Reference values for IGF-I serum concentrations: Comparison of six immunoassays

P. Chanson, A. Arnoux, M. Mavromati, S. Brailly-Tabard, C. Massart, J. Young, M.-L. Piketty, J.-C. Souberbielle

► To cite this version:


HAL Id: hal-01380421
https://hal-univ-rennes1.archives-ouvertes.fr/hal-01380421
Submitted on 18 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Reference values for insulin-like growth factor I (IGF-I) serum concentrations:

comparison of six immunoassays.

(Short title: Reference values for IGF-I with 6 immunoassays)

Philippe Chanson, Armelle Arnoux, Maria Mavromati, Sylvie Brailly-Tabard, Catherine Massart, Jacques Young, Marie-Liesse Piketty, and Jean-Claude Souberbielle for the VARIETE Investigators*. 

Assistance Publique-Hôpitaux de Paris, Hôpitaux universitaires Paris-Sud, Hôpital de Bicêtre, Service d'Endocrinologie et des Maladies de la Reproduction and Centre de Référence des Maladies Endocriniennes Rares de la Croissance (P.C., M.M., J.Y.), and Unité de Recherche Clinique (A.A.), and Service de Génétique Moléculaire, Pharmacogénétique et Hormonologie (S.B.T), Le Kremlin-Bicêtre, F94275, France; Inserm 1185 (P.C., S.B.T., J.Y.), Fac Med Paris Sud, Université Paris-Saclay, Le Kremlin-Bicêtre, F-94276, France; and Centre Hospitalier Universitaire de Rennes, Centre d'Investigation Clinique plurithématique, Inserm 1414, Hôpital Pontchaillou, Laboratoire d'Hormonologie (C.M.), Rennes, F29000, France; and Assistance Publique-Hôpitaux de Paris, Hôpital Necker-Enfants Malades, Service des Explorations Fonctionnelles (M-L.P., J-C.S.), Paris, F75015, France

* a complete list of the Valeurs de Référence de l'IGF-I Et Transformation En Z-score (VARIETE) study investigators is given in Supplemental Appendix

Corresponding author and reprint requests:

Philippe Chanson, MD
Abbreviations: BMI, body mass index; IGFBP, IGF binding protein; SDS, standard deviation score.

Key terms: IGF-I, Z-score, SD score, normative data, reference range, normal healthy population, acromegaly, growth hormone deficiency

Disclosure: The authors have no conflicts of interest to disclose

Word count (for text only): 3581

Abbreviated title: Reference intervals for serum IGF-I
Summary

Context. Measurement of IGF-I is essential for diagnosis and management of patients with disorders affecting the somatotropic axis. However, even when IGF-I kit manufacturers follow recent consensus guidelines, different kits can give very different results for a given sample.

Objectives. We sought to establish normative data for six IGF-I assay kits, based on a large random sample of the French general adult population.

Subjects and Methods: In a cross-sectional multicenter cohort study (ClinicalTrials.gov Identifier: NCT01831648), we measured IGF-I in 911 healthy adults (18-90 years) with six immunoassays (iSYS, LIAISON XL, IMMULITE, IGFI RIACT, Medagnost ELISA, and Medagnost RIA). Pairwise concordance between assays was assessed with Bland-Altman plots for both IGF-I raw data and standard deviation scores (SDS), as well as with the percentage of observed agreement and the weighted Kappa coefficient for categorized IGF-I SDS.

Results: Normative data included the range of values (2.5 to 97.5 percentiles) given by the six IGF-I assays according to age group and sex. A formula for SDS calculation is provided. While the lower limits of the reference intervals of the six assays were similar, the upper limits varied markedly. Pairwise concordances were moderate to good (0.38 to 0.70).

Conclusion. Despite being obtained in the same healthy population, the reference intervals of the six commercial IGF-I assay kits showed noteworthy differences. Agreement between methods was moderate to good.
Growth hormone (GH) exerts its effects on target tissues either directly or via the production of insulin-like growth factor 1 (IGF-I). Accurate measurement of IGF-I in serum is crucial for diagnosis and management of disorders affecting the somatotropic axis, particularly GH excess (acromegaly) and GH deficiency (GHD). However, even if manufacturers follow the recommendations of the Consensus Group on the Standardization and Evaluation of GH and IGF-I Assays (1), the different commercial IGF-I assay kits can give very different results for the same sample, with up to a 2.5-fold difference between the lowest and highest values (2). This inter-method variability is generally explained by calibration against different IGF-I reference preparations (3), and differences in the efficiency of methods used to remove IGF-binding proteins (IGFBPs) (4). In theory, this should not be a problem in clinical practice, as kits that give higher values should have higher normal limits, and patients should thus be consistently classified.

