Clinical Chemistry 57:4
555–559 (2011)

Special Reports

Consensus Statement on the Standardization and Evaluation of Growth Hormone and Insulin-like Growth Factor Assays

David R. Clemmons,[1*] on behalf of the conference participants

Growth hormone (GH) and insulin-like growth factor I (IGF-I) measurements are widely used in the diagnosis of disorders of GH secretion, evaluation of children with short stature from multiple causes, management of disorders that lead to nutritional insufficiency or catabolism, and monitoring both GH and IGF-I replacement therapy. Therefore, there is an ongoing need for accurate and precise measurements of these 2 peptide hormones. Representatives of the Growth Hormone Research Society, the IGF Society, and the IFCC convened an international workshop to review assay standardization, requirements for improving assay comparability, variables that affect assay interpretation, technical factors affecting assay performance, assay validation criteria, and the development and use of normative data. Special attention was given to preanalytical conditions, the use of international commutable reference standards, antibody specificity, matrix requirements, QC analysis, and interference by binding proteins. Recommendations for each of these variables were made for measurements of each peptide. Additionally, specific criteria for IGF-I were recommended for age ranges of normative data, consideration of Tanner staging, and consideration of the effect of body mass index. The consensus statement concludes that major improvements are necessary in the areas of assay performance and comparability. This group recommends that a commutable standard for each assay be implemented for worldwide use and that its recommendations be applied to accomplish the task of providing reliable and clinically useful results.

Measurement of circulating growth hormone (GH)[2] and insulin-like growth factor I (IGF-I) concentrations represents a cornerstone of the diagnosis and monitoring of pituitary and other medical disorders, which include such GH-related disorders as acromegaly and GH deficiency, as well as numerous metabolic disorders *(1–5)*.

Considerable differences exist between the currently available assays with respect to the results of GH and IGF-I measurements *(6–9)*. These differences are largely due to heterogeneity in assay characteristics. Harmonization and comparability of GH and IGF-I assays are required to provide transparent and useful guidelines for the management of disorders of the GH–IGF-I axis *(10)*. The Growth Hormone Research Society, in collaboration with the IFCC, the International Society for IGF Research, and the Pituitary Society, organized an expert workshop to define criteria, strategies, and ways to implement harmonization of GH and IGF-I assays. This document reflects the recommendations of that workshop.

Current Obstacles to GH and IGF-I Assay Standardization

To address the fundamental obstacles to assay standardization requires an understanding of the properties and limitations inherent in immunoassays. The lack of GH and IGF-I assay standardization has led to major differences in the values of hormone concentrations obtained with different assays *(6, 7)*. An important contributor is the use of different calibrator materials *(11)*. Currently, not all GH assays are calibrated to a common international reference preparation *(7)*. Furthermore, the available standard for IGF-I assays, 87/518 (available from the WHO), is recognized as having an incorrectly assigned concentration *(12)*. Heterogeneous analytes pose an additional problem. Circulating GH, for example, consists of various forms, including monomers, dimers, and other posttranslationally modified products, the detection of which varies among assays *(13)*. Therefore, an individual GH or

[1] Department of Medicine, University of North Carolina School of Medicine, Chapel Hill, NC.
* Address correspondence to the author at: CB# 7170, 8025 Burnett Womack, Division of Endocrinology, University of North Carolina, Chapel Hill, NC 27599-7170. Fax 919-966-6025; e-mail endo@med.unc.edu.

Received May 17, 2010; accepted December 14, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.150631

[2] Nonstandard abbreviations: GH, growth hormone; IGF-I, insulin-like growth factor I; GHBP, GH-binding protein; IGFBP, IGF-binding protein; IS, international standard; NIBSC, National Institute of Biological Standards and Control.

