

## Reference Values for IGF-I Serum Concentrations: Comparison of Six 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\*

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.), Unité de Recherche Clinique (A.A.), and Service de Génétique Moléculaire, Pharmacogénétique et Hormonologie (S.B.T.), Assistance Publique-Hôpitaux de Paris, Hôpitaux Universitaires Paris-Sud, Hôpital de Bicêtre, 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 Laboratoire d'Hormonologie (C.M.), Centre Hospitalier Universitaire de Rennes, Centre d'Investigation Clinique Plurithématique, Inserm 1414, Hôpital Pontchaillou, Rennes, F29000, France; Service des Explorations Fonctionnelles (M-L.P., J.-C.S.), Assistance Publique-Hôpitaux de Paris, Hôpital Necker-Enfants Malades, Paris, F75015, France

**Context:** Measurement of IGF-I is essential for diagnosis and management of patients with disorders affecting the somatotrophic 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, we measured IGF-I in 911 healthy adults (18–90 years) with six immunoassays (iSYS, LIAISON XL, IMMULITE, IGFI RIAC, Mediagnost ELISA, and Mediagnost RIA). Pairwise concordance between assays was assessed with Bland-Altman plots for both IGF-1 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 (2.5–97.5 percentiles) given by the six IGF-I assays according to age group and sex. A formula for SDS calculation is provided. Although the lower limits of the reference intervals of the six assays were similar, the upper limits varied markedly. Pairwise concordances were moderate to good (0.38–0.70).

**Conclusion:** Despite being obtained in the same healthy population, the reference intervals of the six commercial IGF-1 assay kits showed noteworthy differences. Agreement between methods was moderate to good. (*J Clin Endocrinol Metab* 101: 3450–3458, 2016)

Growth hormone 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 somatotrophic axis, particularly GH excess (acromegaly) and GH deficiency. However, even if manufacturers follow the recommendations of the Con-

sensus 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 intermethod 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 because 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, whereas 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 body mass index (BMI), use of hormone replacement therapy by postmenopausal 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 because 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 interlaboratory 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 (2.5–97.5 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, IGFI RIAC, Mediagnost ELISA, and Mediagnost 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) as provided by the manufacturer, 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, nonrandomized 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 18–90 years) ([ClinicalTrials.gov Identifier: NCT01831648](https://clinicaltrials.gov/ct2/show/study/NCT01831648)). A total of 972 healthy subjects with BMI values between 19 and 28 kg/m<sup>2</sup> were recruited in 10 centers throughout France between 2010 and 2011. Our objective of including 1000 subjects was not achieved because of 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 ([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 hepatitis C virus 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-I 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-I levels, centile curves were constructed separately for each sex.

