Clinical Biochemistry 114 (2023) 95–102

Contents lists available at ScienceDirect

Clinical Biochemistry

journal homepage: www.elsevier.com/locate/clinbiochem

The impact of different calibration matrices on the determination of insulin-like growth factor 1 by high-resolution-LC-MS in acromegalic and growth hormone deficient patients

Sebastian Simstich [a,1], Thomas Züllig [b,i,1], Federica D'Aurizio [c], Alessia Biasotto [d], Annamaria Colao [e,f], Andrea M. Isidori [g], Andrea Lenzi [h], Günter Fauler [a], Harald C. Köfeler [b], Francesco Curcio [c,d,1], Markus Herrmann [a,*,1]

[a] Clinical Institute of Medical and Chemical Laboratory Diagnosis, Medical University of Graz, Austria
[b] Core Facility Mass Spectrometry, Medical University of Graz, Austria
[c] Department of Laboratory Medicine, Institute of Clinical Pathology, Academic Hospital of Udine, Italy
[d] Department of Medicine, University of Udine, Italy
[e] Department of Clinical Medicine and Surgery, Section of Endocrinology, University Federico II, Naples, Italy
[f] UNESCO Chair for Health Education and Sustainable Development, University Federico II, Naples, Italy
[g] Department of Experimental Medicine, Policlinico Umberto I Hospital, Sapienza University of Rome, Italy
[h] Department of Experimental Medicine, Sapienza University of Rome, Italy
[i] Institute of Molecular Biosciences, University of Graz, Austria

ARTICLE INFO

Keywords:
Insulin-like Growth Factor 1 (IGF-1)
Liquid Chromatography High-Resolution Mass Spectrometry (LC-HRMS)
Online Solid Phase Extraction (OSPE)
Quantification
Hormone
Peptide
Serum

ABSTRACT

Objectives: Calibration is an important source of variability in liquid chromatography mass spectrometry (LC-MS) methods for insulin-like growth factor 1 (IGF-1). This study investigated the impact of different calibrator matrices on IGF-1 measurements by LC-MS. Moreover, the comparability of immunoassays and LC-MS was assessed.

Design & Methods: Calibrators from 12.5 to 2009 ng/ml were prepared by spiking WHO international Standard (ID 02/254 NIBSC, UK) into the following matrices: native human plasma, fresh charcoal-treated human plasma (FCTHP), old charcoal-treated human plasma, deionized water, bovine serum albumin (BSA), and rat plasma (RP). A validated in-house LC-MS method was calibrated repeatedly with these calibrators. Then, serum samples from 197 growth hormone excess and deficiency patients were analysed with each calibration.

Results: The seven calibration curves had different slopes leading to markedly different patient results. The largest differences in IGF-1 concentration from the median (interquartile range) was observed with the calibrator in water and the calibrator in RP (336.4 [279.6–417.0] vs. 112.5 [71.2–171.2], $p < 0.001$). The smallest difference was observed with calibrators in FCTHP and BSA (141.8 [102.0–198.5] vs. 127.9 [86.9–186.0], $p < 0.049$). Compared to LC-MS with calibrators in FCTHP, immunoassays showed relevant proportional bias (range: −43% to −68%), constant bias (range: 22.84 to 57.29 ng/ml) and pronounced scatter. Comparing the immunoassays with each other revealed proportional bias of up to 24%.

Conclusions: The calibrator matrix is critical for the measurement of IGF-1 by LC-MS. Regardless of the calibrator matrix, LC-MS shows poor agreement with immunoassays. Also, the agreement between different immunoassays is variable.

* Corresponding authors at: Clinical Institute of Medical and Chemical Laboratory Diagnosis, Medical University of Graz, Auenbruggerplatz 15/1, 8036 Graz, Austria.
E-mail address: markus.herrmann@medunigraz.at (M. Herrmann).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.clinbiochem.2023.02.008
Received 4 October 2022; Received in revised form 1 February 2023; Accepted 20 February 2023
Available online 26 February 2023

1. Introduction

Serum IGF-1, a 70-amino acid peptide hormone, is a critical analyte for the diagnosis and management of patients with growth hormone excess (acromegaly) and deficiency [1]. It is a key mediator of the anabolic and growth-promoting effects of hGH. In contrast to the pulsatile secretion of hGH by the pituitary gland, IGF-1 is released more steadily resulting in much less volatile serum concentrations [2]. Although the hormone can be produced in many tissues, the liver is the primary source of circulating IGF-1 [3].

In serum, IGF-1 is bound to one of six structurally and evolutionarily related IGF binding proteins (IGFBP-1 through IGFBP-6) and an acid labile subunit (ALS). The majority of the circulating ternary complexes is formed with IGFBP-3. In human plasma, only less than five percent of the IGF-1 molecules are available in its free and unbound form with a mass of 7.649 kDa [4]. Therefore, IGFBP-3 constitutes a possible adjunct to IGF-1 and hGH in the diagnosis and follow-up of acromegaly. In addition to serum analyses of IGF-1 for the diagnosis and management of acromegaly and growth hormone deficiency, IGF-1 can also be measured in alternative matrices, such as dried blood spots [5], or urine samples [6], to detect abuse in athletes, for example.

In clinical practice, quantitative measurement of IGF-1 is typically performed by immunoassays, which are known for their inter-laboratory variation [7,8]. Therefore, results obtained by different assays are difficult to compare and can lead to different clinical interpretations and inappropriate patient management [7-12]. Variable IGF-1 results obtained by immunoassays are mainly caused by different specificities of the capture antibodies, lot-to-lot variations, interferences from binding proteins, and naturally occurring IGF-1 variants with variable biological activity [8,13–15].

Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) is a technology that has the potential to overcome the above-mentioned limitations of IGF-1 immunoassays [16,17]. Similar to immunoassays, also LC-MS/MS methods require pre-analytical sample acidification to dissociate IGF-1 from binding proteins and to precipitate abundant proteins. Then, two different approaches can be used to quantitate IGF-1 by LC-MS/MS. The top-down approach measures intact IGF-1 [6,18-20], whereas a bottom-up approach analyzes IGF-1-specific fragments after digestion with trypsin [21]. In principle, both approaches can deliver comparable results [22]. In addition to the two strategies mentioned before, also immune purification methods have been published, where IGF-1 is isolated from serum by the use of antibody-coated pipette tips [23,24]. Despite superior sensitivity and specificity, also LC-MS/MS methods for IGF-1 can vary significantly [25]. In a recent comparison between five laboratories that measured the same two IGF-1 peptides 1–21 and 22–36 with 5 different instruments, Cox et al showed that the primary source of variability was the method of calibration [25].