IGF-I sample measured by different assays can yield very different results, because assays differ with respect to the epitope specificity of the antibodies used and because the different types of antibodies vary in their ability to bind to the different molecular forms of GH and IGF-I. The reporting of assay results also varies *(14)*. For example, GH assay results have been expressed not only in mass units but also in international units, which have been arbitrarily defined and do not have a clear relationship to mass. GH measurements are subject to interference by a number of factors, including GH-binding protein (GHBP), which corresponds to the circulating extracellular domain of the GH receptor *(15)*. Similarly, IGF-binding proteins (IGFBPs) may interfere with IGF-I assays *(6)*. As with any assay, matrix components can also variably interfere with both GH and IGF-I measurements. The effects of these variables on assay results are not always apparent to the assay user.

General Requirements for Improving Assay Comparability

The foremost fundamental requirement for improving the comparability of GH and IGF-I assay methods is to use a single universally accepted standard for each hormone assay. Each GH or IGF-I assay should also specify the analyte being measured, the specificities of the antibodies used, and the assay's susceptibility to binding protein interference. The epitope of the antibody should be stated if it is known. Normative data are needed for each GH and IGF-I assay. Assay comparability also requires standardization of preanalytical conditions, an understanding of matrix properties, implementation of appropriate internal QC procedures, and participation in an external QC program.

GH Assay

PREANALYTICAL CONDITIONS

Although some analytes are susceptible to many physical conditions, GH and IGF-I are relatively stable *(16)*. Their stability in blood is such that separation of serum from blood cells is recommended within 2 h of collection. GH and IGF-I are stable in serum for several weeks at −20°C. The clinical conditions under which GH should be sampled during a dynamic test for GH deficiency or acromegaly will vary with the test and the patient population being assessed. Performance of these dynamic tests should be guided by published protocols *(17–19)*.

STANDARDS AND ANTIBODY SPECIFICITY

Discrepancies between GH assays can be reduced if a common pure standard preparation is used for calibration *(11)*. A new preparation has recently become available, the Second International Standard for somatropin (a recombinant DNA–derived human GH, international standard (IS) 98/574), and is being used by manufacturers for standardization of GH assays. It is available from National Institute of Biological Standards and Control (NIBSC). This material is of high purity (>96% 22-kDa GH) with desirable characteristics (bioactivity, stability, availability, and so forth) *(20)*, but a demonstration of its commutability has not yet been published *(21)*. In view of the importance of the validation of commutability *(22)*, publication of a peer-reviewed commutability study is strongly recommended. If it is commutable among the various assay methods with patients' clinical samples, we recommend that all manufacturers adopt this standard to provide consistency in standardization. IS 98/574 has been assigned a unitage of 1.95 mg per ampoule. By definition, 1 mg equals 3 IU. We recommended that GH concentrations be reported in mass units.

The ideal assay should use antibodies of high affinity and specificity. Assays should be specific for the 22-kDa form of human GH. Reproducibility may be the most critical near the lower limit of the analytical interval. Growth hormone is secreted in a pulsatile fashion, and some values can be very low. The ability of glucose to suppress GH is used in the diagnosis and monitoring of acromegaly; therefore, it is important that assay manufacturers are able to provide accurate measurements at the lower end of the GH assay interval. Presently, we recommend that assays achieve a lower limit of quantification of 0.05 μg/L with a CV of <20%.

MATRIX REQUIREMENTS

The matrix used in calibrators should mimic the properties of nonpathologic human serum as closely as possible; however, that may still leave a residual matrix-related bias requiring compensation in the value assigned to the calibrator used with a specific assay procedure. Because no reference measurement procedure exists for GH (or IGF-I), the demonstration of equivalent results among a group of routine methods for a panel of patient samples will be necessary when method-specific calibrators are used. If such a panel is unavailable, it may be necessary to use serum from another species that maintains optimal linearity and recovery of known quantities of the hormone.

GHBP INTERFERENCE

Each assay should specify its degree of interference by GHBP within its physiological interval *(23)*. To investigate this interference requires that samples be spiked with recombinant GHBP and incubated for at least 12 hours to allow formation of the GH/GHBP complex before measurement *(24)*.