**Table 1.** Characteristics of the Tested IGF-I Assays as Provided by the Manufacturers

Assay Name	Manufacturer	Automated	Tracer	International Standard Against Which the Assay Calibrated	Intra-assay CV	Inter-assay CV	LOQ or LOD (ng/ml)	Highest Measurable Value Without Dilution (ng/ml)	Reference Adult Population Recruited by the Manufacturer
iSYS	IDS	Yes	Acridinium ester	WHO/NIBSC 02/254	2.9% at 22 ng/ml 1.9% at 163 ng/ml 4.2% at 304 ng/ml	5.4% at 22 ng/ml 3.9% at 163 ng/ml 7.2% at 304 ng/ml	8.8 (LOQ)	1200	6500 adults; reference values provided according to the method of Cole and Green (12)
LIAISON XL	DiaSorin	Yes	Isoluminol	WHO/NIBSC 02/254	5.1% at 70 ng/ml 3.5% at 183 ng/ml 3% at 589 ng/ml	9.6% at 80 ng/ml 7.1% at 187 ng/ml 5.6% at 317 ng/ml	3 (LOD) 10 (LOQ)	1500	1606 adults; reference values provided by age according to the method of Royston and Wright (14)
IMMULITE 2000	Siemens	Yes	Alkaline phosphatase	WHO/NIBSC First IRR 87/518	3.9% at 77 ng/ml 6.5% at 169 ng/ml 2.9% at 380 ng/ml 3.0% at 689 ng/ml 2.3% at 1053 ng/ml 2.4% at 1358 ng/ml	7.7% at 77 ng/ml 5.4% at 169 ng/ml 7.4% at 380 ng/ml 8.1% at 689 ng/ml 3.7% at 1053 ng/ml 4.7% at 1358 ng/ml	20 (LOQ)	1600	1499 pediatric and adult samples from an apparently healthy population (no indication is given concerning the respective numbers of adult and children)
IGFI-RIACT	Cisbio	No	<sup>125</sup> I	WHO/NIBSC First IRR 87/518	3.8% at 49 ng/ml 3.4% at 162 ng/ml 3.2% at 496 ng/ml	3.8% at 39 ng/ml 8.2% at 352 ng/ml 5.9% at 509 ng/ml	1 (LOD)	900	693 adults 29–70 y
Mediagnost	MEDIA	No	Peroxidase enzyme conjugate	WHO/NIBSC 02/254	5.7% at 138 ng/ml	6.1% at 142 ng/ml	1.9 (LOD)	1050	Based on the data reported by Blum and Breier (13)
ELISA	GNOST				5.1% at 141 ng/ml 6.6% at 145 ng/ml	6.8% at 174 ng/ml 2.2% at 494 ng/ml			
Mediagnost	MEDIA	No	<sup>125</sup> I	WHO/NIBSC 02/254	4.6% at 56 ng/ml	4.9% at 55 ng/ml	2.6 (LOD)	780	Based on the data reported by Blum and Breier (13)
RIA	GNOST				3.4% at 140 ng/ml 2.5% at 180 ng/ml	6.2% at 140 ng/ml 4.5% at 186 ng/ml			The reference values for the different age ranges are the same as those used for the Mediagnost ELISA kit

Abbreviations: CV, coefficient of variation; LOD, limit of detection; LOQ, limit of quantification; NICSC, National Institute for Biological Standards and Control; WHO, World Health Organization.

These six assays are sandwich assays that use monoclonal antibodies directed against epitopes, whose exact nature is not disclosed by the manufacturers. In all cases, IGF-BPs 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 LOQ is the lowest amount of IGF-I that can be accurately quantified with an allowable error <20%. The 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.

Age- and sex-specific centile curves were constructed for each assay by using the LMS (parameters L for skewness, M for median, and S for the coefficient of variation) method (12) implemented in the GAMLSS software package version 4.3–1 (15) of R software, version 3.1.2 (R Core Team, 2014; R: A language and environment for statistical computing; R Foundation for Statistical Computing; <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, M, and S were also computed for each age and sex class. SDS were calculated as  $z = [(IGF-I/M)^L - 1]/(L \times S)$ , 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 SDS formula.

Pairwise concordance between assays was assessed with scatter plots and Bland-Altman plots for both IGF-I raw values and SDS values, as well as with the percentage of observed agreement (total number of agreements divided by the total number of pa-

tients tested with both assays) and the linearly weighted Kappa coefficient for categorized IGF-I 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).

## Results

### 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 nine subjects were excluded because of missing information on pregnancy status or viral serology. The study population thus consisted of 911 subjects (470 males), comprising 101, 118, 99, 98, 103, 102, 108, 97, and 85 subjects in the 18–20, 21–23, 24–26, 27–29, 30–39, 40–

**Table 2.** Normative Reference Intervals (95% CI) of IGF-I Measured by Six Assay Methods According to Age Range and Sex in a Cohort of 899 Healthy Subjects