However, it is very difficult to establish reference values for IGF-I. Indeed, serum IGF-I concentrations increase with children's age and pubertal stage, while they fall with age in adults (5). Furthermore, the distribution of IGF-I values in an apparently healthy population is non-Gaussian, and this necessitates complex mathematical transformation to obtain reference intervals for each age group. For this reason, it is essential to generate reference values after stratifying a large healthy population into age groups. Another problem is that IGF-I concentrations are influenced by many factors other than GH concentrations, including nutritional status and BMI, use of hormone replacement therapy by post-menopausal women, depending on the administration route (6-8), kidney and liver function, and diabetic status (9). Reference IGF-I values may therefore be influenced by the inclusion criteria used to select the reference population sample. This could have important implications for diagnosis and therapeutic decision-making, as a given patient could be classified as having a normal IGF-I concentration with one method but an abnormal value with another method. Several studies
suggest that the main reason for inter-laboratory variability in patient classification is the use of different populations to establish reference values for the different IGF-I assays (2,10,11).

It is currently difficult to monitor an individual patient with different IGF-I assays, even if the results are all expressed in the same units (ng/ml). It is thus recommended to establish specific reference ranges for each assay, and to apply common, well-defined inclusion criteria to the reference population (1). It is also recommended, for the comparison of values obtained with different assays in the same patient, to express each IGF-I result as an SD score (SDS) with reference to the normative data for the assay in question, after appropriate transformation for data non normality. We reasoned that the best way to overcome this variability would be to apply all the commercial kits used in clinical laboratories to a battery of samples from the same well-defined reference population, and to use the same mathematical transformation to calculate reference ranges from the raw data.

The aim of this study was thus to establish normative data for six commercial IGF-I assays in a large random sample of healthy subjects from the French general population representing all adult age groups (about 100 subjects per decade), as recommended by the Consensus Group on the Standardization and Evaluation of GH and IGF-I assays (1). Serum samples from the reference population were tested with six commercial assay kits available in France at the time of this study, after careful exclusion of subjects with medical conditions or medications that might affect their IGF-I concentration. The data were analyzed to obtain the range (2.5 to 97.5 percentiles) in mass units. The standard deviation scores were used to compare the six assays.

Subjects and Methods

IGF-I assay characteristics
Six immunoassays (iSYS, LIAISON XL, IMMULITE, IGF RI A CT, Medig nst ELISA, and Medig nst RIA) were used to measure the IGF-I concentration in each healthy subject. The main characteristics of the assays, and the mathematical models used to determine normative data, where relevant (12-14), are shown in Table 1.

Healthy subjects

The subjects were part of a large cohort of French healthy adults (VARIETE). The VARIETE cohort was an open, prospective, national, multicenter, non randomized study of healthy volunteers, designed to establish normative data for IGF-I and other hormones in the French general adult population representing all age groups (about 100 subjects per decade from 18 to 90 years) (ClinicalTrials.gov Identifier: NCT01831648). A total of 972 healthy subjects with BMI values between 19 and 28 kg/m² were recruited in 10 centers throughout France between 2010 and 2011. Our objective of including 1000 subjects was not achieved due to difficulties for obtaining an accurate number of subjects in the older age categories (>70 years) fulfilling all the inclusion criteria and without exclusion criteria before the end of our inclusion period. Subjects with medical conditions or medications that might affect IGF-I serum levels were excluded (see Supplemental Appendix). Each subject had a clinical examination, personal medical history-taking and general examination, including careful evaluation of nutritional and gonadal status. Standard laboratory tests (plasma sodium, potassium, calcium, phosphate and creatinine, glycemia, total cholesterol, liver enzymes, TSH, blood cell count, albuminemia, prothrombin time, as well as HIV and HCV serologies) were then performed, and 80 mL of blood (50 mL without anticoagulant and 30 mL in EDTA-containing tubes) was sampled and promptly centrifuged (2000 g, 4°C). Serum and plasma were aliquoted, frozen, and stored at -80°C until hormone measurements.
All healthy subjects gave their written informed consent to participate in the study, which was approved by the Paris-Sud Ethics committee before the beginning of the study.