QUALITY CONTROL

Laboratories should use internal QC materials independent of those provided by the assay manufacturer. All laboratories measuring GH should participate in an accredited proficiency testing/external quality assessment program that uses materials that have been proved commutable at a national, and ultimately at an international, level. The availability of such a program at an international level would accelerate harmonization.

When commutable materials are available, we encourage the IFCC, in collaboration with national colleges of pathology (including the College of American Pathologists), to facilitate the worldwide exchange of proficiency testing/external quality assessment information and reports for the sharing of best practices. A desirable goal is to develop a centralized advisory group that will facilitate such sharing at an international level. Best practice for thresholds for variation between methods needs to be established to ensure comparability of quality. The characterization of each assay must include studies that define the reactivity of each of the major forms of GH in that assay.

Large pools of longitudinally available QC sera with documented commutability and covering concentrations relevant to clinically important cutoff points, such as those for the diagnosis of GH deficiency, need to be available and distributed across laboratories for the purpose of comparative studies.

PRESENTATION OF ASSAY RESULTS

The laboratory report presented to the clinician should give GH results in mass units (i.e., micrograms per liter). Because discrepancies can remain despite the adoption of a uniform standard, potentially because of differences in reagents, some laboratories have considered adopting conversion factors to compare the results of different assays. Such conversion factors do not account for all assay differences, however, and their use is discouraged.

COOPERATIVE EFFORTS AND EDUCATION

Implementation of the use of IS 98/574 requires the bringing together of representatives of professional organizations, manufacturers, proficiency-testing providers, and key opinion leaders to endorse this effort. Use of the reference standard can also be encouraged by educational initiatives that will disseminate information on its use, as well as by the adoption of policies by journals that require use of the internationally adopted standard in GH assays in studies they publish.

Information available in a manufacturer's assay literature should specify the following: the identity of the reference preparation; calibration information, including traceability to the international reference standard; assay cross-reactivity with 20-kDa GH, placental GH, and therapeutic GH analogs; GHBP interference; limits of quantification; assay performance at the upper and lower limits of the normative reference intervals; and QC data. A manufacturer must communicate to the laboratories major changes in assay components or performance that could affect clinical diagnostic or therapeutic decisions. Laboratories then must communicate this information to clinicians in a timely manner.

IGF-I Assay

CHARACTERISTICS OF THE IGF-I REFERENCE STANDARD

The IS 02/254 WHO reference standard has recently become available. It is a >97%-pure recombinant standard and has been well characterized by the NIBSC. IS 02/254 has been analyzed in >20 laboratories and has been shown to be bioactive and stable *(12, 21, 25)*. If this standard is documented to be commutable with patients' clinical samples among the various assay methods, then journals, societies, and regulatory authorities should advocate the use of the new standard to minimize the interassay differences that arise from the use of different standards. Although the use of conversion factors is not encouraged, their use will be necessary during the transition period. Conversion factors should be used only within the same reagent set, however, not for conversion of IGF-I concentrations between methods. If the standard is commutable, assays should be recalibrated with the new standard, and the normative interval reestablished.

CHARACTERISTICS OF QC SAMPLES

As part of the validation procedure, we recommend the establishment of a set of reference samples, to be distributed internationally to monitor the concordance between assay methods. These reference samples will consist of human serum pools that contain low, medium, and high concentrations of IGF-I and GH. Samples from patients with a variety of conditions in which IGFBPs can influence IGF-I assay results (e.g., type 1 diabetes mellitus, chronic renal failure, cirrhosis, GH disorders) should be used to directly compare the results of IGF-I measurements obtained with different assays *(6)*. This information should be made available to end users, as should information regarding assay performance characteristics and interference by IGFBPs.