The strong binding of IGF-1 to IGFBPs complicates the preparation of valid calibrators. Previous studies used calibrators that were prepared in serum from rats [20,25], mice [26], or a solution of ovalbumin in phosphate buffered saline [27]. Considering that calibration seems to be a major source of variability, harmonization of LC-MS/MS assay calibration is desirable. However, a head-to-head comparison of different artificial calibration matrices has not been performed yet. Therefore, the present study aimed to investigate the impact of different artificial matrices for calibrating a high-resolution LC-MS method for IGF-1. After validating this method with the best performing calibrators, a set of samples from acromegaly and GH-deficient patients was analyzed by LC-MS and compared to different commercial immunoassays.

2. Material and methods

2.1. Implementation and validation of a LC-MS-method

For measurement of IGF-1 an in-house LC-MS method has been established and validated. This method is based on a previously published method from Bystrom et al. [21], with the main difference that we used an Orbitrap instrument instead of a quadrupole time-of-flight (qTOF) mass spectrometer. Also, the internal standard of our method was a ^{15}N labelled IGF-1 (^{15}N labeled IGF-1, 8.0 µg/ml, CYT-128, PROSPEC, Rehovot, Israel) instead of an in-house oxidized rat IGF-1.

For online sample purification and subsequent chromatography, a Vanquish Flex UHPLC (Thermo Fisher Scientific, Waltham, MA, USA) was used. A ZORBAX Eclipse AAA C18 (4.6 × 12.5 mm, 5 µm) was employed for online solid phase extraction (OSPE), whereas analytical separation was performed with a Phenomenex Kinetex core–shell EVO C18 (150 × 2.1 mm, 2.6 µm) column. The following eluents were used: A: 0.2 % formic acid (FA) in deionized water and B: 0.2 % FA in acetonitrile/H_2O (80/20, v/v).

A Vanquish quaternary pump (VF-P20_A, Thermo Fisher Scientific, Waltham, MA, USA) and a Vanquish binary pump (VF-P10_A, Thermo Fisher Scientific) in combination with two 6 port-valves were used for sample clean-up and subsequent analytical separation. Flow rates and gradients are shown in Table 1. After each analysis, a blank injection was performed to reduce carry-over.

10 µl of sample, including IS, were injected into the instrument and loaded on the OSPE column for clean-up. Then, the purified sample extracts were transferred to the analytical column for separation, and subsequently introduced into the ion source of the Q-Exactive Focus (Thermo Fisher Scientific, Waltham, MA, USA) mass spectrometer for analysis. The system was operated in data dependent positive full-scan mode using the following instrument parameters:

Aux gas temperature was set at 400 °C, Spray voltage 4 kV, Sheath gas flow rate 48, Aux gas flow rate 15, Capillary temperature 320 °C, Probe heater 350 °C, and S-Lens RF level 60. Resolution was 70,000 (m/z 200) and scan range was from $m/z = 600$ to $m/z = 1800$, AGC target was 2e5, and maximum IT was 200 ms.

Following liquid chromatography, samples were introduced into the ion source of the mass spectrometer, where ionization was performed with electro-spray-ionization (ESI) in positive ion mode. From the total ion current, the three most abundant isotopic ions of the IGF-1 were selected for quantitation. These peaks were derived from the 7 fold charged IGF-1 ion ($[M + 7H]^{+7}$) and have the following m/z-values: 1093.521, 1093.665, 1093.807. In addition, the ^{15}N-labeled ion of the IS with an $m/z = 1106.769$ was used (Fig. 1).

Raw data were exported and processed using the Lipid Data Analyzer 2.6.3_9 software [28,29]. All detected chromatographic peaks were manually controlled and reintegrated, if necessary. Quantitation was performed within a concentration range of 12.5 ng/ml to 2009 ng/ml in any matrix.

The validation of the method included the determination of the following performance indices: limit of detection (LOD), limit of

Table 1
Gradient elution of the IGF-1 method with the valve switch procedure.

Quaternary pump		Binary pump		Time	Valve 1	Valve 2	Steps
ml/min	% A	ml/min	% A	min	position	position	
0.5	80	0.4	80	0	01.Feb	01.Feb	OSPE Loading
0.5	80	0.4	80	1	01.Jun	01.Feb	start elution to analytical column
0.4	80	0.4	80	1.01	01.Jun	01.Feb	
0.4	80	0.4	80	3.5	01.Jun	01.Jun	analytical column to waste
0.4	0	0.4	0	4	01.Jun	01.Jun	
0.4	0	0.4	0	5	01.Jun	01.Jun	
0.4	80	0.4	80	5.01	01.Jun	01.Jun	
0.4	80	0.4	80	5.5	01.Feb	01.Jun	clean up OSPE
0.4	80	0.4	80	5.6	01.Feb	01.Jun	
0.4	80	0.4	80	5.9	01.Jun	01.Jun	
0.4	80	0.4	80	6	01.Jun	01.Jun	