**Statistical methods**

The distribution of IGF-1 values obtained with each assay was skewed, and was thus first normalized by means of sex- and age-specific Box-Cox power transformation. Student’s t test and Levene’s test were then used to assess equality of means and homogeneity of variances between men and women in each age group. As men and women had significantly different IGF-1 levels, centile curves were constructed separately for each sex.

Age- and sex-specific centile curves were constructed for each assay by using the LMS method (12) implemented in the GAMLSS software package version 4.3-1 (15) of R software version 3.1.2 (2014-10-31) (R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL http://www.R-project.org/). The LMS method enables smooth curves to be estimated for percentiles after normalization (by Box-Cox power transformation) and standardization of the data. The parameters L (for skewness), M (for median) and S (for the coefficient of variation) were also computed for each age and sex class. SD scores (SDS) were calculated as

\[
z = \frac{(\text{IGF-1} / M)^L - 1}{L \times S}
\]

where IGF-I is the raw value given by the assay (in ng/mL). For each technique, SDS were categorized as low, normal or high according to their positions relative to both the 2.5th and 97.5th percentiles.

Once the L, M and S parameters for each category of age and sex had been obtained, the lower and upper reference interval limits were determined for each assay by fixing z at -1.96 and 1.96, respectively, and then mathematically back-transforming the SD score formula.

Pairwise concordance between assays was assessed with scatter plots and Bland-Altman plots for both IGF-1 raw values and SDS values, as well as with the percentage of observed
agreement (total number of agreements divided by the total number of patients tested with both assays) and the linearly weighted Kappa coefficient for categorized IGF-1 SDS (16,17). An overall kappa coefficient (16) and Friedman’s test were computed for global comparison of all assays at the same time. Landis and Koch’s table was followed for interpretation of Kappa values (18).

Unless otherwise stated, SAS software was used for all statistical analyses (Statistical Analysis System, version 9.4, SAS Institute, Cary, N.C., USA).

Results

1- Description of the population

Nine hundred seventy-two subjects were initially recruited, of whom 52 were excluded because of abnormal values in the standard laboratory screening tests. A further 9 subjects were excluded because of missing information on pregnancy status or viral serology. The study population thus consisted of 911 subjects (470 males), comprising respectively 101, 118, 99, 98, 103, 102, 108, 97 and 85 subjects in the 18-20, 21-23, 24-26, 27-29, 30-39, 40-49, 50-59, 60-69, 70-89 year age groups. Mean BMI was 23.0 ± 2.4 kg/m².

2- IGF-I reference intervals obtained with the six assays

The IGF-I reference intervals (2.5th-97.5th percentiles) obtained with the six immunoassays are shown in Table 2 according to age and sex. Supplemental Figure 1 shows individual points and fitted percentiles (2.5%, 50% and 97.5%) for males and females in each IGF-I assay.

A calculator available online (http://ticemed_sa.upmc.fr/sd_score/) or by using Apps (IGF-I SD_score) downloadable for Android from Google Play and for iOS from Apple Store (free...
of charge) allows to obtain individual IGF-I SDS after entering the name of the assay, the individual IGF-I value obtained with the assay, and the sex and age of the individual.

The six reference intervals for males and females are plotted on the same graph in Figure 1. While the lower limits of the reference intervals (2.5\textsuperscript{th} percentiles) were similar, the upper limits (97.5\textsuperscript{th} percentiles) varied markedly from one assay to another.

3- Comparison of IGF-I levels given by the six assays

The results obtained with each IGF-I assay were compared with those obtained with each of the other five assays. Scatter plots and Bland-Altman plots based on raw values and SDS for each pair of assays are shown in Supplemental Figure 2.

Whatever the assay, IGF-I concentrations were generally higher in women than in men until the age of 59 years (this was significant for the age ranges 18-20 and 24-26 years). From the age of 60 years, IGF-I levels were slightly higher in men than in women, although the gender difference was smaller than in the younger age groups and was only significant for Immulite, Medagnost Elisa and Medagnost RIA.