THE BIOLOGICAL VARIATION OF IGF-I

Given the 3%–36% within-individual biological imprecision of IGF-I assays *(26)*, caution should be exercised in the interpretation of a single IGF-I value, par-

ticularly if it is close to a reference limit. Assays for samples that yield borderline results should be repeated with a separate blood sample. The clinical context of the patient needs to be incorporated into the interpretation of the result. Consideration should be given to performing other tests, such as dynamic GH stimulation/suppression tests.

VARIABLES THAT AFFECT ASSAY INTERPRETATION

Multiple physiological variables and clinical conditions can affect the concentrations of IGF-I. Age, pubertal stage, pregnancy, and extremes of body mass index are especially important and need to be taken into consideration when interpreting results. In adults, sex and ethnicity are minor contributors to variation in IGF-I concentrations. Circadian and meal-related changes do not appear to affect IGF-I measurements *(6)*. Assay validation should be undertaken for such conditions as type I diabetes mellitus and chronic renal failure. Specific diseases, such as diabetes, hepatic and renal disease, and nutritional compromise, can affect IGF-I concentrations; therefore, caution should be used in interpreting results for patients with these conditions. Some medications, such as oral estrogen, can also affect IGF-I concentrations *(27)*.

TECHNICAL FACTORS AFFECTING ASSAY PERFORMANCE

Protocols for sample collection and handling should be provided for each assay method. The use of serum is recommended. The use of anticoagulants in the collection tube will require separate validation. Blood samples should be processed within 2 h to avoid an artifactual increase in results.

A major factor affecting assay performance is the presence of IGFBPs in the sample. An assay that recognizes IGF-I without any interference from binding proteins would be ideal. In general, IGFBPs interfere with IGF-I detection to produce falsely low values. In practice, this problem is usually managed by the use of a reagent that dissociates IGF-I from the IGFBPs, followed by preventing reassociation of the IGFBPs with IGF-I.

The method for preventing IGFBP interference should be validated and stated for each assay. The current reference method for eliminating IGFBP interference is gel chromatography at low pH. All methods must demonstrate that IGFBP interference has been substantially removed (>95%), and the clinical conditions under which residual IGFBPs may still cause interference should be stated.

Physicians should be informed of the technical specifications of the assay with regard to specific clinical conditions. The clinician is also responsible for being aware of any clinical conditions *(6)* that make the interpretation of results for total serum IGF-I challenging.

VALIDATION OF ASSAY PERFORMANCE

The manufacturer should use nonpathologic sera, as well as sera from patients with GH disorders, type 1 diabetes mellitus, chronic renal failure, and cirrhosis, during assay validation, and they should report this information. Assay results should be shown to be equivalent to those obtained after gel chromatography performed at low pH. Before processing (as recommended by the kit manufacturer), samples should be spiked with a range of IGFBP concentrations (IGFBP-2 and IGFBP-3) up to twice the upper limit of the reference interval to ensure the absence of interference. Properly designed studies should document the recovery of added highly purified IGF-I *(6)*.

NORMATIVE DATA

Normative data should be based on a random selection of individuals from the background population, and this population sample should include representation from all age groups. Individuals with medical conditions (e.g., cirrhosis, poorly controlled diabetes, renal failure) and taking medications (e.g., estrogen) that may affect the outcome are excluded. Normative data should include the central 95% interval (2.5–97.5 percentiles) and be reported in mass units and, after appropriate transformation for data nonnormality, as SD scores (equivalent to the *z* score, which represents the number of SDs a given result is above or below the age-adjusted mean).

The IGF-I concentration shows the greatest change during childhood and puberty and then changes more slowly with advancing age. For the generation of normative data, the stratification of age groups should be based on properly designed studies and statistical analysis of normative data *(28)*. Special considerations for children and adolescents should include narrow age ranges (e.g., every 3 years) and Tanner stages. Studies must include adequate numbers of reference individuals in the age intervals with rapidly changing IGF-I concentrations to ensure reliable estimates of reference intervals at these ages *(28)*.