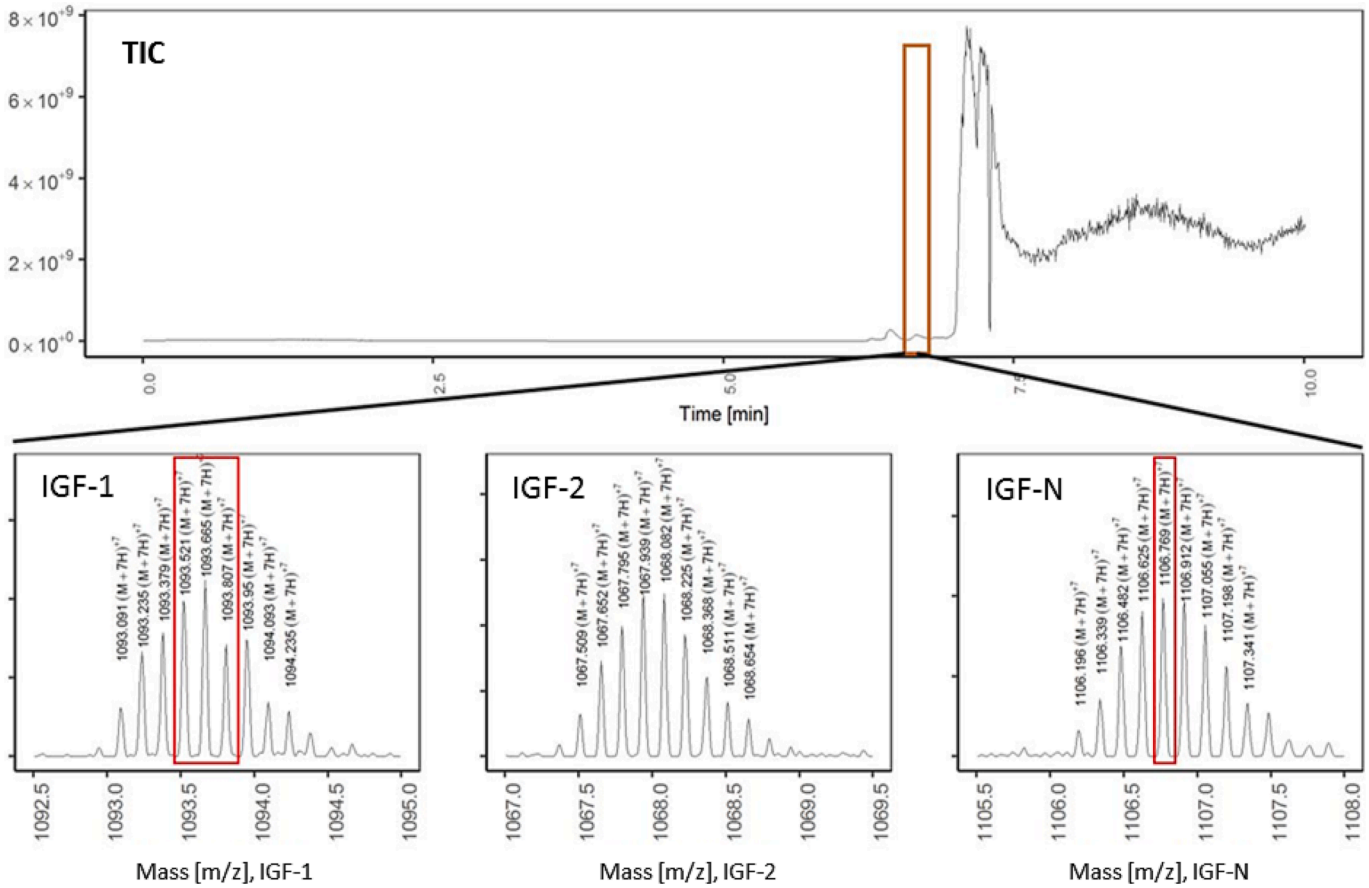

Fig. 1. Total ion chromatogram (TIC) of extracted serum with internal standard, shown as extracted ion chromatograms (XIC) of IGF-1, IGF-2 and IGF-N (IS), the red box highlights the ions used for quantitation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

quantification (LOQ), autosampler stability and imprecision. For calculation of LOD and LOQ a linear regression model was used by determination of standard deviations of five independent calibration sets. Autosampler stability was assessed using two sets of calibrators (eight samples each) with concentrations from 12.5 ng/ml to 2009 ng/ml. Samples were quantified directly after sample preparation and after storage in the autosampler for 24 h and for 48 h, respectively, at 8 °C. Imprecision was determined by repeat measurements of human charcoal stripped plasma samples, spiked with standard concentrations of 62.5 ng/ml, 500 ng/ml, and 1000 ng/ml. Each sample was processed five times at different days. Means, standard deviations and CVs were calculated.

2.2. Preparation of serum and calibration samples

In order to dissociate IGF-1 from binding proteins and to reduce the protein content in the samples, serum has to be acidified and precipitated before sample preparation starts. For this purpose, 100 µl of serum and 10 µl of internal standard (8 µg/ml, ^{15}N labeled IGF-1) were added into a 0.5 ml LoBind Tube (0060108.094, Hamburg, DE-HH). Precipitation was performed by adding 400 µl acidified ethanol (87.5 % EtOH, 12.5 % 1 M HCl). All solvents and chemicals were purchased from Sigma Aldrich, Vienna, unless otherwise specified. After 30 min of incubation at room temperature, samples were spun at 13,000 g for 10 min. Subsequently, 350 µl of the supernatant were transferred to a fresh LoBind tube and mixed with TRIS-Base (60 µl, 1.5 mol/l). After 30 min of incubation at −20 °C, samples were centrifuged again at 13,000 g for 10 min and the supernatant was transferred to a 200 µl glass vial and placed in the autosampler at 8 °C.

Preparation of calibration samples: The WHO international Standard for IGF-1 from the National Institute for Biological Standards and Control (ID 02/254 NIBSC, UK) was used for the preparation of eight different calibrators with concentrations ranging from 12.5 ng/ml to 2009 ng/ml. Similar to patient samples, also calibrators were processed together with 10 µl of IS. For the comparison of calibration matrices, calibrators were prepared in the following matrices: native human plasma, old charcoal-treated human plasma, fresh charcoal-treated human plasma, pure deionized water, bovine serum albumin (BSA) solution, and rat plasma.

Preparation of charcoal-stripped IGF-1-free human plasma: A pool of human plasma was prepared. Charcoal-stripping was performed by adding 6 g of charcoal to 100 ml of plasma and stirring for 90 min at room temperature. After centrifugation for 60 min with 5,000 g, the clear supernatant was transferred and charcoal treatment was repeated.

A new calibration curve in the desired matrix was generated together with every batch of samples.

2.3. Method comparison

In the next step, the comparability of the in-house LC-MS method with four commercial IGF-1 immunoassays was assessed in a set of 197 serum samples from patients with acromegaly and GH-deficiency. These patients were recruited at different endocrine units in Italy (Milan, Rome, Naples, Turin) after approval of the institutional ethics committee. All immunoassays were performed at the Azienda Sanitaria Universitaria Friuli Centrale, Presidio Ospedaliero S. Maria della Misericordia, Udine, Italy. The IGF-1 immunoassays employed were from Siemens Healthineers (Erlangen, Germany), DiaSorin (Saluggia,

Italy), Roche Diagnostics (Mannheim, Germany) and Immunodiagnostic Systems ([IDS], Bolden, UK), and were run in accordance with the manufacturer's instructions on the following instruments: Siemens Immulite 2000 XPi, DiaSorin LIAISON XL, Roche Cobas e 801 and IDS-iSYS. Details of assay design and analytical performance are provided in Table 2. In order to exclude between-assay variability, the 197 serum samples were analysed with each of the four immunoassays in a single analytical session. All assays were validated by the laboratory and reached an analytical performance that was similar to the manufacturer's specifications.

