Comparison of 11 Human Insulin Assays: Implications for Clinical Investigation and Research

Susan E. Manley,1* Irene M. Stratton,2 Penelope M. Clark,3 and Stephen D. Luzio4

Background: The American Diabetes Association task force on standardization of insulin assays in 1996 showed wide variation in assay bias. Newer assays are specific for insulin, with several now available on automated immunoassay analyzers.

Methods: In 2004, we compared 11 commercially available insulin assays by analyzing 150 serum samples (99 fasting/51 postprandial) from study participants with various degrees of glucose intolerance (exclusions being type 1 diabetes, insulin treatment, or presence of insulin antibodies). All assays were calibrated against International Reference Preparation 66/304. One assay was not specific for insulin and another was an RIA; 10 assays used enzyme/chemiluminescent labels. Bland–Altman difference plots were modified to use the mean insulin from all assays on the x-axis as a common comparator.

Results: As in the 1996 study, insulin values from the different assays varied by a factor of 2, with the non-specific assay ranking in the middle of the distribution. Spearman rank correlation coefficients, for ranking samples vs the mean, were 0.983–0.997. Both offsets and concentration-dependent differences were seen in the modified difference plots. Imprecision (mean CV) for automated assays (3%) was not significantly different from the mean CV for manual assays (5%). Similar values were obtained when one automated assay was run in laboratories in both the UK and the US. Results of 1 assay showed lower insulin concentrations in heparinized plasma than in serum.

Conclusions: Assay performance must be considered before comparing insulin results. The 2-fold variation in insulin results may be related to specificity, manufacturer calibration procedures, or conversion factors.

In 1996, the American Diabetes Association (ADA) task force on standardization of insulin assays reported widely disparate results (1). Results for plasma and serum from the 17 assays studied, mostly RIAs, varied by a factor of 2. Use of the same insulin reference preparation did not improve comparability, and the same assay method run in 2 laboratories yielded different results. Only 2 of the assays studied in the ADA comparison, an RIA and ELISA, were available for inclusion in this study.

Because of this bias, the task force did not recommend the use of different insulin assays within a population or research study unless their guidelines for assay validation and comparison with other laboratories were followed. To address these issues, major clinical trials (2–6) have organized comparisons of insulin assays with patient samples or sample exchanges with other laboratories.

Immunometric assays specific for insulin (based on low cross-reactivity to proinsulin/s quoted by the manufacturers) are widely available, commercially, with many assays run on automated platforms (7–10). These assays, however, may not recognize insulin analogs for treatment of diabetes (11–14). No international reference method for insulin has as yet been established (15).

External quality assessment schemes still show large differences in insulin values, e.g., the March 2003 UK National External Quality Assessment Service (UK NEQAS) Guildford, which involved 43 assays and 17 methods, found that the trimmed mean insulin concentration in human serum samples for all laboratories was 87 pmol/L, range 40–145 pmol/L, and for serum with...
added human recombinant insulin was 274 pmol/L, range 160–475 pmol/L (Gwen Wark, personal communication, March 7, 2006).

Specificity, sensitivity, and working ranges are key issues for measurement of circulating insulin for the clinical investigation of hypoglycemia (7). Calculated indices such as those derived from Homeostasis Model Assessment (HOMA) for insulin resistance and β-cell function, widely quoted in research, are dependent on the absolute value for insulin and on specificity in particular (16).

We compared commercially available insulin assays by assaying serum samples from study participants with varying degrees of glucose tolerance.

Materials and Methods
In this study, performed in 2004, insulin concentrations in fasting and postprandial serum samples from study participants with varying degrees of glucose tolerance were measured by use of 11 insulin assays performed in 7 UK and 2 US laboratories. Reconstituted, lyophilized QC material was also circulated. Total proinsulin was measured in one center, insulin compared in serum and heparinized plasma in another, and the same assay compared in 2 laboratories in the UK and US.

Ethics approval was obtained from the South East Wales Local Research Ethics Committee and complies with the current revision of the Helsinki Declaration.

Patients
Fasting or postprandial blood samples were collected from 138 participants, 18–75 years of age. Sample donors were either nondiabetic or had impaired fasting glucose, impaired glucose tolerance, or type 2 diabetes. Patients with type 1 diabetes or type 2 diabetes requiring insulin treatment or who were recorded as insulin antibody–positive were excluded. Anonymized data were collected, including age, ethnicity, height, weight, and whether the sample was taken fasting or postprandial. All patients gave written informed consent.

