list of banned substances that have been classified according to their mode of action. An athlete is considered to have violated the regulations if the prohibited substance is discovered in the athlete’s body fluids or if the athlete attempts to use a prohibited substance or method, fails to submit a sample once requested, or fails to make him- or herself available for out-of-competition testing, unless the athlete can demonstrate that the presence of the substance is the result of a physiologic or pathologic condition (5).

The fear of detection and subsequent disgrace and loss of income is a stronger deterrent than the thought of personal harm. In a survey by Sports Illustrated, 195 of 198 athletes said that they would take a performance-enhancing drug if they were guaranteed to win and not be caught; 50% stated that they would still take the substance even if they would die from a side effect of the drug after 5 years of successful competition (6).

**GH Abuse**

GH is a peptide hormone and is on the list of substances issued by WADA as banned for competitive sports (5). Until recently, there has been no test available to detect GH abuse; therefore, the prevalence of this abuse can be surmised only through anecdotal evidence.

The performance-enhancing potential of GH for use in sports was first advocated in the *Underground Steroid Handbook* in 1983 (2), where it was described as “the most expensive, most fashionable and least understood of the new athletic drugs.”

After Ben Johnson was stripped his 100-m gold medal from the Seoul Olympic Games, he admitted to having taken a cocktail of drugs including GH. A Chinese swimmer, Yuan Yuan, was forced to withdraw from the 1999 world championship after 13 vials of human GH were discovered in her suitcase. More recently, during a grand jury testimony, Tim Montgomery (current 100-m world record holder) admitted receiving an 8-week supply of GH and a steroid compound known as “the clear” (7).

Both pituitary-derived and short-acting and discontinuous depot recombinant human GH (rhGH) are widely available on the internet.

**Why Is GH Abused?**

There are no clinical trials in healthy humans that demonstrate that GH has a performance-enhancing effect. Nevertheless, anecdotal evidence suggests that GH is widely abused for its anabolic and lipolytic properties.

The anabolic actions of GH are mostly mediated through insulin-like growth factor-I (IGF-I) and include increases in total body protein turnover and muscle synthesis as seen adults with GH deficiency and endurance-trained athletes (8, 9). GH and testosterone act through separate mechanisms and have synergistic effects on anabolism. This has not been lost on athletes, who use cocktails of anabolic agents to gain the maximal effect. GH also stimulates proliferation of cartilage in the growing epiphysial plate, stimulates linear growth, and increases bone mass, mineral content, and the number of bone-modeling units (10). GH also induces lipolysis in adipose tissue and leads to a reduction in fat mass (11).

The effect of GH is seen most dramatically in adults with GH deficiency in whom treatment with GH leads to increased muscle mass, enhanced use of lipids as a fuel source, improved thermal regulation, increased cardiac output, and improved wound healing and ligamentous strength (11). The replacement of GH for 6 months in adults can lead to an 8.8% increase in muscle mass and 14.4% loss in fat mass (11).

Excess GH secretion also leads to changes in body composition. Patients with active acromegaly have a decreased fat mass and increased lean body mass, both of which are normalized with successful treatment (12). Although most patients with acromegaly do not exhibit athletic prowess, there have been a few cases of athletic achievement in early acromegaly that have provoked debate about the ability of GH to improve performance (13).

Some authors have suggested that the banning of GH is encouraging its use and that future research to develop methodology to detect its abuse should be stopped (14). These arguments are flawed for several reasons. One reason is that athletes look for an individual response, whereas clinical trials look at mean changes. In addition, similar arguments were made about anabolic steroids, and only recently have clinical trials shown what athletes have known for a long time—anabolic steroids do enhance performance (15). Moreover, research by clinical endocrinologists into the effects of GH has fallen at least a decade behind athletes. Athletes are highly trained to know their performance and evaluate small changes in response to changes in training. By contrast, clinical trials are designed to evaluate relatively large changes. The numbers of study participants needed to detect a 1% change in performance would be huge, whereas the margin for winning an Olympic gold medal is usually less than this. Finally, athletes use a cocktail of drugs that are individually tailored to their requirements. In contrast, clinical trials are designed to evaluate only 1 or 2 interventions at a time with all other variables being kept...
equal. Athletes therefore provide a new paradigm for examining the potential benefits of new anabolic agents that clinical trials could not, and it is incumbent on us to take note of the agents currently being used.

**Further Manipulation of the GH-IGF Axis**

New technologies and ways of manipulating the GH-IGF axis are emerging. Gene doping, GH secretagogues, rhGH-I, and rhGH-I recombinant human IGF-I-binding protein 3 (rHGBP3-3) complex are already in circulation. Certain genotypes confer athletic advantages, and transmission of a genetic code with or without the aid of a vector, such as a virus, allows incorporation of the DNA into target tissues where expression of that gene can leads to enhanced local production of an anabolic substance such as IGF-I. This confers target tissue specificity without altering systemic concentrations of the product and is thus not detected by blood or urine testing. Detection requires tissue biopsy, which would not be feasible in the sporting setting.