After completion of immunoassay testing, the samples were transferred on dry ice to the Clinical Institute of Medical and Chemical Laboratory Diagnostics at the Medical University of Graz (Austria), where LC-MS testing was performed. For this comparison, the LC-MS method was calibrated with calibrators prepared in the best performing matrix (fresh charcoal-stripped human plasma, see below).

2.4. Statistics

LOD and LOQ were calculated as $3S_a\ b^{-1}$ and as $10S_a\ b^{-1}$, respectively, from the linear regression line: $y = -2.99 + 0.0243x$, ($R = 0.9882$).

Normalized peak area ratios of two sets of calibrators (eight samples each) were compared, by measuring directly after sample preparation and after storage in the autosampler (at 8 °C) for 24 h and for 48 h, respectively. A Mann-Whitney *U* test was used to show differences between direct sample analysis and autosampler storage.

The IGF-1-results obtained in the 197 patient samples by LC-MS and immunoassay were compared by Deming regression analyses and Bland-Altman plots.

3. Results

3.1. Validation of LC-MS method

The in-house LC-MS method produced chromatograms of good quality with well separated peaks for the target ions ($m/z = 1093.521$, 1093.665, 1093.807, respectively) and the IS ($m/z = 1106.769$, Fig. 1). As shown in Fig. 2, the use of an IS had a substantial impact on the slope of the calibration curve and the precision of repeat measurements. The method showed a linear behavior between 12.5 ng/mL and 2009 ng/mL in a charcoal-treated plasma matrix, and this range was used for comparison with the other calibration variants. LOD and LOQ were 1.72 ng/ml and 5.2 ng/ml, respectively. Imprecision was ≤7.3% at all tested concentrations (Table 3). When kept in the autosampler, the samples did not show degradation after 48 h.

3.2. Impact of different calibration matrices on IGF-1

Preparation of the calibrators in various matrices resulted in markedly different slopes of the calibration curves (Fig. 2). The highest and the lowest slope were observed in fresh and old charcoal-treated human plasma, respectively. Introducing an IS in the sample preparation improved the situation, but substantial differences remained. However, calibrators in matrices that are closely related to native human plasma, but free of human IGF-1, produced calibration curves with reasonable agreement. In contrast, calibrators prepared in water and native human plasma yielded markedly different calibration curves. These differences impacted the results of patient samples (Fig. 2). The greatest difference between the median (IQR) IGF-1 concentration in the 197 patient samples was seen when calibrators were prepared in water and rat plasma (336.4 [279.6–417.0] vs. 112.5 [71.2–171.2], $p < 0.001$). In contrast, the smallest difference between the patient results was observed when calibration was performed with calibrators in fresh charcoal treated human plasma and BSA (141.8 [102.0–198.5] vs. 127.9 [86.9–186.0], $p < 0.049$). Calibrators in old charcoal treated human plasma yielded significantly higher IGF-1 results when compared to calibrators in fresh charcoal treated human plasma (207.0 [132.4–313.0] vs. 141.8 [102.0–198.5], $p < 0.001$). Also, the preparation of calibrators in native human plasma caused significantly lower IGF-1 results than fresh charcoal treated human plasma based calibrators.

3.3. Comparison of patient samples in different methods

In the next step, we compared the IGF-1 results of the 197 patients samples obtained with four different commercial immunoassays and our LC-MS method calibrated with fresh charcoal treated human plasma based calibrators. When compared to LC-MS, all immunoassays showed great negative proportional bias, ranging from 53 to 72% (Fig. 3). In addition, the assays from DiaSorin, IDS and Siemens had significant constant bias between 46.24 and 51.76 ng/ml. Visual examination of the regression analyses and the Bland-Altman plots showed pronounced scatter of the immunoassay results around the line of identity, when compared to LC-MS.

The IGF-1 results obtained with the four immunoassays also exhibited substantial variability (Fig. 4). Despite pronounced scatter at high concentrations, the assays from DiaSorin, IDS and Roche did not have significant proportional or constant bias against each other in the Deming regression analysis. In contrast, the Siemens Immulite assay exhibited significant proportional bias between −18% and −24% against the other immunoassays.

4. Discussion

LC-MS allows the measurement of IGF-1 in human serum with high specificity and sensitivity. However, the present results demonstrate that the calibration matrix is critical for the IGF-1 concentrations obtained with this technique. IGF-1 results in patient samples agreed best, when calibrators were prepared in IGF-1-free matrices that are closely related to human plasma, such as fresh charcoal treated human plasma or BSA. The immunoassays showed very different performances, with results not being comparable to those of LC-MS in terms of systematic and proportional bias, as shown by Deeming regression. When compared amongst themselves, the immunoassays from DiaSorin, IDS and Roche agreed reasonably well, whereas the SIEMENS test had substantial negative proportional bias against all other methods. With all assays, the scatter around the regression line increased with higher concentrations suggesting limited accuracy at this concentration range.

Our results support previous studies suggesting that calibration is the primary source of variability when measuring IGF-1 by LC-MS. In a comparison study of IGF-1 measurements by LC-MS in five different

Table 2
Relevant key features of four commercial immunoassays.

	sample volume	measurement time	antibody	LOD	LOQ	analytical range	imprecision
	µl	min	min	ng/ml	ng/ml	ng/ml	CV %
Immulite, Siemens	20	60	mono-poly-clonal	13.3	24.9	15–1000	≤7.6
LIAISON, DiaSorin	20		2x monoclonal	3	10	10–1000	≤8.5
Elecsys, Roche	6	18	2x monoclonal	7	15	7–1600	≤5
IDS-iSYS, Immunodiagnosticsystems	10		2x monoclonal	4.4	8.8	10–1200	≤7.2