Two examples of inter-assay comparisons are shown in Figure 2. The results obtained with iSYS and Medagnost RIA were in good overall agreement, with no significant bias as assessed by Bland-Altman plots (Figure 2 A, B, C and D). In contrast, the results obtained with LIAISON XL and Medagnost RIA were not in good agreement (Figure 2 E, F, G and H).

Pairwise assay concordances assessed with the weighted Kappa coefficient for categorized IGF-1 SDS are shown in Table 3. The concordances were moderate to good (0.38 to 0.70), although the percentages of observed agreement were quite high (94\% to 97\%).

Overall agreement was moderate as overall Kappa coefficient was 0.55. Both in men and women, global inter-assay comparison showed significant differences (p<0.0001) on raw values but not on SDS values (p=0.26 and p=0.36, respectively).
Table 4 shows pairwise concordances between the reference intervals provided by the manufacturer and those obtained in the VARIETE cohort, as assessed by the Kappa coefficient and the percentage agreement for each IGF-I assay. The concordances and percentages of observed agreement were generally poor.

Discussion

We report reference intervals for IGF-I concentrations obtained with six immunoassays in the same population of nearly 900 French healthy subjects aged from 18 to 90 years, in keeping with the 2011 recommendations of the Consensus Group on the Standardization and Evaluation of GH and IGF-I assays (1). The population comprised about 100 subjects per age decade, and specific reference intervals were calculated for each sex and age group. The reference intervals varied from one assay to another: the lower limits of the normal range (2.5\textsuperscript{th} percentile) were quite similar with the six methods, but the upper limits (97.5\textsuperscript{th} percentile) varied widely from one assay to another, in both men and women (Figure 1). Although the pre-analytic conditions were the same for the six kits, and although four of the six kits were calibrated against the international reference standard 02/254, concordance between the assays, as assessed with Bland-Altman plots and the Kappa coefficient, remained quite variable, not only for raw IGF-I values but also for IGF-I SDS. This latter result was somewhat surprising, as we expected that, by using the same healthy population, we would obtain similar SDS.

In table 2, which shows the reference ranges for each assay, we have deliberately omitted the mean and SD calculated for each age category from the raw values, in order to avoid erroneous calculations of SDS. Indeed, the Box-Cox power transformation, which is necessary because of the non Gaussian distribution in each age category, uses parameters (L for skewness, M for median and S for the coefficient of variation) that are specific to each
assay and also to each age group and gender. We thus propose an online calculator available either following this link (http://ticemed_sa.upmc.fr/sd_score/) or by using Apps (IGF-I SD_score) downloadable for Android from Google Play and for iOS from Apple Store (free of charge) which allows to determine SDS as a function of the assay method, the measured IGF-I value, gender, and age. L, M and S parameters are also provided in Supplemental Table 1.

Reliable reference intervals are crucial for interpreting IGF-I values in adults with acromegaly (for diagnosis and assessment of disease control during treatment), and also for diagnosing GH deficiency and monitoring GH therapy (4,5,19,20). Reference intervals obtained with the IGF-I Nichols Advantage assay in a very large population of healthy subjects (21) were once widely used for research and clinical practice. However, market withdrawal of this assay, together with the availability of numerous automated methods with considerable heterogeneity, led to calls for improved comparability and reliable normative data. One important first step was the creation of the recombinant international IGF-I standard preparation 02/254 (22). A consensus conference held in 2011 proposed that all assays be calibrated against this standard, and advocated precise pre-analytical and analytical conditions (1). Another recommendation was to establish normative data based on a random selection of individuals from the background population, with representation of all age groups (1). The first normative data for the iSYS IGF-I assay, based on these recommendations and on a very large healthy population, were published by Bidlingmaier et al (23). We now propose reference intervals for six IGF-I assays also based on a large population of healthy subjects. It should be noted that we used very stringent inclusion criteria. Indeed, despite the large sample size (almost one thousand healthy subjects, with about 100 subjects per age group), all the subjects had a clinical examination, including assessment of gonadal status, and also a careful medical history taking that included ongoing medications. Furthermore, all the subjects had
an extensive standard biological work-up in order to exclude those with disorders capable of influencing IGF-I levels or their measurement. These very strict inclusion and exclusion criteria allow to define a population as “healthy” as possible; however this implies that these normative data will not be strictly applicable to patients with BMI > 28 kg/m² or to patients with oral treatment with estrogens.