Age Range	N	iSYS IGF-I (ng/ml) 95% CI	LIAISON XL IGF-I (ng/ml) 95% CI	IMMULITE 2000 IGF-I (ng/ml) 95% CI	IGFI-RIACT IGF-I (ng/ml) 95% CI	Mediagnost ELISA IGF-I (ng/ml) 95% CI	Mediagnost RIA IGF-I (ng/ml) 95% CI
Males (y)							
18–20	56	168–391	186–453	195–537	197–486	177–430	168–374
21–23	61	147–346	168–411	171–477	173–430	159–388	150–337
24–26	53	132–313	153–377	152–430	155–389	144–355	135–308
27–29	49	122–292	142–351	138–396	143–363	133–331	126–289
30–39	56	108–265	124–310	118–348	127–329	115–295	112–265
40–49	51	91–233	106–271	98–301	107–286	98–261	97–237
50–59	54	81–214	97–252	85–273	94–262	88–245	86–218
60–69	49	75–208	92–245	77–260	87–250	80–237	82–214
70–89	34	64–192	80–220	66–242	75–231	71–233	72–200
Females (y)							
18–20	41	155–421	191–483	180–586	169–517	169–487	161–412
21–23	54	144–383	176–448	166–541	159–476	156–446	149–379
24–26	45	134–353	163–418	153–501	150–440	144–412	139–353
27–29	48	126–330	152–391	142–467	142–410	134–385	131–332
30–39	47	113–294	131–345	121–403	126–356	118–341	118–298
40–49	50	97–253	109–296	98–331	107–297	100–296	103–258
50–59	54	80–209	93–253	80–271	90–247	82–248	97–220
60–69	47	64–170	84–222	68–227	76–209	68–208	75–190
70–89	50	56–154	81–204	60–188	67–189	60–187	68–175

49, 50–59, 60–69, and 70–89 year age groups, respectively. Mean BMI was  $23.0 \pm 2.4$  kg/m<sup>2</sup>.

### 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/](http://ticemed_sa.upmc.fr/sd_score/)) or by using Apps (IGF1 SD\_score) downloadable for Android from Google Play and for iOS from Apple Store (free of charge) allows the obtaining of 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. Although the lower limits of the reference intervals (2.5th percentiles) were similar, the upper limits (97.5th 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, Mediagnost ELISA, and Mediagnost RIA.

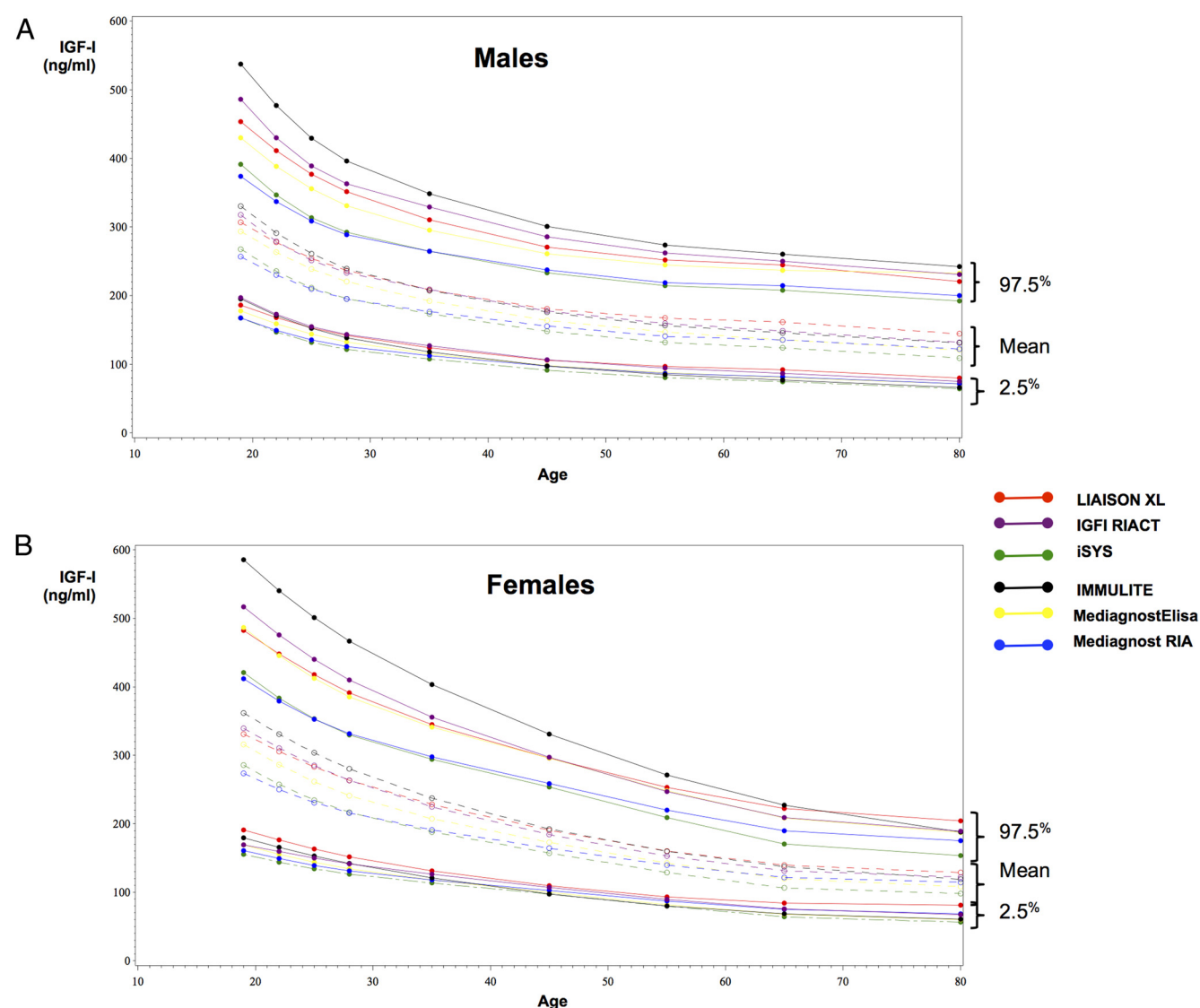
Two examples of interassay comparisons are shown in Figure 2. The results obtained with iSYS and Mediagnost RIA were in good overall agreement, with no significant bias as assessed by Bland-Altman plots (Figure 2, A–D). In contrast, the results obtained with LIAISON XL and Mediagnost RIA were not in good agreement (Figure 2, E–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 (0.38–0.70), although the percentages of observed agreement were quite high (94–97%).

