Pilot Study for the Standardization of Insulin Immunoassays with Isotope Dilution–Liquid Chromatography/Tandem Mass Spectrometry

Diego Rodríguez-Cabaleiro, Katleen Van Uytfanghe, Veronique Stove, Tom Fiers, and Linda M. Thienpont

Background: An international working group convened by the American Diabetes Association (ADA) called for a reference measurement procedure for use in a true-ness-based standardization project of insulin immunoassays. In view of this demand, we conducted a pilot study to investigate the feasibility of such a standardization project with our isotope dilution-liquid chromatography/tandem mass spectrometry (ID-LC/tandem MS) procedure.

Methods: We evaluated the precision, accuracy, and limit of quantification (LoQ) of the ID-LC/tandem MS procedure by use of insulin-free serum supplemented with insulin to give 3 pools with concentrations of 0.0796, 0.769, and 5.56 μg/L. We conducted a pilot method comparison study with 4 immunoassays and 80 samples from fasting and glucose-stimulated patients.

Results: The within-run and total imprecision (CV) ranged from 3.2% to 6.3% and from 4.9% to 12.1% (listing sequence from the high to the low pool). The recovery from supplemented insulin-free sera ranged from 101.8% to 104.1%, and the LoQ was 0.07 μg/L (12 pmol/L). Weighted Deming regression and correlation analysis of the method-comparison data showed considerable between-assay variation for the immunoassays but, with the exception of one assay, excellent correlation with ID-LC/tandem MS. Recalibration of the immunoassay results considerably reduced the between-assay variation. Moreover, after recalibration, 3 of the 4 assays fulfilled the total error specification of 32% proposed by the ADA Workgroup.

Conclusions: Recalibration of insulin assays by regression equations established from method comparison with ID-LC/tandem MS can result in successful standardization and fulfillment of the total error criterion proposed by the ADA Workgroup.

© 2007 American Association for Clinical Chemistry

In 1996, the American Diabetes Association (ADA) task force on standardization of insulin assays reported disparate results from immunoassays. This lack of standardization hinders efforts toward achieving consistent measures for treatment guidelines. To address this issue, the ADA, in conjunction with the National Institute of Diabetes and Digestive and Kidney Diseases and the CDC, convened an international working group in 2004 to evaluate the specificity of different assays, to establish guidelines for assay acceptability, and to develop a standardization program to achieve uniform, accuracy-based values. A recent report on the progress made by the working group concluded that “a common insulin reference preparation did not change the among-assay CV and failed to improve harmonization of results among assays” and that the “investigation of a reference measurement procedure for insulin should be a priority to provide a metrologically appropriate basis to evaluate the accuracy of routine methods.” The ADA Workgroup proposed the following performance specifications for such a procedure: a limit of quantification (LoQ) of 0.07 μg/L (= 12 pmol/L or 1.73 mIU/L), a total maximum CV of 3% (6%–7% at the LoQ), and a bias limit of 5% (3). Similarly,
a European group stated in a recent report that “a reference method for measurement of insulin is required with an international standard defined in terms of mass” (4).

Here we report on a pilot study that investigated the standardization of 4 insulin immunoassays by method comparison with an isotope dilution–liquid chromatography/tandem mass spectrometry (ID-LC/tandem MS) procedure. We purposely undertook the study in an early developmental stage of the isotope dilution–mass spectrometry procedure to have a fast answer on questions about the feasibility of the project and the expected impact on immunoassay results after recalibration, and, also, to do a focused refinement of our isotope dilution–mass spectrometry procedure.

Materials and Methods

Materials
All solvents used were at least analytical grade. For immunoadfinity chromatography, a commercial immunoaffinity gel (Laboratoire d’Hormonologie, C.E.R. Groupe) with a capacity of 10 g IgG/L [mouse monoclonal antibody (IgG1,k) coupled to CNBr Sepharose] was used (5). The gel was stored in a buffer at pH 7.4 containing 50 mmol/L Tris and 0.5 g/L sodium azide in water (Trizma® hydrochloride, Trizma base, and sodium azide, all from Sigma). Sep-Pak C18 1-cc cartridges (30 mg; Waters) were used for solid phase extraction.