Proof-of-concept experiments have been undertaken in animals in which injection of a recombinant adenovirus genetically manipulated to induce myocytes overexpression of IGF-I in young mice induced a 15% increase in muscle mass and a 14% increase in muscle strength without inducing a systemic increase in IGF-I (16).

**Challenges of Detecting GH Abuse**

Important considerations in GH measurement for antidoping include the amino acid sequence identity between the main fraction of pituitary-derived GH and recombinant GH, the heterogeneous nature of GH, the presence of GH-binding proteins in plasma, the potential cross-reactivity with homologous polypeptide hormones (i.e., prolactin), the heterogeneous immunoreactivity of (monoclonal) antibodies used for commercial immunoassays, and the short half-life in circulation.

Detection of abuse with GH poses many challenges. Unlike many substances of abuse, such as synthetic anabolic steroids, GH is a naturally occurring substance; thus, demonstration of exogenous administration must rely on detecting concentrations exceeding established reference intervals and the exclusion of a pathologic cause such as acromegaly. Possible solutions include repeat testing after a period of known abstinence and detailed clinical examination and investigation. Detection is hampered by the fact that recombinant and endogenous GH have identical amino acid sequences.

Physiologic challenges include a pulsatile release pattern, a short half-life of ~20 min, and increased concentrations 2 h after exercise (17, 18). Although researchers can perform repeated sampling over a 24-h period to overcome the issue of pulsatility, this is not feasible in the sporting setting (19).

Traditional drug testing in sport has involved urinary sampling, but it is not viable for rhGH detection because neither GH itself nor markers of GH, which are also peptides, are secreted into the urine in sufficient and reliable quantities (20). Consequently, blood sampling is required for the detection of GH abuse. This is minimally invasive and has been accepted for use in competitive events for blood doping and erythropoietin detection.

**Methodologies for Detecting GH**

Two main approaches have been investigated to detect GH abuse; the first is based on the detection of different pituitary GH isoforms, whereas the second relies on measurement of GH-dependent proteins. Both methods use immunologic assays and are therefore subject to the WADA requirement that 2 separate assays detecting different epitopes are needed to verify the presence of the isoform or marker (5).

**GH Isoforms**

The analysis of GH isoforms was originally termed the “direct method”, but it is now more accurately referred to as the “isoform assay method”. Endogenous GH exists in several forms: the 22-kDa isoform is the most abundant, constituting of 75% of the circulating GH, and others forms, collectively termed “non–22-kDa”, include the 20- and 17-kDa isoforms and many fragments of isoforms.

Recombinant GH contains only the 22-kDa isoform, and exogenous rhGH administration leads to a marked decreased in the endogenous pituitary-derived non–22-kDa isoforms by negative feedback mechanisms. Hence, a high ratio of 22- to non–22-kDa has been proposed as a mechanism of detecting exogenous GH usage (Fig. 1) (21).

Age, sex, physiologic stimulus, and pathologic state do not affect the relative proportions of GH isoforms, but it is unclear whether ethnicity could affect the isoform ratio (22). Exercise causes an increase in the 22-kDa isoform; therefore, post race concentrations of this isoform relative to the other isoforms may lower the test’s sensitivity (23).

GH isoforms have short half-lives; therefore, the window of opportunity for detection is short, up to (at the very most) 36 h (21). Spontaneous GH secretion returns to

---

**Fig. 1.** Effect of administration of one dose of rhGH on 22-kDa and 20-kDa isoforms of GH, expressed as a percentage of pretreatment values.

Adapted from Leung et al. (43) and Thorner et al. (22).
baseline values 48 h after the last dose of GH treatment (24).

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.

**GH ISOFORM ASSAYS**

RIA and IRMA will detect circulating whole isoforms, whereas the immunofunctional assay methodology (IFA) has the advantage of detecting only biologically active GH isoforms. The values obtained from the IFA methods are 26.8% lower than those obtained from conventional IRMAs, reflecting the proportion of GH that is biologically active. However, the use of immunoassay techniques in forensic toxicology remains controversial despite widespread acceptance in laboratory scientific practice (4).