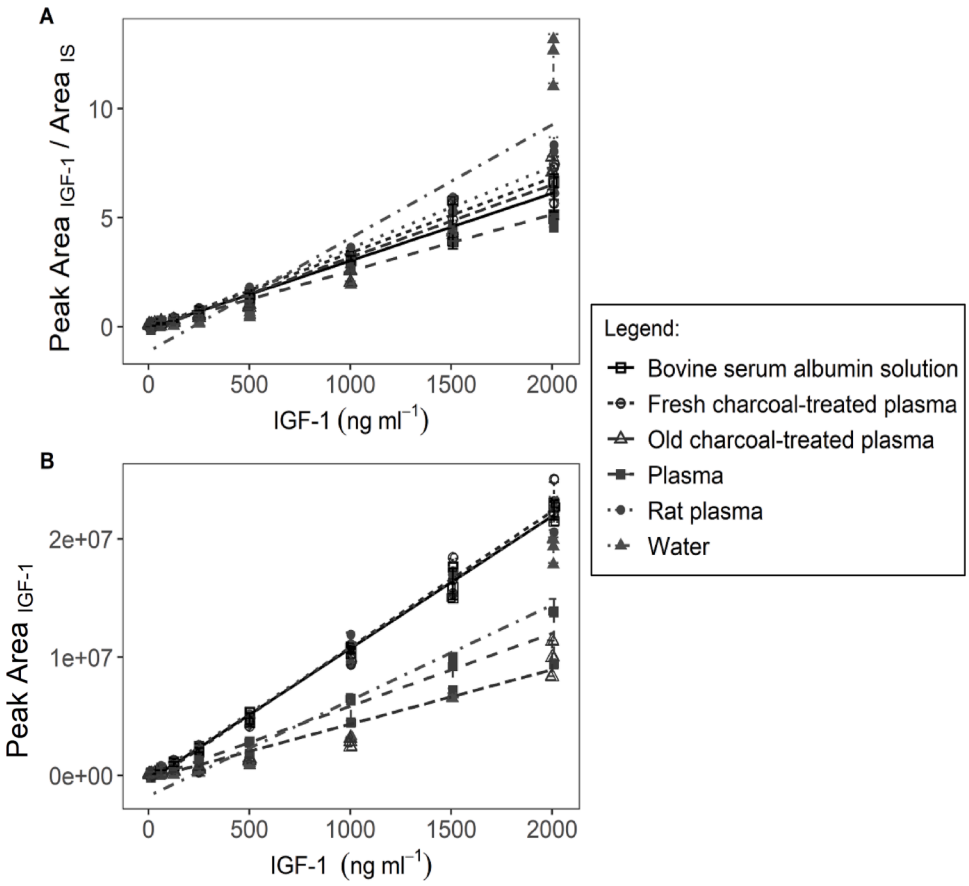

Fig. 2. IGF-1-calibration curves in different matrices, **A**, with (peak area IGF-1/peak area IS) and **B**, without (IGF-1 peak areas) correlation to internal standard.

Table 3
Individual standard sample preparations in three different concentrations and five individual repetitions. Mean SD and CV of three concentrations.

st. conc ng/ml	n1	n2	n3	n4	n5	mean	SD	CV%
62.5	63.1	66	70.6	74.5	74.3	69.7	5.1	7.3
500	443.8	446.9	511.3	506	466.1	474.8	32.1	6.8
1000	893.6	993.6	913.9	951.6	1022.6	955.1	53.7	5.6

laboratories, Cox et al. showed that reconstitution of the reference material and method of calibration were the main source of variability, whereas the impact of different instruments and sample preparation procedures on patients IGF-1 results was limited [25]. However, this study did not investigate the effect of different calibration matrices. Nevertheless, interlaboratory agreement improved substantially with a single-point human calibrator rather than a calibration curve constructed in each laboratory. The present study identified IGF-1 free human plasma as the preferred calibrator matrix. Cox et al. argue that traceable native human serum calibrators might be preferable, as they avoid variability due to the reconstitution of purified reference material and ensure that the calibrator protein is present in its native state with naturally occurring posttranslational modifications and interactions with other serum proteins, lipids, and other small molecules [25]. In line with this argument, Agger et al. have shown a better calibration for apolipoprotein A-I and B with native serum rather than with spiked purified apolipoproteins in an animal serum matrix [30]. Until today, a head-to-head comparison of calibrators in charcoal treated human plasma that is free of endogenous IGF-1 with native serum-based calibrators has not been performed. In the present study, the use of calibrators, prepared by spiking IGF-1 reference material in native human serum, resulted in 30% higher IGF-1 results than calibrators where IGF-1 was spiked in charcoal treated human plasma. This supports the concept that matrix composition, particularly the presence of various proteins and lipids, is critical for the calibration of LC-MS assays. However, on the basis of the present results it is not possible to decide if one approach is preferable over the other.

Another key finding of the present study is that IGF-1 immunoassays yielded systematically lower results than our LC-MS method calibrated with IGF-1 free human calibrators. Similar results have been reported by others before [25–27]. Interestingly, Cox et al. reported a −40% proportional bias of the SIEMENS Immulite assay when compared with their in-house LC-MS method, but only in healthy individuals. In acromegaly patients with markedly higher IGF-1 concentrations, no proportional bias was detected. This is in contrast to our data, where the SIEMENS Immulite assay and all other immunoassays tested exhibited pronounced negative proportional bias even at very high IGF-1 concentrations. In line with our results, also Pratt et al. reported an underestimation of IGF-1 with the immunoassay from IDS-ISYS across the entire concentration range [27]. The systematic underestimation of IGF-1 by automated immunoassays may have various explanations. Due to technical reasons, the separation of IGF-1 from its carriers used by immunoassays may by incomplete. In fact, the antibodies used in immunoassays and the analyzers themselves require a more gentle separation procedure than LC-MS, where strong organic solvents can be used. In addition, automated immunoassays contain detergents that prevent the formation of foam, which may compromise sample aspiration. Another aspect is that mass spectrometric methods calculate the concentration on the basis of more than one ion-specific isotopic peak, which may imply a superior sensitivity and specificity. The method used in the present study summed up the three dominant isotopic peaks of the IGF-1 ion.