Calibration Materials and Procedures
Recombinant human insulin expressed in yeast (Sigma-Aldrich) was used as the calibration material. It was delivered in vials containing 50 mg of freeze-dried material and had a purity and potency of >98% and 27.5 kIU/g, respectively (according to the manufacturer’s information). 4-[D10]Leu-insulin, labeled in positions 6, 11, 15, and 17 of the β-chain, was used as internal standard (IS; Peptide Institute). In addition, all volumetric steps involved in the preparation/sampling of the calibration solutions were gravimetrically controlled, so that the concentrations were accurately known to 4 significant digits.

Procedure 1. The calibration stock solution was prepared by weighing and dissolving a minimum of 5 mg of calibration material in 1 mL of a mixture of 50 mmol/L hydrochloric acid in water (600 mL/L) and glycerol (400 mL/L). The stock solution was diluted in 2 consecutive steps with the same solvent to obtain a final concentration of 25 mg/L (0.62 kIU/L). Immediately after preparation, the latter solution was divided into 50-µL aliquots and frozen at −20 °C until the day of analysis. These portions were never used more than 4 times. Each day of analysis, fresh working solutions with a concentration of 20 µg/L (0.5 IU/L) were prepared by diluting a portion of the stock solution of 25 mg/L with a mixture of 600 mL/L water and 400 mL/L glycerol, containing 0.1 g/L protease-free BSA. Note that in every case 2 individually weighed stock solutions were diluted to prepare a total of 4 different working solutions (2 from each stock). The stock and working solutions of the IS were prepared in an identical way. We pipetted 30 µL of a mixture of 600 mL/L water and 400 mL/L glycerol, containing 0.1 g/L BSA and 10 mg/L glargine (Lantus® from Sanofi-Aventis) into the empty vial, which was vortex-mixed before adding the calibration working solutions. Glargine was used in this and other steps (see also procedure 2) because it proved to have a carrier effect.

Procedure 2. The calibration stock solution was exactly prepared as described in procedure 1. Subsequently, it was diluted in 2 consecutive steps with insulin-free serum (Scipac) to obtain a final concentration of 25 mg/L (0.62 kIU/L). The pH of the insulin-free serum was adjusted to 7.4 by adding 100 µL of HEPES buffer (185 g/L HEPES acid and 8.8 g/L NaOH, both from Sigma) per mL of serum. Immediately after preparation, the latter solution was divided into 50-µL aliquots and frozen at −20 °C until the day of analysis. On that occasion, a frozen vial of the stock solution of 25 mg/L was thawed and diluted with a mixture of 600 mL/L water and 400 mL/L glycerol containing 0.1 g/L BSA and 10 mg/L glargine to obtain working solutions of 20 µg/L (0.5 IU/L). This dilution was never reused. Note that in all cases 2 individually weighed stock solutions were diluted to prepare a total of 4 different working solutions (2 from each stock). The stock and working solutions of the IS were prepared as described in procedure 1. We pipetted 100 µL of a mixture of 600 mg/L water and 400 mL/L glycerol, containing 0.1 g/L BSA and 10 mg/L glargine into the empty vial and vortex-mixed before adding the calibration working solutions.

In both procedures low protein-binding Sorensen® vials (BioScience) were used. In addition, all volumetric steps involved in the preparation/sampling of the calibration solutions were gravimetrically controlled, so that the concentrations were accurately known to 4 significant digits. Finally, the calibrators were prepared by mixing of equivalent amounts of the insulin and IS working solutions. The calibration mixtures were stored at −20 °C until the time of analysis. We verified against calibrators through analysis that they could be analyzed as such with LC/tandem MS. A one-point calibration procedure was used (6).