Overall agreement was moderate as overall Kappa coefficient was 0.55. Both in men and women, global inter-assay comparison showed significant differences ( $P < .0001$ ) on raw values but not on SDS values ( $P = .26$  and  $P = .36$ , respectively).

Table 4 shows pairwise concordances between the reference intervals provided by the manufacturer and those





**Figure 1.** Reference intervals for (A) males and (B) females according to the age intervals of the six IGF-I immunoassays tested. Lower limits (2.5th percentile) and upper limits (97.5th percentile) of the normal range are drawn as full lines and means as dotted lines.

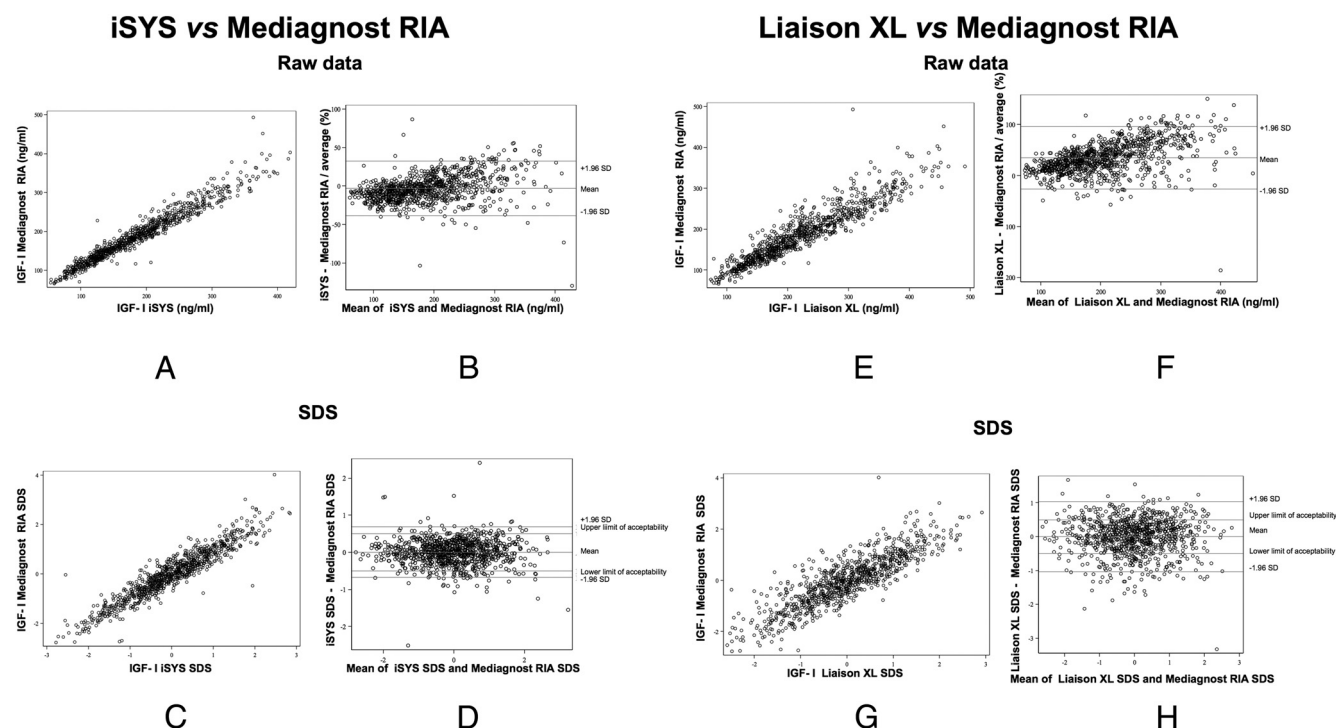
obtained in the VARIETE cohort, as assessed by the Kappa coefficient and the percentage agreement for each IGF-I assay. The concordances and percentages of observed agreement were generally poor.