Sex-specific IGF-I reference intervals are required for ages between 6 and 18 years *(6)*. There is little change in IGF-I across the interval of body mass index values from approximately 22 to 37 kg/m^2, but significantly lower IGF-I values have been observed above and below these values. This fact should be taken into account when interpreting results *(6)*.

GH stimulation and suppression tests are valuable tools in the diagnosis of GH deficiency and excess, respectively. As long as different assays give different results for GH, however, it is mandatory to use appropri-

ately validated, assay-specific reference data for these studies. These considerations should also include the effect of body mass index and age. The development of normative data for children is recognized to be challenging, for a variety of reasons, and is limited by the ability to recruit healthy children into such studies.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures or Potential Conflicts of Interest: *No authors declared any potential conflicts of interest.*

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: The participants thank Jane Fruchtnicht for her help in organizing the conference and Laura Lindsey for her help in preparing the manuscript. Contributors to the Consensus Statement: Alecia Algeciras-Schimnich, Margherita Banci, Gerhard Baumann, Robert Baxter, Martin Bidlingmaier, Beverly Biller, David E. Bruns, Felipe Casanueva, Phillippe Chanson, Jens Christiansen, Peter Clayton, David Clemmons, Pinchas Cohen, Cheri Deal, Andy Ellis, Robin A. Felder, Pamela Freda, Jan Frystyk, Richard Fulanetto, H. Mario Geysen, Shelia Hanna, Philip Harris, Ken K.Y. Ho, Andrew Hoffman, Jeff Holly, Reiko Horikawa, Gudmundur Johnannsson, Anna-Marie Kappelgaard, Susan Kirshner, David Kleinberg, John Kopchick, Derek LeRoith, Saul Malozowski, Kathy Maugh, A. Wayne Meikle, Shlomo Melmed, Pharis Mohideen, Clement Olivier, Bernhard Saller, Rudolf Schemer, Akira Shimatsu, Christian Strasburger, Catharine M. Sturgeon, Mario Thevis, Michael Thorner, Peter Trainer, Yi Tsong, Mary Lee Vance, Gwen Wark, and Ingrid Zegers. We thank James C. Boyd for discussions regarding the determination of reference intervals and *z* scores.