Sample Purification Procedure
The procedure started with gravimetric sampling of the serum, between 0.3 mL (minimum) and 4 mL (maximum) depending on the concentration (on the basis of data available from the Roche immunoassay) and available volume. Ideally an absolute amount of insulin of 2 ng was processed. A known amount of IS was added to obtain an isotope ratio of 1:1 within ±15%, and the samples were equilibrated for at least 30 min. Samples with a total amount of insulin <0.6 ng were analyzed at an isotope ratio of 0.5:1 together with a set of calibration mixtures prepared at the same ratio.
Samples were purified by a combined immunoaffinity chromatography–solid phase extraction procedure as previously described (7). The immunoaffinity gel-containing columns were used in the gravity feed mode. Conditioning was done by adding 3 mL of binding buffer (50 mmol/L Tris, pH 7.8, in water) for immediate application of up to 4 mL of serum diluted 1:1 with binding buffer. Then, the gel was rinsed with 5 mL of binding buffer, 4 mL of 10 mmol/L ammonium acetate (pH 4.5, Fluka), and 0.2 mL of 1 mL/L trifluoroacetic acid in water (this first rinse with trifluoroacetic acid did not cause elution because the volume corresponds to the dead volume of the column; Fluka). Human insulin and 4-[D10]Leu-human insulin were subsequently eluted with 2 times 0.4 mL of 1 mL/L trifluoroacetic acid in water, taking care that a time interval of 5 min elapsed between application and elution of the trifluoroacetic acid solutions. Finally, the immunoaffinity chromatography columns were reequilibrated with 5 mL of binding buffer. The columns were repeatedly used (15–20 times), which necessitated that care was taken to never let the gel run dry. To this end, columns equipped with a 2-way stopcock were used. If not used for a longer period, the binding buffer was replaced by a buffer at pH 7.4 containing 50 mmol/L Tris and 0.5 g/L sodium azide in water. Carryover upon reuse was checked and found negative by analysis of several blanks after processing highly concentrated sera.

The eluate from immunoaffinity chromatography was loaded into a Sep-Pak C18 cartridge, mounted upside down, and preconditioned with 2 mL of acetonitrile and 2 mL of 1 mL/L trifluoroacetic acid in water. The upside-down mounting was done by connecting (with tubing) the loading side of a cartridge containing the solid phase to the vacuum manifold. Subsequently, the cartridge’s elution side was connected to the same side of an empty cartridge, serving as solvent container. The idea behind this reverse loading was to restrict the adsorption of insulin to a small part of the solid phase, hence minimizing losses and reducing the elution volume. After being loaded, the cartridge was washed with 2 mL of a 10/90/0.1 (vol/vol/vol) mixture of acetonitrile, water, and trifluoroacetic acid. Finally, the cartridge was mounted in the normal flow direction for elution with 0.25 mL of a 70/30/0.1 (vol/vol/vol) mixture of acetonitrile, water, and trifluoroacetic acid. The extracts were evaporated to dryness at 50 °C with N₂ and the residues stored at −20 °C until analysis. Before processing by LC/tandem MS, 0.25 mL of cyclohexane (Merck) was added, vortex-mixed, and removed. The remaining residue was dried and reconstituted in 50 µL of a mixture of 100 mL/L acetonitrile and 900 mL/L water for injection into the chromatographic system of 15 to 45 µL (depending on the concentration).

**ID-LC/Tandem MS Conditions**

MS was done with an API4000™ system (Applied Biosystems), operated in the positive electrospray ionization mode and connected with a HP 1100 HPLC system (Agilent). The latter was equipped with a Waters XBridge C18 column (2.1 × 10 mm, 5 µm) maintained at room temperature. Details on the gradient HPLC and tandem MS conditions are described elsewhere (7). Briefly, gradient elution was done with eluents A and B containing a mixture of acetonitrile, water, ammonium hydroxide (solution of >25% ammonia in water; A: 10/90/0.025 (vol/vol/vol); B: 70/30/0.025 (vol/vol/vol)). The percentage of acetonitrile was increased from 10% to 40% in 3 min at a flow rate of 0.25 mL/min and kept constant during 3 min. Afterward, it was increased to 70% in the next 0.1 min and kept constant again for 3 min to clean the column. Finally, the column was reequilibrated with eluent A for 5 min. Positive electrospray-tandem MS was done by monitoring the transitions of the 4X-protonated molecular ions [M+4H]⁺ from m/z 1453 to 226 (insulin) and 1463 to 226 (4-[D10]Leu-insulin) with a dwell time of 0.5 s. The declustering potential was 50 V, the collision-induced dissociation energy 85 eV, and the N₂ collision gas setting 12.