Comparison of the LC-MS results with those from the four

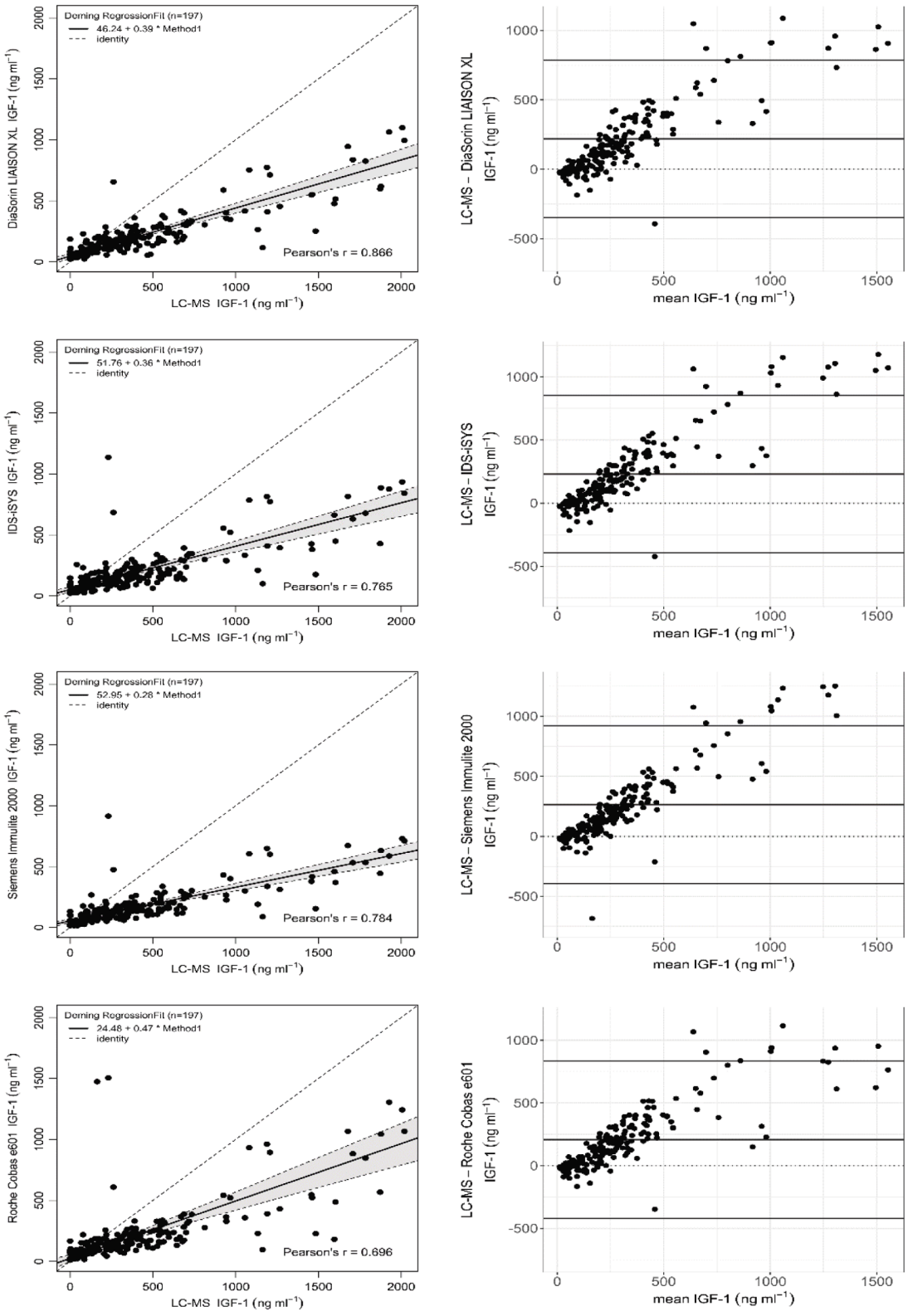

Fig. 3. Deming regression, as well as Bland-Altman plots of the data of 197 patients with acromegaly or growth hormone deficiency. Comparison of the results of four commercial immunoassays with MS-data.