As expected, IGF-I concentrations fell gradually with age in both sexes, irrespective of the assay. Contrary to previous reports (21,23), we found a gender difference, with higher IGF-I levels in women than in men, whatever the assay, until the 5th decade. After 50 years of age, however, IGF-I levels were higher in men than in women, as reported elsewhere (21,23). We therefore propose separate normative data for men and women. One possible explanation for the discrepancy between this work and previous reports is that we excluded all subjects receiving steroid hormones such as estrogens. Indeed, oral estrogen is known to lower IGF-I levels (6-8). In premenopausal women, for example, contraceptive pills containing ethinyl estradiol reduce IGF-I levels by up to an average of 30% (24-27). Another explanation might be the size of our population. Indeed, in their study involving a larger number of subjects (15,000), Bidlingmaier et al. did not find differences in terms of gender differences (23).

Inter-assay differences in IGF-I reference intervals are a well-known issue that has previously been underlined by one of us (28,29) and by many other researchers (2,11,23,30,31). In this study, as expected, the largest inter-centile intervals (and highest values) were obtained with the two assays calibrated with the old standard IRP 87/518 (IMMULITE and IGFI RIACT). Moreover, the three automated methods (iSYS, Liaison XL and IMMULITE), which should theoretically be the most reproducible, did not yield narrower reference intervals. For example, the iSYS automated method and the Mediagnost RIA manual method gave very similar intervals for both men and women in all age groups. Thus, the main source of variation does not appear to be analytical reproducibility. Using the same
iSYS method and a similar transformation for normalizing data and constructing specific
centile curves in the LMS method, our 2.5th and 97.5th percentiles were generally slightly
higher and our intervals generally narrower than those reported by Bidlingmaier et al. (23).
Although inter-laboratory variability may play a role in these discrepancies, they are likely
due mainly to differences in the population samples (our population was smaller, and the
inclusion criteria were different). Another issue raised by our study is the poor concordance
between our reference intervals and those proposed by the assay manufacturers. Once again
this might be related to the use of different background populations: indeed, those used by
manufacturers may not fulfill all the criteria recommended by the consensus group in 2011,
particularly with respect to their size, the definition of healthy subjects, and the use of
hormonal contraceptives (Supplemental Material).
Likewise, one obvious explanation for the discordance between assays is the use of different
populations to establish reference intervals. This is why we used the same reference
population for all the kits. However, although the six assays showed comparable analytical
performance in terms of their reproducibility and detection limits (Table 1), and despite the
fact that they use the same non-competitive “sandwich” format and similar methods to avoid
IGFBP interference (IGF-II addition), the reference values obtained in our well-controlled
adult population differed strikingly from one assay to another. Two of the six assays
(IMMULITE and IGF-I RIACT) are still calibrated against the old IRR 87/518 standard,
whereas the other four are calibrated against the new IRR 02/254 standard, as currently
recommended (1). As expected, the former two assays gave the highest upper reference range
for both sexes until the age of 50 (Table 2, Figure I). However, the reference ranges of two
differently calibrated kits may be either similar (e.g. LIAISON XL and IGFRIACT in men),
or significantly different (e.g. iSYS lower than IMMULITE) (Table 2). Likewise, reference
ranges determined with kits calibrated against the same reference preparation may also be
significantly different, even for kits from the same manufacturer (e.g. the RIA and ELISA kits from Mediagnost). It therefore seems likely that the observed differences are related to other analytical factors, such as the efficiency of IGFBP interference removal and the specificity and/or affinity of the antibody used. For example, since the 2.5th percentile is at least similar between the assays, the broadening of the interval for the IMMULITE assay is probably not related to the calibrator, but to relatively higher measurement results at the upper end: an explanation could be that IMMULITE assay preferentially recognizes the high free IGF-I at high concentrations, while the other 2 assays more efficiently remove the impact of BPs.

This could have important implications in patients with disorders affecting their IGFBP profile, such as acromegaly and chronic kidney disease. If confirmed in further studies, this implies that a given individual must be monitored with the same IGF-I assay.

Another limitation of our study is that it lies on a single measurement of IGF-I while it is well known that there is some within-subject variability when an individual is sampled on different days (32,33).