IFA is an ELISA method that relies on the two GH-binding sites of the transmembrane GH receptor: one binding GH and the other binding a radiolabeled monoclonal antibody (25). The solution is flooded with GH-binding protein (GHBP) to bind any remaining unoccupied site 1 epitopes. Nonfunctional isoforms do not bind the GHBP or induce dimerization, which is a critical step in bioactivity, and are thus not detected by the assay. The assay is technically easy to perform, has high capacity, and allows radiolabeling and metric or colorimetric quantification; however, the structural format of the molecule may be such that the monoclonal antibody site is unavailable for the radiolabel to attach, leading to underestimation of the true concentration. Sites already bound by GHBP will not be detected, but in vivo these molecules are too large to pass through the capillary membrane and achieve biological activity (26). The reported interassay and intraassay CVs are 10.3% and 7.3%, respectively, with no interference from endogenous GHBP (25).

The presence of fragments of isoforms that may contain binding sites for antibodies can lead to overestimation of the concentrations of whole and functional isoforms present. For example, the 22-kDa exclusion assays demonstrate incomplete exclusion of 22-kDa isoforms at high concentrations (27).

**GH-DEPENDENT MARKERS**

The anabolic actions of GH lead to generation of several proteins, and the serum concentrations or ratios of these proteins can be used as a means of detecting exogenous GH. This method of GH detection is known as the “marker method”. The measurement of multiple markers in conjunction with specific equations, “discriminant functions”, can be used to detect GH abuse with enhanced sensitivity and specificity compared with single-marker analysis. Two groups of potential markers were identified by the GH-2000 research team: one group includes members of the IGF-IGFBP axis, and the other includes markers of bone and collagen turnover and mineralization (17).

IGF-I is an ideal candidate marker because it has little diurnal or day-to-day variation (22), increases 1.3- to 2.3-fold in a uniform dose-dependent fashion after GH administration (28), and has low basalar scatter (29) and minimal changes with exercise. Ninety-five percent of circulating IGF-I is bound to binding proteins (IGFBP-1 to -6), predominantly IGFBP-3, which modulate its actions and bioavailability (22, 30). IGFBP-3 increases with GH administration but has a less uniform and dose-dependent curve than IGF-I (28). Total IGF-I is used because free IGF-I is less responsive to GH (31).

Similarly, several bone and soft tissue markers change in response to GH administration. Procollagen III terminal peptide (P-III-P) is a marker of type 3 collagen formation (mainly soft tissues); exhibits little day-to-day, diurnal, or gender variation in basal concentrations; and increases in a dose-dependent fashion after GH administration (17). As a soft tissue marker, P-III-P may increase after injury and with several pathologic conditions, including fibrotic lung or liver disease or alcoholic liver disease (32). The GH-2004 team, based in Southampton, is currently evaluating the effects of different types of injury on plasma concentrations of P-III-P and other markers of GH action. C-Terminal propeptide of collagen type I (PICP) is involved in the early scaffold of collagen formation and callus formation with bone remodeling and increases in a non–dose-dependent manner after GH administration. C-Terminal cross-linked telopeptide of type I collagen (ICTP) is a marker of bone resorption and exhibits a short phase increase of 9.7% with males exhibiting an increase greater in magnitude (29).

Osteocalcin and bone alkaline phosphatase are both markers of bone mineralization (Table 1). After GH administration, osteocalcin concentrations increase in a non–dose-dependent manner, and males exhibit a greater increase than females (31). There is a small increase 60 min after exercise. Bone-specific alkaline phosphatase is posture dependent and increases slowly after GH administration (33). Leptin is derived from adipocytes in response to GH, but changes in this marker are too variable for practical use (28).

<table>
<thead>
<tr>
<th>Marker</th>
<th>Preexercise disappearance half-life, h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osteocalcin</td>
<td>770</td>
</tr>
<tr>
<td>PICP</td>
<td>433</td>
</tr>
<tr>
<td>P-III-P</td>
<td>693</td>
</tr>
<tr>
<td>ICTP</td>
<td>248</td>
</tr>
<tr>
<td>IGF-I</td>
<td>89.5</td>
</tr>
<tr>
<td>IGFBP-3</td>
<td>179</td>
</tr>
<tr>
<td>ALS</td>
<td>119</td>
</tr>
</tbody>
</table>

*a* Adapted from Wallace et al. (17) and Longobardi et al. (31).

*b* ALS, bone-specific alkaline phosphatase.
DISCRIMINANT FUNCTION ANALYSIS

Single-marker analysis lacks sufficient specificity to detect exogenous GH abuse; therefore, a combination of markers is used in discriminant function analysis to improve sensitivity and specificity. To generate these equations, a training set of data is applied to create the discriminant model, and a confirmatory set of data is then similarly applied to ensure that the model is applicable to the population in general and not just the sample set.

Kicman et al. (34) proposed the use of ratios of IGF-I to IGFBP-2 and IGFBP-3 to IGFBP-2, both of which were augmented at 30 h after GH administration.