**VALIDATION OF THE ID-LC/TANDEM MS PROCEDURE**

The performance of the ID-LC/tandem MS procedure (imprecision, accuracy, and trueness) was validated with insulin-free serum supplemented with human insulin to obtain 3 pools, a low pool at 0.0796 µg/L (= 1.97 mIU/L, the LoQ specification by the ADA Workgroup) and medium and high level pools at 0.769 µg/L (= 19.0 mIU/L) and 5.56 µg/L (= 137.8 mIU/L), respectively. Aliquots were stored at −20 °C and analyzed in duplicate on 6 different days (n = 12). The insulin content (95% CI) of the insulin-free serum was determined beforehand.

**METHOD COMPARISON**

For the method comparison, serum samples from 16 ambulatory patients (4 males and 12 females, age 21–74 years) from Ghent University Hospital were used. The patients had been subjected to an oral glucose (75 g) tolerance test after overnight fasting. The insulin tests were ordered by the endocrinologist for assessment of glucose tolerance, glucose handling, β cell function, and degree of insulin sensitivity in subjects with morbid obesity and/or other risk factors for insulin resistance and glucose intolerance. Handling of patient samples was done according to the local Ethics Committee guidelines. Blood was drawn into Venosafe VF-106SAS tubes (Terumo) before the glucose load and after 30, 60, 120, and 180 min. The collected blood was allowed to clot for at least 30 min and centrifuged at 1500 g for 10 min. The fasting blood samples were transported to the laboratory within 30 min after withdrawal, whereas the remaining samples were received within 30 min after the last collection. Then, the analysis was immediately started with the 4 immunoassays. The remaining serum was transported ice-cooled to the MS laboratory. There, the samples were immediately processed by combined immunoaffinity chromatography-solid phase extraction as described above. The evaporated extracts were stored −20 °C until LC/tandem MS analysis on the next day.
The method comparison was done with the 4 following assays: electrochemiluminescence immunoassay for use on the Immulite® 2000 Insulin chemiluminescent immunoassay from Diagnostic Products Corporation (DPC), the Modular Analytics E170 (Roche Diagnostics GmbH) immunoassay analyzer, the AsSYM Insulin microparticle enzyme immunoassay from Abbott, and the Access Ultra-sensitive Insulin chemiluminescent immunoassay from Beckman Coulter. All assays were calibrated against the WHO Insulin 1st International Reference Preparation 66/304 (National Institute for Biological Standards and Control). All measurements with the commercial insulin immunoassays were performed with the same batch and in accordance with the respective manufacturer’s instructions for use. Each assay’s performance was controlled either with assay-specific control samples (for DPC the Insulin Controls LINC1 and LINC2 with assigned target values of 9.2 and 48 mIU/L; for Abbott the AsSYM Insulin L, M, and H with targets of 8, 40, and 120 mIU/L), with commercial control materials (for Beckman the Lyphochek Immunoassay Plus Control Levels 1, 2, and 3 from Bio-Rad with assigned values of 7.1, 36, and 95 mIU/L) or with in-house prepared serum pools (for Roche) targeted at 11.1 and 60.0 mIU/L.

The measurements in the ID-LC/tandem MS laboratory were done singly according to a measurement protocol with bracketing of the patient and control samples (maximum 6) between 2 times 3 calibrators, which allowed analysis of samples of 2 stimulated patients per working day. The values of all calibrators were used for the calculation of the sample concentration. Measurement of the 16 stimulation curves by both the ID-LC/tandem MS and routine laboratory was done in 8 different analytical runs (batches of 10 samples and 1 control sample) spread over 3 weeks. The quality of each analytical run was controlled by the use of a high-level lyophilized control material (estimated mean 4.426 µg/L; total CV 2.6%; n = 8; INSTAND e.V.).