## Discussion

We report reference intervals for IGF-I concentrations obtained with six immunoassays in the same population of nearly 900 French healthy subjects aged 18–90 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 composed 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 (2.5th percentile) were quite similar

with the six methods, but the upper limits (97.5th percentile) varied widely from one assay to another, in both men and women (Figure 1). Although the preanalytic 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, because 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 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 pa-



**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 IGF-I 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.

rameters (L, M, and S) 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/](http://ticemed_sa.upmc.fr/sd_score/)) or by using Apps (IGF-1SD\_score) downloadable for Android from Google Play and for iOS from Apple Store (free of charge), which allows the determination of 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

**Table 3.** Agreement of Each IGF-I Assay Method Against Each of the Others, Expressed as Weighted Kappa and Percentages of Observed Agreement

	LIAISON XL	iSYS	IMMULITE 2000	Mediagnost ELISA	Mediagnost RIA	IGF-I RIAC
LIAISON XL	—	0.49 94.86%	0.50 94.83%	0.47 94.95%	0.38 94.05%	0.48 95.22%
iSYS	0.49 94.86%	—	0.64 96.08%	0.61 96.11%	0.70 97.00%	0.64 96.46%
IMMULITE 2000	0.50 94.83%	0.64 96.08%	—	0.61 95.95%	0.58 95.73%	0.64 96.32%
Mediagnost ELISA	0.47 94.95%	0.61 96.11%	0.61 95.95%	—	0.59 96.00%	0.53 95.66%
Mediagnost RIA	0.38 94.05%	0.70 97.00%	0.58 95.73%	0.59 96.00%	—	0.48 95.22%
IGF-I RIAC	0.48 95.22%	0.64 96.46%	0.64 96.32%	0.53 95.66%	0.48 95.22%	—

**Table 4.** Concordance Between IGF-I VARIETE Cohort Reference Intervals and IGF-I Reference Intervals Provided by Each Manufacturer, Expressed as Kappa and Percentages of Observed Agreement

	LIAISON XL	iSYS	IMMULITE 2000	Mediagnost ELISA	Mediagnost RIA	IGFI -RIACT
Weighted Kappa	0.19	0.35	0.38	0.18	0.17	0.22
% of agreement	83.28	93.36	86.97	93.55	94.77	88.21

that all assays be calibrated against this standard, and advocated precise preanalytical 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 1000 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 workup to exclude those with disorders capable of influencing IGF-I levels or their measurement. These very strict inclusion and exclusion criteria allow us to define a population as “healthy” as possible; however, this implies that these normative data will not be strictly applicable to patients with BMI higher than 28 kg/m<sup>2</sup> 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 fifth 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-1 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).

Interassay 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 intercentile 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 interlaboratory 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 international reference reagent 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 years (Table 2, Figure 1). However, the reference ranges of two differently calibrated kits may be either similar (eg, LIAISON XL and IGF1 RIACT in men) or significantly different (eg, 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 (eg, 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, whereas the other two assays more efficiently remove the impact of binding proteins.