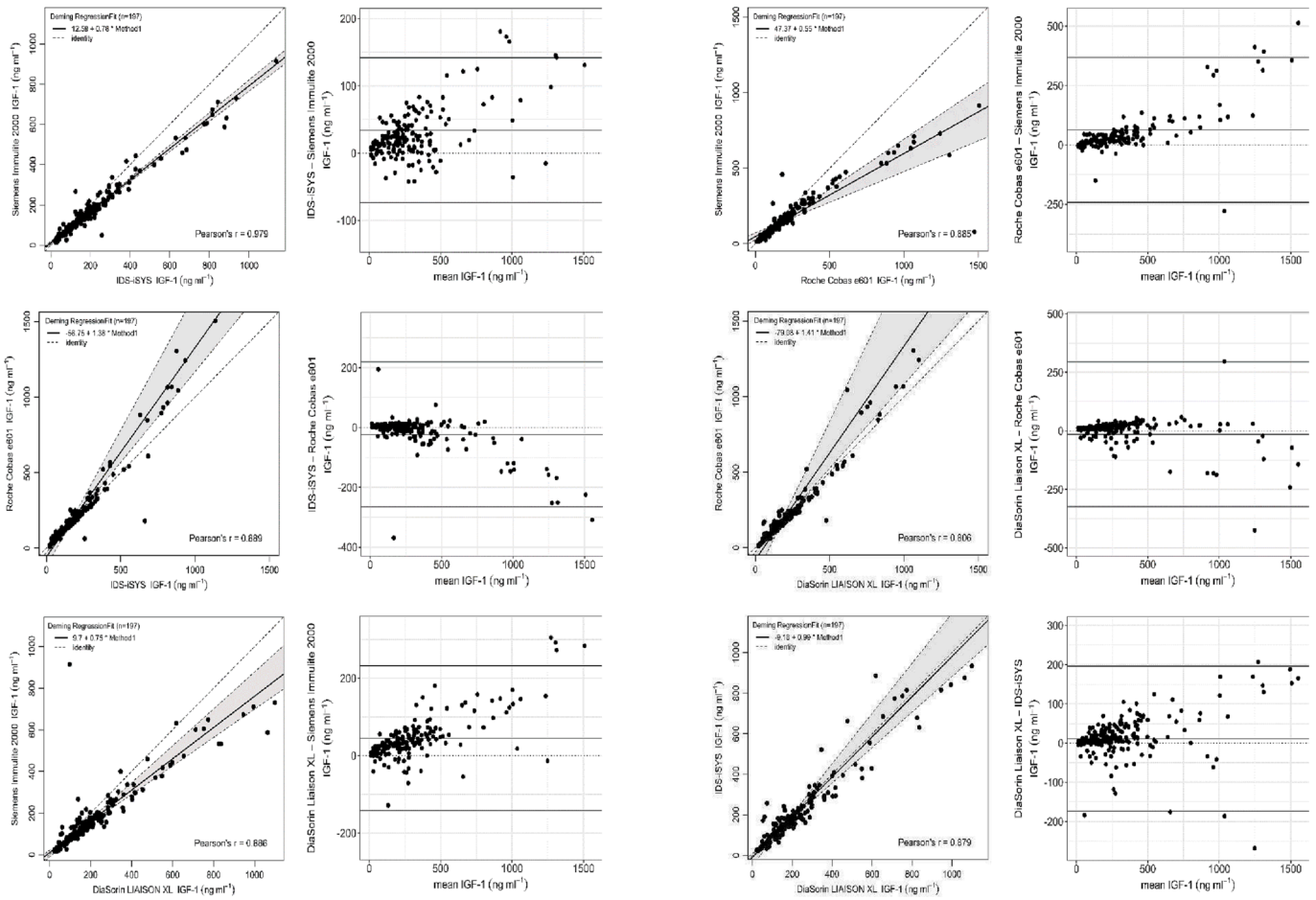

Fig. 4. Deming regression, as well as Bland-Altman plots of the data of 197 patients with acromegaly or growth hormone deficiency. Comparison of the data of four commercial immunoassays with each other.

immunoassays also showed pronounced scatter around the regression line. This phenomenon cannot be explained by differences in calibration. Such deviations may be due to the presence of IGF-1 variants in individual samples that are not detected in equimolar fashion by the antibodies employed in the different immunoassays [17]. Furthermore, interfering heterophilic antibodies and human anti-mouse antibodies (HAMAs) are a common sources of error in immunoassays that can cause falsely high and falsely low results. The prevalence of heterophilic antibodies and HAMAs has been reported to range between 0.05% and 6% [31]. Due to the efficient removal of large proteins during the sample preparation for LC-MS, this technology is rather immune against such interferences. Other common causes of aberrant results in immunoassays include macro-complexes, biotin interferences, anti-streptavidin antibodies and anti-Ruthenium antibodies.

Great systematic bias and pronounced scatter of the results were not only observed when comparing immunoassay results with those obtained by LC-MS, but also when comparing individual immunoassays with each other. In particular, the SIEMENS assay showed substantial negative proportional bias against all other immunoassays. With all assays, the scatter around the line of regression increased with higher concentrations suggesting limited accuracy at this concentration range. Bias amongst immunoassays may be the result of different calibration methods, different assay architecture, and the use of antibodies that are directed against different epitopes of the IGF-I molecule.

This study has several strengths and limitations that should be considered when interpreting the results. Although it is the first systematic comparison of different calibration matrices for the measurement of IGF-1 by LC-MS, it was not possible to include traceable human serum-based calibrators. Although such calibrators have been recommended by Cox et al., commercial products with certified concentrations are not yet available. Furthermore, the present study investigated only the impact of different calibration matrices on the comparability of results, but did not consider other sources of variation that are related to calibration, such as the number of calibrators or the quality of pipetting. The lack of therapeutic information on the patients studied here precluded a separate analysis of patients treated for growth hormone deficiency. Also, LC-MS analyses were performed months after the immunoassay measurements so that some degree of analyte degradation cannot be excluded completely. However, considering that also the different immunoassays show pronounced systemic bias that are similar in magnitude, substantial analyte degradation is rather unlikely. Although method comparisons with immunoassays have been published before, this is the most comprehensive study where four common immunoassays and an in-house LC-MS method were compared with each other.

In summary, calibration is a critical aspect for the measurement of serum IGF-1 by LC-MS. In addition to previously described factors, the type of calibrator matrix has a profound impact on patient results. The present study identified charcoal treated human plasma that is free of IGF-1 as the preferred matrix for the preparation calibrators. In addition, measurement of IGF-1 by LC-MS yields numerically different results that do not correspond to those obtained by immunoassays. In addition to systemic differences between LC-MS and immunoassays, matrix effects are rather common and may cause aberrant results in an unpredictable fashion. Therefore, laboratories should carefully select the assay that they use and perform a thorough validation that supports its use in the

local population.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C.C. van Bunderen, I.C. van Nieuwpoort, L.I. Arwert, M.W. Heymans, A.A. M. Franken, H.P.F. Koppeschaar, A.J. van der Lely, M.L. Drent, Does growth hormone replacement therapy reduce mortality in adults with growth hormone deficiency? Data from the Dutch National Registry of Growth Hormone Treatment in adults, J Clin Endocrinol Metab 96 (10) (2011) 3151–3159.

[2] S. Kannan, L. Kennedy, Diagnosis of acromegaly: state of the art, Expert Opin Med Diagn 7 (5) (2013) 443–453.

[3] J.E. Puche, I. Castilla-Cortazar, Human conditions of insulin-like growth factor-I (IGF-I) deficiency, J Transl Med 10 (2012) 224.

[4] S. Yakar, C.J. Rosen, W.G. Beamer, C.L. Ackert-Bicknell, Y. Wu, J.-L. Liu, G.T. Ooi, J. Setser, J. Frystyk, Y.R. Boisclair, D. LeRoith, Circulating levels of IGF-1 directly regulate bone growth and density, J Clin Invest 110 (6) (2002) 771–781.

[5] H.D. Cox, C.M. Hughes, D. Eichner, Sensitive quantification of IGF-1 and its synthetic analogs in dried blood spots, Bioanalysis 6 (19) (2014) 2651–2662.

[6] A. Thomas, M. Kohler, W. Schänzer, P. Delahaut, M. Thevis, Determination of IGF-1 and IGF-2, their degradation products and synthetic analogues in urine by LC-MS/MS, Analyst 136 (5) (2011) 1003–1012.

[7] P. Chanson, A. Arnoux, M. Mavromati, S. Brailly-Tabard, C. Massart, J. Young, M.-L. Piketty, J.-C. Souberbielle, Reference values for IGF-I serum concentrations: comparison of six immunoassays, J Clin Endocrinol Metab 101 (9) (2016) 3450–3458.

[8] M.L. Granada, A. Ulied, F.F. Casanueva, A. Pico, T. Lucas, E. Torres, A. Sanmartí, Serum IGF-I measured by four different immunoassays in patients with adult GH deficiency or acromegaly and in a control population, Clin Endocrinol (Oxf) 68 (6) (2008) 942–950.

[9] M.C. Broeren, J. Krabbe, L. Boesten, A.S. Hokken-Koelega, Y. de, Rijke, Impact of the choice of IGF-I assay and normative dataset on the diagnosis and treatment of growth hormone deficiency in children, Horm Res Paediatr 90 (3) (2018) 181–189.