**STATISTICAL DATA ANALYSIS**

Statistical data analysis was done with CBstat, Version 5.1 (CBstat software) and MedCalc, Version 9.2.0.0 (MedCalc software). Model II ANOVA was used to calculate the within-, between-, and total CV (%). For the recovery data, CIs (95%, 2-sided) were calculated and a 1-sample t-test was performed. For analysis of the method comparison data, weighted Deming regression and correlation were done. The 2 highest concentration values were not taken into account for the regression and correlation analysis, because they are outliers with respect to the distribution of the insulin concentrations. In the graphical plots, however, the regression line is extrapolated to show the performance of the assays in the high concentration range.

**Results**

The preliminary performance characteristics of the ID-LC/tandem MS procedure are presented in Table 1. The insulin content (95% CI) of the insulin-free serum used to prepare the 3 pools was estimated to be 8 ng/L (1–15 ng/L; n = 6). The within-run imprecision ranged from 3.2% to 6.3%, the total imprecision from 4.9% to 12.1% (listing sequence from the high to the low pool). In all cases the between-run precision was greater than the within-run precision. Analysis of the low pool (target concentration, 0.0796 µg/L) was used to confirm that our measurement procedure met the LoQ demand of the ADA Workgroup, i.e., 0.07 µg/L (12 pmol/L) as the lowest concentration that the measurement procedure should be able to measure with a total CV of 6%–7%. At that LoQ concentration, our measurement procedure had a mean signal-to-noise ratio of 25 with an imprecision and 95% CI around the mean as shown in Table 1. Achieve-

<table>
<thead>
<tr>
<th>Sample</th>
<th>Low (LoQ)</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean, µg/L</td>
<td>0.0829</td>
<td>0.792</td>
<td>5.656</td>
</tr>
<tr>
<td>Within-run CV, %</td>
<td>6.3</td>
<td>3.5</td>
<td>3.2</td>
</tr>
<tr>
<td>Between-run CV, %</td>
<td>10.3</td>
<td>5.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Total CV, %</td>
<td>12.1</td>
<td>6.1</td>
<td>4.9</td>
</tr>
</tbody>
</table>

**Trueness data (design: duplicates on 6 days; n = 12)**

| Target, µg/L | 0.0796 | 0.769 | 5.556 |
| Mean, µg/L | 0.0829 | 0.792 | 5.656 |
| 95% CI, % | 11.8 | 5.8 | 4.6 |
| Recovery, % | 104.1 | 102.9 | 101.8 |
| LCL, % | 92.4 | 97.1 | 97.2 |
| UCL, % | 115.9 | 108.8 | 106.4 |
| P<sup>a</sup> | 0.424 | 0.258 | 0.369 |

<sup>a</sup> Spike plus content of insulin-free serum.

<sup>b</sup> LCL, Lower confidence limit; UCL, upper confidence limit.

<sup>c</sup> Probability of the 1-sample t-test.
ment of this LoQ on patient samples was confirmed from analysis of the lowest concentrated patient sample (1.83 mIU/L or 0.073 μg/L). The recovery of insulin added to insulin-free sera ranged from 101.8% to 104.1% (identical listing sequence as before).

The results of the method comparison study are summarized in Figs. 1–3. Fig. 1 shows for each immunoassay the scatter and absolute difference plots before recalibration. In Fig. 2 the combined scatter and absolute difference plots before (A and B) and after recalibration with the

---

**Fig. 1.** Method comparison study between 4 immunoassays and ID-LC/tandem MS. For each immunoassay the scatter and absolute difference plots before recalibration are shown.
respective weighted Deming equation (C and D) are given. Fig. 3 documents for the individual assays the percentage difference plots after recalibration. The units for the reported results (mIU/L) were selected because the 4 manufacturers use different conversion factors for the calculation of results in mass units (pmol/L). To convert the ID-LC/tandem MS results into international units (IU), we used a relative molecular mass for insulin of 5807.6 and a factor of 6.945 for the conversion (mIU/L to pmol/L). The difference plots of the recalibrated data include for the results of the individual manufacturers the 32% total error limit as proposed by the ADA Work-

Fig. 3. Percentage difference plots after recalibration for each individual immunoassay.

The dotted lines at ±32% represent the total error specification proposed by the ADA Workgroup.