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 liquid chromatography (LC) tandem mass spectrometry (MS) method may prove to be a valid alternative and is now being used to assess interlaboratory 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, the 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, tandem LC-MS 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 (eg, 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. Interassay 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 authors thank Dr. Hélène Agostini and Prof. Bruno Falissard for their helpful comments. The authors also thank Cisbio International, DiaSorin, IDS, Mediagnost, and Siemens for the kind donation of IGF-I kits.

Address all correspondence and requests for reprints to: Philippe Chanson, MD, Service d'Endocrinologie et des Maladies de la Reproduction, Hôpital de Bicêtre, 78 rue du Général Leclerc, 94275 Le Kremlin-Bicêtre, France. E-mail: [philippe.chanson@bct.aphp.fr](mailto:philippe.chanson@bct.aphp.fr).

This study was supported by a grant from Programme Hospitalier de Recherche Clinique, French Ministry of Health, no. P081216/IDRCB 2009-A00892-55, to Drs Chanson and Soubrier, and by a grant from Fond National Suisse, P1GEP3 155694, to Dr. Mavromati.

\*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](#).

Trial Registration: [ClinicalTrials.gov](https://clinicaltrials.gov): NCT01831648.

Disclosure Summary: The authors have nothing to disclose.

## References

1. Clemmons DR. Consensus statement on the standardization and evaluation of growth hormone and insulin-like growth factor assays. *Clin Chem*. 2011;57:555–559.
2. Pokrajac A, Wark G, Ellis AR, Wear J, Wieringa GE, 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.
3. Quarmby V, Quan C, Ling V, Compton P, Canova-Davis E. How much insulin-like growth factor I (IGF-I) circulates? Impact of standardization on IGF-I assay accuracy. *J Clin Endocrinol Metab*. 1998;83:1211–1216.
4. Frystyk J, Freda P, Clemmons DR. The current status of IGF-I assays—a 2009 update. *Growth Horm IGF Res*. 2010;20:8–18.
5. Brabant G, Wallaschofski H. Normal levels of serum IGF-I: determinants and validity of current reference ranges. *Pituitary*. 2007;10:129–133.