References

1. LeRoith D, Yakar S. Mechanisms of disease. Metabolic effects of growth hormone and insulin-like growth factor 1. Nat Clin Pract Endocrinol Metab 2007;3:302–10.
2. Clemmons DR. Modifying IGF1 activity: an approach to treat endocrine disorders, atherosclerosis and cancer. Nat Rev Drug Discov 2007;6:821–33.
3. Ho KKY, on behalf of the 2007 GH Deficiency Consensus Workshop participants. Consensus guidelines for the diagnosis and treatment of adults with GH deficiency II: a statement of the GH Research Society in association with the European Society for Pediatric Endocrinology, Lawson Wilkins Society, European Society of Endocrinology, Japan Endocrine Society, and Endocrine Society of Australia. Eur J Endocrinol 2007;157:695–700.
4. Melmed S, Casanueva F, Cavagnini F, Chanson P, Frohman L, Grossman A, et al. Guidelines for acromegaly management. J Clin Endocrinol Metab 2002;87:4054–8.
5. Giustina A, Barkan A, Channson P, Grossman A, Hoffman A, Ghigo E, et al. Guidelines for the treatment of growth hormone excess and growth hormone deficiency in adults. J Endocrinol Invest 2008;31:820–38.
6. Frystyk J, Freda P, Clemmons DR. The current status of IGF-I assays—a 2009 update. Growth Horm IGF Res 2009;20:8–18.
7. Bidlingmaier M, Freda PU. Measurement of human growth hormone by immunoassays: current status, unsolved problems and clinical consequences. Growth Horm IGF Res 2009;20:19–25.
8. Seth J, Ellis A, Al-Sadie R. Serum growth hormone measurements in clinical practice: an audit of performance from the UK National External Quality Assessment scheme. Horm Res 1999;51:13–9.
9. Morsky P, Tiikkainen U, Ruokonen A, Markkanen H. Problematic determination of serum growth hormone: experience from external quality assurance surveys 1998-2003. Scand J Clin Lab Invest 2005;65:377–86.
10. Trainer PJ, Barth J, Sturgeon C, Wieringaon G. Consensus statement on the standardisation of GH assays. Eur J Endocrinol 2006;155:1–2.
11. Tanaka T, Tachibana K, Shimatsu A, Katsumata N, Tsushima T, Hizuka N, et al. A nationwide attempt to standardize growth hormone assays. Horm Res 2005;64(Suppl 2):6–11.
12. Quarmby V, Quan C. How much insulin-like growth factor (IGF-I) circulates?: impact of standardization on IGF-I assay accuracy. Dev Biol Stand 1999;97:111–8.
13. Baumann G. Growth hormone heterogeneity: genes, isohormones, variants, and binding proteins. Endocr Rev 1991;12:424–49.
14. Pokrajac A, Wark G, Ellis AR, Wear J, Wieringa JE, Trainer PJ. Variation in GH and IGF-I assays limits the applicability of international consensus criteria to local practice. Clin Endocrinol (Oxf) 2007;67:65–70.
15. Ebdrup L, Fisker S, Sorensen HH, Ranke MB, Ørskov H. Variety in growth hormone determinations due to use of different immunoassays and to the interference of growth hormone-binding protein. Horm Res 1999;51:20–6.
16. Derr RL, Cameron SJ, Golden SH. Pre-analytic considerations for the proper assessment of hormones of the hypothalamic-pituitary axis in epidemiological research. Eur J Epidemiol 2006;21:217–26.
17. Molitch ME, Clemmons DR, Malozowski S, Merriam GR, Shalet SM, Vance ML, Stephens PA. Evaluation and treatment of adult growth hormone deficiency: an Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metab 2006;91:1621–34.
18. Growth Hormone Research Society. Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: summary statement of the GH Research Society. J Clin Endocrinol Metab 2000;85:3990–3.
19. Giustina A, Barkan A, Casanueva FF, Cavagnini F, Frohman L, Ho K, et al. Criteria for cure of acromegaly: a consensus statement. J Clin Endocrinol Metab 2000;85:526–9.
20. Bristow AF, Jespersen AM. The Second International Standard for somatropin (recombinant DNA-derived human growth hormone): preparation and calibration in an international collaborative study. Biologicals 2001;29:97–106.
21. Sturgeon C, Ellis AR, Al-Sadie R. UK NEQAS for peptide hormones: annual review for 2009. Edinburgh (UK): National External Quality Assessment Service; 2010. p 22–6.
22. Miller WG, Myers GL, Rej R. Why commutability matters. Clin Chem 2006;52:553–4.
23. Jan T, Shaw MA, Baumann G. Effects of growth hormone-binding proteins on serum growth hormone measurements. J Clin Endocrinol Metab 1991;72:387–91.
24. Popii V, Baumann G. Laboratory measurements of growth hormone. Clin Chim Acta 2004;350:1–16.
25. Burns C, Rigsby P, Moore M, Rafferty B. The First International Standard for Insulin-like Growth Factor-I (IGF-I) for immunoassay: preparation and calibration in an international collaborative study. Growth Horm IGF Res 2009;19:457–62.
26. Milani D, Carmichael JD, Welkowitz J, Ferris S, Reitz RE, Danoff A, et al. Variability and reliability of single serum IGF-I measurements: impact on determining predictability of risk ratios in disease development. J Clin Endocrinol Metab 2004;89:2271–4.
27. Clemmons DR. IGF-I assays: current assay methodologies and their limitations. Pituitary 2007;10:121–8.
28. CLSI (formerly NCCLS). Defining, establishing and verifying reference intervals in the clinical laboratory: approved guidelines – third edition. CLSI document C28A3E. Wayne (PA): CLSI; 2008.