What refinements may be expected in the measurement of this very demanding analyte? The LC-MSMS method may prove to be a valid alternative and is now being used to assess inter-laboratory agreement on IGF-I concentrations (34) or for validation of IGF-I measures (35). Reference intervals for IGF-I provided with this LC-MS (36) seem very comparable with those obtained with immunoassays. When compared with our data, lower limit of normal range is similar and upper limit corresponds more or less with those observed with Liaison XL or IGF1 RIACT immunoassays. However, LC-MSMS is a time-consuming and complex method that requires expensive machines and high technical expertise, because many variables need to be controlled for providing accurate quantitative results (e.g. extraction strategies, approaches to detect and quantify IGF-I, calibration protocols…) (37). Furthermore, a recent preliminary study of an LC-MS method suggested
that it might miss some IGF-I protein variants (pathogenic or physiological), which are present in 0.6% of the population (38). Thus, despite their limitations, immunoassays will continue to be widely used, at least in the near future (39).

In conclusion, we have established reference intervals for six commercial IGF-I assays, in a study conforming to recent international recommendations. Despite being obtained in the same large population of French healthy subjects, the reference intervals differed somewhat from one assay to another, and agreement between assays was moderate to good. Finally, concordances between the manufacturers’ reference intervals and those obtained in our cohort were generally poor. These findings confirm the need to establish reference intervals for each commercial IGF-I assay in a large background population. Inter-assay concordance with respect to the classification of patients with acromegaly or GH deficiency remains to be determined, and the IGF-I standard deviation scores obtained with the six assays in these subjects need to be compared.

Acknowledgments

The study was supported by a grant from Programme Hospitalier de Recherche Clinique, French Ministry of Health, N° P081216 / IDRCB 2009-A00892-55 and by a grant from Fond National Suisse, P1GEP3_155694, to Dr Maria Mavromati.

We thank for Dr Hélène Agostini and Prof. Bruno Falissard for their helpful comments. We also thank Cisbio International, DiaSorin, IDS, Mediagnost, and Siemens for the kind donation of IGF-I kits.
Legends of Figures

Figure 1.
Reference intervals (Upper panel, males; lower panel, females) according to the age intervals of the 6 immunoassays tested. Lower limits (2.5\textsuperscript{th} percentile) and upper limits (97.5\textsuperscript{th} percentile) of the normal range are drawn as full lines and means as dotted lines.

Figure 2.
Comparisons between iSYS and Mediagnost RIA expressed as scatter plots (A) or Bland-Altman plots (B) for raw data, or scatter plots (C) and Bland-Altman plots (D) for SDS showing a good overall agreement between both immunoassays, with no significant bias. Comparisons between Liaison XL and Mediagnost RIA expressed as scatter plots (E) or Bland-Altman plots (F) for raw data, or scatter plots (G) and Bland-Altman plots (H) for SDS showing a bad overall agreement between these two immunoassays.
References


8. Meinhardt UJ, Ho KK. Modulation of growth hormone action by sex steroids. Clin Endocrinol (Oxf) 2006; 65:413-422


large number of healthy subjects: clinical utility in the follow-up of patients with

The introduction of the IDS-iSYS total IGF-1 assay may have far-reaching
consequences for diagnosis and treatment of GH deficiency. J Clin Endocrinol Metab
2015; 100:309-316

12. Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and

13. Blum WF, Breier BH. Radioimmunoassays for IGFs and IGFBPs. Growth Regul
1994; 4 Suppl 1:11-19

based on fractional polynomials and exponential transformation. J Royal Stat Soc:
Series A 1998; 161:79-101

15. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and
54:507-554


18. Landis JR, Koch GG. The measurement of observer agreement for categorical data.
Biometrics 1977; 33:159-174

Pituitary 2007; 10:121-128


**Table 1:** Characteristics of the tested IGF-I assays as provided by the manufacturers. These 6 assays are sandwich assays that use a couple of monoclonal antibodies directed against epitopes whose exact nature is not disclosed by the manufacturers. In all cases, IGFBPs are said to be removed by displacement of endogenous IGF-I by an excess of IGF-II (or analog) as initially proposed by Blum and Breier (13). The limit of quantification (LOQ) is the lowest amount of IGF-I that can be accurately quantified with an allowable error ≤ 20%. The limit of detection (LOD) is the IGF-I concentration corresponding to the 95th percentile value from a number of determinations of IGF-I concentration in free serum samples.