The GH-2000 team reported that IGF-I and P-III-P were the simplest combination of the numerous markers studied providing the best sensitivity and specificity during GH therapy (29). Other markers of bone and collagen turnover also provided useful discrimination between placebo and treatment groups (Fig. 2).

Kniess et al. (28) reported significant increases in the concentrations of the products of the terms (IGF-I × P-III-P) and (IGF-I × IGFBP-3) in a GH administration study (Fig. 3).

The differences in time course of the markers can be exploited by mathematical modeling to improve the sensitivity of the test to detect GH abuse. Comparing the time-dependent proportionate decreases in concentrations of the markers can allow determination of the time since doping last occurred.

CREATION OF APPROPRIATE REFERENCE INTERVALS
FOR GH-DEPENDENT MARKERS

It is important to construct appropriate reference intervals for each of the proposed GH markers. Natural differences in responses of the GH-IGF axis among individuals and among differing body habitus may confer an advantage in sport and lead to differences in marker concentrations.

Exercise is a potent stimulus for GH secretion, and both acute and chronic physical activity increase the concentrations of some of the GH markers. This effect is augmented by GH administration, but the pattern of the increase is unchanged. Total but not free IGF-I and IGFBP-1 both increase after acute exercise (33). Acute exercise also induces a metabolic acidosis that stimulates osteoclasts and inhibits osteoblastic activity, whereas in the longer term, resistance training induces a greater increase in bone remodeling than does endurance training (35). Postcompetition concentrations of GH markers in elite athletes are different from the values in standard reference intervals (36).

An appropriate sex- and age-specific reference interval is important. Women have lower GH peaks and higher troughs and are relatively more GH resistant (10, 22). GH secretion increases through childhood and early adulthood and thereafter decreases by ~14% per decade. There is also a decrease in the secretion of IGF-I and IGFBP-3 with aging (37). Athletes exhibit the same decrease in GH-dependent markers as the general population, and any differences in markers between sporting events are largely explained by differences in the ages of the competitors. Therefore, any antidoping tests for GH must take age into account. This is particularly important in adolescent competitors because during that age period, pubertal staging rather than chronologic age determines GH secretion.

The GH-2004 team based in Southampton and an Australian and Japanese Consortium are currently evaluating the possibility of ethnic differences in the GH-dependent markers. Although there are small differences in the mean concentrations of some markers, notably
IGFBP-3, among ethnic groups, preliminary data suggest that these are not large enough to affect the performance of the test proposed by the GH-2000 team.

**GH Marker Assays**

Several key areas 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 (38, 39).

To improve the accuracy of the IGF-I assays, high-affinity antibodies should be used wherever possible because IGFBPs have similar affinities and compete with conventional antibodies. IGFBPs should be dissociated and separated from the IGF-I before assay, e.g., by acidification (40). Acid ethanol extraction followed by cryoprecipitation is a refinement of the size-exclusion method that has a considerably improved IGF-I recovery of 90%–95%.

The authors of a recent report comparing IGF-I assays (Nichols RIA, Medialogist RIA, and R&D ELISA) reported variability attributable to IGF-I dissociation procedure, calibrators used, IGF-I concentrations, and antibody specificity (41). The authors reported that interlaboratory correlation was best for the ELISA for IGF-I and the Orion Diagnostics RIA for P-III-P (41).

Potential alterations in commercial assays by manufacturers could mean that a new reference interval is required for each new assay (42). There thus is a specific need for WADA to develop its own assays over which it would have complete control to ensure standardization.

**Legal Issues**

The Court of Arbitration of Sport has yet to decide on an acceptable false-positive rate, but it is expected to be in the range of 1 in 10 000 tests. Current medical practice accepts as "normal" values within 2 SD from the mean on a calibration curve. By definition, this means that 5% of the population lie outside the "normal range" and would create an unacceptably high false-positive rate if applied to athletes.

**Ethical Issues**

Although the aim of antidoping organizations and committees is to ensure that competition is fair and free of drugs, antidoping testing has its own ethical considerations. Although largely beyond the scope of this review, it is important to ensure that testing for GH does not cause undesirable ethical and medical consequences that may outweigh the advantages for sports. It is important that any test is consistent and fair to all athletes to whom the test is administered. For this reason, WADA has its own code of ethics and ethics committee to oversee the activities of WADA and to ensure that any testing is transparent and accountable to the public.

**Conclusion**

Anecdotal evidence suggests that GH has been widely abused by sportsmen and -women for the last 20 years because of its anabolic and lipolytic properties. Recent advances in methodologies to detect GH have led to the introduction of testing for GH for the first time at the Athens Olympic Games.

The GH-2004 project is funded through grants from the US Anti-Doping Agency and World Anti-Doping Agency.

**References**