6. Juul A. Serum levels of insulin-like growth factor I and its binding proteins in health and disease. *Growth Horm IGF Res.* 2003;13:113–170.
7. Leung KC, Johannsson G, Leong GM, Ho KK. Estrogen regulation of growth hormone action. *Endocr Rev.* 2004;25:693–721.
8. Meinhardt UJ, Ho KK. Modulation of growth hormone action by sex steroids. *Clin Endocrinol (Oxf).* 2006;65:413–422.
9. Clemmons DR. Value of insulin-like growth factor system markers in the assessment of growth hormone status. *Endocrinol Metab Clin North Am.* 2007;36:109–129.
10. Massart C, Poirier JY, Jard C, Pouchard M, Vigier MP. Determination of serum insulin-like growth factor-I reference values for the immunometric Cisbio method on a large number of healthy subjects: clinical utility in the follow-up of patients with treated acromegaly. *Clin Chim Acta.* 2007;381:176–178.
11. Varewijck AJ, Lamberts SW, van der Lely AJ, Neggers SJ, Hofland LJ, Janssen JA. 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 penalized likelihood. *Stat Med.* 1992;11:1305–1319.
13. Blum WF, Breier BH. Radioimmunoassays for IGFs and IGF-BPs. *Growth Regul.* 1994;4 Suppl 1:11–19.
14. Royston P, Wright EM. A method for estimating age-specific reference intervals 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 shape. *J Royal Stat Soc Series C (Applied Statistics).* 2005;54:507–554.
16. Fleiss J-L, Levin B, Paik MC. *Statistical Methods for Rates and Proportions.* 3rd ed. New York: Wiley.
17. Cicchetti DV, Allison T. A new procedure for assessing reliability of scoring EEG sleep recordings. *Am J EEG Technol.* 1971;11:101–109.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
19. Clemmons DR. IGF-I assays: current assay methodologies and their limitations. *Pituitary.* 2007;10:121–128.
20. Junnila RK, Strasburger CJ, Bidlingmaier M. Pitfalls of insulin-like growth factor-I and growth hormone assays. *Endocrinol Metab Clin North Am.* 2015;44:27–34.
21. Brabant G, von zur Muhlen A, Wuster C, et al. Serum insulin-like growth factor I reference values for an automated chemiluminescence immunoassay system: results from a multicenter study. *Horm Res.* 2003;60:53–60.
22. Burns C, Rigby P, Moore M, Rafferty B. The First International Standard For Insulin-like Growth Factor-1 (IGF-1) for immunoassay: preparation and calibration in an international collaborative study. *Growth Horm IGF Res.* 2009;19:457–462.
23. Bidlingmaier M, Friedrich N, Emeny RT, et al. Reference intervals for insulin-like growth factor-1 (IGF-I) from birth to senescence: results from a multicenter study using a new automated chemiluminescence IGF-I immunoassay conforming to recent international recommendations. *J Clin Endocrinol Metab.* 2014;99:1712–1721.
24. Jernstrom H, Deal C, Wilkin F, et al. Genetic and nongenetic factors associated with variation of plasma levels of insulin-like growth factor-I and insulin-like growth factor-binding protein-3 in healthy premenopausal women. *Cancer Epidemiol Biomarkers Prev.* 2001;10:377–384.
25. Elkazaz AY, Salama K. The effect of oral contraceptive different patterns of use on circulating IGF-1 and bone mineral density in healthy premenopausal women. *Endocrine.* 2015;48:272–278.
26. Blackmore KM, Wong J, Knight JA. A cross-sectional study of different patterns of oral contraceptive use among premenopausal women and circulating IGF-1: implications for disease risk. *BMC Womens Health.* 2011;11:15.
27. Balogh A, Kauf E, Vollandt R, et al. Effects of two oral contraceptives on plasma levels of insulin-like growth factor I (IGF-I) and growth hormone (hGH). *Contraception.* 2000;62:259–269.
28. Massart C, Poirier JY. Serum insulin-like growth factor-I measurement in the follow-up of treated acromegaly: comparison of four immunoassays. *Clin Chim Acta.* 2006;373:176–179.
29. Massart C, Poirier JY. Determination of serum insulin-like growth factor-I reference values for the automated chemiluminescent Liaison(R) assay. Clinical utility in the follow-up of patients with treated acromegaly. *Clin Chim Acta.* 2011;412:398–399.
30. Krebs A, Wallaschofski H, Spilcke-Liss E, et al. Five commercially available insulin-like growth factor I (IGF-I) assays in comparison to the former Nichols Advantage IGF-I in a growth hormone treated population. *Clin Chem Lab Med.* 2008;46:1776–1783.
31. Cowan DA, Bartlett C. Laboratory issues in the implementation of the marker method. *Growth Horm IGF Res.* 2009;19:357–360.
32. Milani D, Carmichael JD, Welkowitz J, 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;89:2271–2274.
33. Nguyen TV, Nelson AE, Howe CJ, et al. Within-subject variability and analytic imprecision of insulinlike growth factor axis and collagen markers: implications for clinical diagnosis and doping tests. *Clin Chem.* 2008;54:1268–1276.
34. Cox HD, Lopes F, Woldemariam GA, et al. Interlaboratory agreement of insulin-like growth factor 1 concentrations measured by mass spectrometry. *Clin Chem.* 2014;60:541–548.
35. Kay R, Halsall DJ, Annamalai AK, et al. A novel mass spectrometry-based method for determining insulin-like growth factor 1: assessment in a cohort of subjects with newly diagnosed acromegaly. *Clin Endocrinol (Oxf).* 2013;78:424–430.
36. Bystrom C, Sheng S, Zhang K, et al. Clinical utility of insulin-like growth factor 1 and 2; determination by high resolution mass spectrometry. *PLoS One.* 2012;7:e43457.
37. Hoofnagle AN, Whiteaker JR, Carr SA, et al. Recommendations for the generation, quantification, storage, and handling of peptides used for mass spectrometry-based assays. *Clin Chem.* 2016;62:48–69.
38. Hines J, Milosevic D, Ketha H, et al. Detection of IGF-1 protein variants by use of LC-MS with high-resolution accurate mass in routine clinical analysis. *Clin Chem.* 2015;61:990–991.
39. Ketha H, Singh RJ. Clinical assays for quantitation of insulin-like growth-factor-1 (IGF1). *Methods.* 2015;81:93–98.