[10] M.P. Brugts, M.B. Ranke, L.J. Hofland, K. van der Wansem, K. Weber, J. Frystyk, S. W.J. Lamberts, J.A.M.J.L. Janssen, Normal values of circulating insulin-like growth factor-I bioactivity in the healthy population: comparison with five widely used IGF-I immunoassays, J Clin Endocrinol Metab 93 (7) (2008) 2539–2545.

[11] M.B. Ranke, K.J. Osterziel, R. Schweizer, B. Schuett, K. Weber, P. Röbbel, A. Vornwald, G. Blumenstock, M.W. Elmlinger, Reference levels of insulin-like growth factor I in the serum of healthy adults: comparison of four immunoassays, Clin Chem Lab Med 41 (10) (2003).

[12] A.J. Varewijck, A.J. van der Lely, S.J.C.M.M. Neggers, L.J. Hofland, J.A.M.J. L. Janssen, Disagreement in normative IGF-I levels may lead to different clinical interpretations and GH dose adjustments in GH deficiency, Clin Endocrinol (Oxf) 88 (3) (2018) 409–414.

[13] Valid measurements of total IGF concentrations in biological fluids. Recommendations from the 3rd International Symposium on Insulin-like Growth Factors. Endocrinology, 1995; **136**(2): p. 816-7.

[14] J. Hines, et al., Detection of IGF-1 protein variants by use of LC-MS with high-resolution accurate mass in routine clinical analysis, Clin Chem 61 (7) (2015) 990–991.

[15] Irène Netchine, S. Azzi, M. Houang, D. Seurin, L. Perin, J.-M. Ricort, C. Daubas, C. Legay, J. Mester, R. Herich, Francois Godeau, Y. Le Bouc, Partial primary deficiency of insulin-like growth factor (IGF)-I activity associated with IGF1 mutation demonstrates its critical role in growth and brain development, J Clin Endocrinol Metab 94 (10) (2009) 3913–3921.

[16] R.K. Kam, C.S. Ho, M.H. Chan, Serum insulin-like growth factor I quantitation by mass spectrometry: insights for protein quantitation with this technology, Ejifcc 27 (4) (2016) 318–330.

[17] S. Kirsch, J. Widart, J. Louette, J.-F. Focant, E. De Pauw, Development of an absolute quantification method targeting growth hormone biomarkers using liquid chromatography coupled to isotope dilution mass spectrometry, J Chromatogr A 1153 (1-2) (2007) 300–306.

[18] M. Bredehöft, W. Schänzer, M. Thevis, Quantification of human insulin-like growth factor-1 and qualitative detection of its analogues in plasma using liquid chromatography/electrospray ionisation tandem mass spectrometry, Rapid Commun Mass Spectrom 22 (4) (2008) 477–485.

[19] C.E. Bystrom, S. Sheng, N.J. Clarke, Narrow mass extraction of time-of-flight data for quantitative analysis of proteins: determination of insulin-like growth factor-1, Anal Chem 83 (23) (2011) 9005–9010.

[20] G. Coppieters, P. Judák, N. Van Haecke, P. Van Renterghem, P. Van Eenoo, K. Deventer, A high-throughput assay for the quantification of intact Insulin-like Growth Factor I in human serum using online SPE-LC-HRMS, Clin Chim Acta 510 (2020) 391–399.

[21] K.J. Bronsema, F. Klont, F.B. Schalk, R. Bischoff, I.P. Kema, N.C. van de Merbel, A quantitative LC-MS/MS method for insulin-like growth factor 1 in human plasma, Clin Chem Lab Med 56 (11) (2018) 1905–1912.

[22] D. Moncrieffe, et al., Inter-laboratory agreement of insulin-like growth factor 1 concentrations measured intact by mass spectrometry, Clin Chem 66 (4) (2020) 579–586.

[23] C. Mongongu, et al., Detection of LongR, Drug Test Anal 13 (7) (2021) 1256–1269.

[24] R.G. Kay, et al., High-throughput ultra-high-performance liquid chromatography/tandem mass spectrometry quantitation of insulin-like growth factor-I and leucine-rich alpha-2-glycoprotein in serum as biomarkers of recombinant human growth hormone administration, Rapid Commun Mass Spectrom 23 (19) (2009) 3173–3182.

[25] H.D. Cox, F. Lopes, G.A. Woldemariam, J.O. Becker, M.C. Parkin, A. Thomas, A. W. Butch, D.A. Cowan, M. Thevis, L.D. Bowers, A.N. Hoofnagle, Interlaboratory agreement of insulin-like growth factor 1 concentrations measured by mass spectrometry, Clin Chem 60 (3) (2014) 541–548.

[26] Y. Seo, J. Park, M. Kim, C. Sung, O.-S. Kwon, H.J. Lee, H. Min, Optimization, validation, and comparison of a rapid method for the quantification of insulin-like growth factor 1 in serum using liquid chromatography-high-resolution mass spectrometry, Drug Test Anal 13 (2) (2021) 451–459.

[27] M.S. Pratt, M. van Faassen, N. Remmelts, R. Bischoff, I.P. Kema, An antibody-free LC-MS/MS method for the quantification of intact insulin-like growth factors 1 and 2 in human plasma, Anal Bioanal Chem 413 (8) (2021) 2035–2044.

[28] J. Hartler, A. Triebl, A. Ziegl, M. Trötzmüller, G.N. Rechberger, O.A. Zeleznik, K. A. Zierler, F. Torta, A. Cazenave-Gassiot, M.R. Wenk, A. Fauland, C.E. Wheelock, A. M. Armando, O. Quehenberger, Q. Zhang, M.J.O. Wakelam, G. Haemmerle, F. Spener, H.C. Köfeler, G.G. Thallinger, Deciphering lipid structures based on platform-independent decision rules, Nat Methods 14 (12) (2017) 1171–1174.

[29] J. Hartler, et al., Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data, Bioinformatics 27 (4) (2011) 572–577.

[30] S.A. Agger, L.C. Marney, A.N. Hoofnagle, Simultaneous quantification of apolipoprotein A-I and apolipoprotein B by liquid-chromatography-multiple-reaction-monitoring mass spectrometry, Clin Chem 56 (12) (2010) 1804–1813.

[31] J. Favresse, et al., Interferences with thyroid function immunoassays: clinical implications and detection algorithm, Endocr Rev 39 (5) (2018) 830–850.