<table>
<thead>
<tr>
<th>Name of the assay</th>
<th>Manufacturer</th>
<th>Automated</th>
<th>Tracer</th>
<th>International standard against which the assay is calibrated</th>
<th>Intra-assay CV</th>
<th>Inter-assay CV</th>
<th>LOQ or LOD in ng/mL</th>
<th>Highest measurable value without dilution (ng/mL)</th>
<th>Reference adult population recruited by the manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>iSYS</td>
<td>IDS</td>
<td>Yes</td>
<td>Acridinium ester</td>
<td>WHO/NIBSC 02/254</td>
<td>2.9% at 22 ng/mL</td>
<td>5.4% at 22 ng/mL</td>
<td>8.8 (LOQ)</td>
<td>1200</td>
<td>6500 adults. Reference values provided according to the method of Cole and Green (12)</td>
</tr>
<tr>
<td>LIAISON XL</td>
<td>DiaSorin</td>
<td>Yes</td>
<td>Ioluminol</td>
<td>WHO/NIBSC 02/254</td>
<td>5.1% at 70 ng/mL</td>
<td>9.6% at 80 ng/mL</td>
<td>3 (LOD)</td>
<td>10 (LOQ)</td>
<td>1606 adults. Reference values provided by age according to the method of Royston and Wright (14)</td>
</tr>
<tr>
<td>IMMULITE 2000</td>
<td>Siemens</td>
<td>Yes</td>
<td>Alkaline phosphatase</td>
<td>WHO/NIBSC 1st IRR 87/518</td>
<td>3.9% at 77 ng/mL</td>
<td>7.7% at 77 ng/mL</td>
<td>20 (LOQ)</td>
<td>1600</td>
<td>1499 pediatric and adult samples from an apparently healthy population (no indication is given concerning the respective numbers of adult and children)</td>
</tr>
<tr>
<td>IGFI-RIACT</td>
<td>Cisbio</td>
<td>No</td>
<td>125I</td>
<td>WHO/NIBSC 1st IRR 87/518</td>
<td>3.8% at 49 ng/mL</td>
<td>3.8% at 39 ng/mL</td>
<td>1 (LOD)</td>
<td>900</td>
<td>693 adults 29-70 years</td>
</tr>
<tr>
<td>Mediagnost ELISA</td>
<td>MEDIA-GNOST</td>
<td>No</td>
<td>Peroxidase enzyme conjugate</td>
<td>WHO/NIBSC 02/254</td>
<td>5.7% at 138 ng/mL</td>
<td>6.1 % at 142 ng/mL</td>
<td>1.9 (LOD)</td>
<td>1050</td>
<td>Based on the data reported by Blum and Breier (13)</td>
</tr>
<tr>
<td>Mediagnost RIA</td>
<td>MEDIA-GNOST</td>
<td>No</td>
<td>125I</td>
<td>WHO/NIBSC 02/254</td>
<td>4.6% at 56 ng/mL</td>
<td>4.9% at 55 ng/mL</td>
<td>2.6 (LOD)</td>
<td>780</td>
<td>Based on the data reported by Blum and Breier (13)</td>
</tr>
</tbody>
</table>

The reference values for the different age ranges are the same as those used for the Mediagnost ELISA kit.
Table 2. Normative reference intervals (95% confidence interval: CI) of IGF-I measured by 6 assay methods according to age range and sex in a cohort of 899 healthy subjects