The symbols used are those of Fig. 1: ○ for DPC, × for Roche, □ for Abbott, and △ for Beckman.
group for a routine procedure (2). With ID-LC/tandem MS, the basal insulin concentrations ranged from 3.74 to 14.7 mIU/L and the stimulated concentrations from 1.83 to 247.6 mIU/L. Fig. 4 shows typical ion chromatograms obtained for one of the processed serum samples. With an insulin concentration of 3.2 mIU/L (22 pmol/L), the sample is 1 of the 4 sera with the lowest concentrations of the 80 sera analyzed in the study. The weighted Deming regression results for the method comparisons are shown in Table 2. The observed bias was +20% for Roche, +42% for DPC, −17% for Beckman, and +6% for Abbott. The SD of the differences of all assay results from ID-LC/tandem MS was 9.6 mIU/L (8.9 mIU/L without the DPC assay). The weighted correlation data were \( r = 0.991 \) for Roche, \( r = 0.920 \) for DPC, \( r = 0.993 \) for Beckman, and \( r = 0.992 \) for Abbott.

**Discussion**

This pilot study demonstrated the feasibility of standardization of insulin immunoassays by method comparison with an ID-LC/tandem MS procedure. In comparison with the performance specifications proposed by the ADA Workgroup for a reference measurement procedure used for that purpose, the CV data demonstrate that some improvement is needed, in particular, in terms of the imprecision at the LoQ and the between-run component of imprecision. This problem will be addressed by increasing the number of independent calibrators and improvement of the absolute recovery of insulin during the overall procedure. On the other hand, the data show that the method achieved the proposed LoQ of 0.07 μg/L with a reasonable signal-to-noise-ratio and satisfies the bias limit of 5%.

The results for the method comparison with 4 immunoassays (Figs. 1–3) show a considerable between-assay variation; however, with the exception of the DPC assay, they document an excellent correlation with ID-LC/tandem MS, which was a valid basis for recalibration on the basis of the weighted Deming regression equation. This decreased the SD of the differences of all assay results from ID-LC/tandem MS from 9.6 mIU/L to 5 mIU/L (from 8.9 to 3.4 mIU/L without the DPC assay; compare also Fig. 2, A and B, with Fig. 2, C and D). These data show that the variations in the recalibrated data are dominated by the DPC assay. This fact is most obvious in the percentage difference plot of this assay (Fig. 3). The plots in Fig. 3 further show that after recalibration all assays, with the exception of DPC, fulfill the total error specification of 32% set by the ADA Workgroup (2).

In conclusion, our pilot study demonstrated that it is feasible to standardize insulin immunoassays by method comparison with an ID-LC/tandem MS procedure. For the investigated immunoassays the impact of recalibration on the results was considered moderate.

**Grant/funding support:** This research was in part supported by the American Diabetes Association (Award no. 10-06-Insulin to the laboratory of L.M.T.). Purchase of isotopically labeled insulin was funded by a gift from Prof. H. Reinauer (INSTAND e.V., Düsseldorf) to the laboratory of L.M.T.

**Financial disclosures:** None declared.

**Acknowledgments:** We thank Prof. W. Lambert (Laboratory for Toxicology, Ghent University) for offering measurement time with the LC/MS infrastructure in his laboratory [GOA 12051204 (Bijzonder Onderzoeksfonds UGent)]. Prof. J.M. Kaufman is acknowledged for his interest in the study and permission to do the immunoassay analyses in the laboratory of his department.

**References**


---

**Table 2. Agreement of selected methods (y) with mass spectrometric method (x).**

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Slope (SE)</th>
<th>Intercept (SE), mIU/L</th>
<th>( S_{x,y} ), mIU/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbott</td>
<td>1.08 (0.02)</td>
<td>−0.33 (0.21)</td>
<td>0.11</td>
</tr>
<tr>
<td>Beckman</td>
<td>0.84 (0.01)</td>
<td>−0.20 (0.12)</td>
<td>0.09</td>
</tr>
<tr>
<td>DPC</td>
<td>1.27 (0.07)</td>
<td>+0.74 (1.28)</td>
<td>0.34</td>
</tr>
<tr>
<td>Roche</td>
<td>1.21 (0.02)</td>
<td>−0.33 (0.31)</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*The slopes and intercepts of the regression equations were derived by weighted Deming regression analysis.*