<table>
<thead>
<tr>
<th>Age range</th>
<th>N</th>
<th>isYS IGF-I (ng/mL)</th>
<th>IGAISON XL IGF-I (ng/mL)</th>
<th>IMMULITE 2000 IGF-I (ng/mL)</th>
<th>IGFI-RIACT IGF-I (ng/mL)</th>
<th>Medigast ELISA IGF-I (ng/mL)</th>
<th>Medigast RIA IGF-I (ng/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td></td>
<td>18-20 years:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>56</td>
<td>168-391</td>
<td>195-537</td>
<td>197-486</td>
<td>177-430</td>
<td>168-374</td>
<td></td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>132-313</td>
<td>152-430</td>
<td>155-389</td>
<td>144-355</td>
<td>135-308</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>108-265</td>
<td>118-348</td>
<td>127-329</td>
<td>115-295</td>
<td>112-265</td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>81-214</td>
<td>85-273</td>
<td>94-262</td>
<td>88-245</td>
<td>86-218</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>75-208</td>
<td>77-260</td>
<td>87-250</td>
<td>80-237</td>
<td>82-214</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>64-192</td>
<td>66-242</td>
<td>75-231</td>
<td>71-233</td>
<td>72-200</td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td>18-20 years:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>155-421</td>
<td>180-586</td>
<td>169-517</td>
<td>169-487</td>
<td>161-412</td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>144-383</td>
<td>166-541</td>
<td>159-476</td>
<td>156-446</td>
<td>149-379</td>
<td></td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>134-353</td>
<td>153-501</td>
<td>150-440</td>
<td>144-412</td>
<td>139-353</td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>126-330</td>
<td>142-467</td>
<td>134-385</td>
<td>131-332</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>113-294</td>
<td>121-403</td>
<td>126-356</td>
<td>118-341</td>
<td>118-298</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>97-253</td>
<td>109-296</td>
<td>107-297</td>
<td>100-296</td>
<td>103-258</td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>80-209</td>
<td>90-247</td>
<td>82-248</td>
<td>97-220</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>64-170</td>
<td>68-227</td>
<td>76-209</td>
<td>68-208</td>
<td>75-190</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>56-154</td>
<td>60-188</td>
<td>67-189</td>
<td>60-187</td>
<td>68-175</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Agreement of each IGF-1 assay method against each of the other, expressed as weighted kappa and percentages of observed agreement.

<table>
<thead>
<tr>
<th></th>
<th>LIAISON XL</th>
<th>iSYS</th>
<th>IMMULITE 2000</th>
<th>Medagnost ELISA</th>
<th>Medagnost RIA</th>
<th>IGFI-RIACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIAISON XL</td>
<td>-</td>
<td>0.49</td>
<td>0.50</td>
<td>0.47</td>
<td>0.38</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>94.86%</td>
<td>94.83%</td>
<td>94.95%</td>
<td>94.05%</td>
<td>95.22%</td>
</tr>
<tr>
<td>iSYS</td>
<td>0.49</td>
<td>-</td>
<td>0.64</td>
<td>0.61</td>
<td>0.70</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>94.86%</td>
<td></td>
<td>96.08%</td>
<td>96.11%</td>
<td>97.00%</td>
<td>96.46%</td>
</tr>
<tr>
<td>IMMULITE 2000</td>
<td>0.50</td>
<td>0.64</td>
<td>-</td>
<td>0.61</td>
<td>0.58</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>94.83%</td>
<td>96.08%</td>
<td></td>
<td>95.95%</td>
<td>95.73%</td>
<td>96.32%</td>
</tr>
<tr>
<td>Medagnost ELISA</td>
<td>0.47</td>
<td>0.61</td>
<td>0.61</td>
<td>-</td>
<td>0.59</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>94.95%</td>
<td>96.11%</td>
<td>95.95%</td>
<td></td>
<td>96.00%</td>
<td>95.66%</td>
</tr>
<tr>
<td>Medagnost RIA</td>
<td>0.38</td>
<td>0.70</td>
<td>0.58</td>
<td>0.59</td>
<td>-</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>94.05%</td>
<td>97.00%</td>
<td>95.73%</td>
<td>96.00%</td>
<td></td>
<td>95.22%</td>
</tr>
<tr>
<td>IGFI-RIACT</td>
<td>0.48</td>
<td>0.64</td>
<td>0.64</td>
<td>0.53</td>
<td>0.48</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>95.22%</td>
<td>96.46%</td>
<td>96.32%</td>
<td>95.66%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Concordance between VARIETE cohort reference intervals and reference intervals provided by each manufacturer, expressed as Kappa and percentages of observed agreement

<table>
<thead>
<tr>
<th>Weighted Kappa</th>
<th>LIAISON XL</th>
<th>iSYS</th>
<th>IMMULITE 2000</th>
<th>Medignost ELISA</th>
<th>Medignost RIA</th>
<th>IGFI - RIACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of agreement</td>
<td>83.28</td>
<td>93.36</td>
<td>86.97</td>
<td>93.55</td>
<td>94.77</td>
<td>88.21</td>
</tr>
</tbody>
</table>
Click here to access/download

Supplemental Material
Supplemental material VF .docx
Click here to access/download
Supplemental Material
Suppl Table 1.docx
Click here to access/download
Supplemental Material
Suppl Figure 1VF.docx