The 150 blood samples (fasting/nonfasting 99/51) were collected from 43 healthy persons (fasting/nonfasting 14/29), 20 patients with impaired fasting glucose or impaired glucose tolerance (fasting/nonfasting 18/2), and 87 patients with type 2 diabetes (fasting/nonfasting 67/20). Both fasting and nonfasting samples were obtained from 12 study participants. The mean (SD) age and body mass index of the normoglycemic study participants were 43.8 (14.3) years and 28.5 (6.2) kg/m². The corresponding values for those with impaired fasting glucose or impaired glucose tolerance were 62.0 (10.1) years and 31.0 (4.9) kg/m², and for those with type 2 diabetes 60.0 (9.4) years and 31.9 (5.6) kg/m².

Collection of Blood
Over a 12-week period in clinics at the Diabetes Research Unit, Cardiff University, UK, blood samples were obtained from study participants via a vacutainer or intravenous cannula with a 3-way tap and collected into a plain vacutainer (yellow, 3.5 mL, Greiner Vacuette Z Serum Sep Clot Activator, Greiner Bio-One Ltd) or lithium heparin vacutainer (green, 4 mL, Greiner Vacuette Lithium Heparin). Samples were placed on a roller immediately and serum left to clot for 30 min. Blood samples were spun at 2000 g for 5 min at 4 °C. No samples with hemolysis or lipemia were collected. Immediately after centrifugation aliquots (0.5 mL) were pipetted into tubes (Sarstedt Ltd) labeled with laboratory and sample codes. The tubes were frozen immediately at −20 °C and transferred to a −70 °C freezer within 24 h.

QC Materials
Trilevel, lyophilized QC material (Lyphochek Immunoassay Plus Control, Bio-Rad Labs Ltd) was reconstituted according to manufacturer instructions, divided into aliquots, and stored at −20 °C.

Shipment and Measurement of Samples and QC Samples
Samples were shipped on dry ice by an overnight delivery service. On receipt, laboratories stored samples at −20 °C (or −70 °C if available) immediately and did not to allow the samples to thaw.

Insulin and Proinsulin Assays
Study organizers selected the most commonly used, commercially available insulin assays from those used by participants in the UK NEQAS Guildford. Assay information obtained from the manufacturers is shown in Table 1. Manual and automated assays for clinical investigation and research were included; some were approved for research use only. Assays from the US were included, one from the previous comparison (1) and another used for major epidemiological studies (17, 18). The Tosoh ST AIA assay was included twice, once from a US research laboratory and once from a routine UK hospital laboratory. All laboratories requested serum when given the option of serum or heparinized plasma. Both types of samples were analyzed with the Biosource-Invitrogen (previously Medigenix) assay. Participating centers were asked to use their laboratory’s standard operating procedures to analyze samples soon after receipt. Serum samples were measured in batches of 30 over a 5-day period. Reconstituted trilevel QC material provided by the study was measured with each batch of samples. Total proinsulin was measured in serum by an ELISA (Linco Total Proinsulin ELISA, Linco Research Inc) at Linco Research Inc, St Charles, Missouri, US, with interassay imprecision (CVs) of 2.9%–8.3% at concentrations of 6.52–80.6 pmol/L.

Role of the Sponsors
The funding organizations had no role in the design of the study, data collection, analysis, or interpretation, or preparation of the manuscript other than diagnostic companies confirming the information published in leaflets for Table 1, and did not approve or disapprove of, or delay publication of the work.
Data were analyzed by use of the Statistical Analysis System (SAS Institute) (19). Because there is no consensus regarding statistical methods for comparison of this number of insulin assays without a reference method, we compared individual assays. We found no aberrant assays, so we then compared each of the 12 sets of assay results with the overall mean. The method of Carstensen (20) was not appropriate, as it requires replicate measurements within an assay. Difference plots (21, 22) were modified to use the mean of all results for the x-axis as a common comparator in the Figs. in this paper. Bland–Altman difference plots for pairs of assays (one assay compared with one other) are presented in the online Data Supplement (see the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol53/issue5). Each data point, zero (nil difference) lines, mean differences, and 1.96 SD lines are shown. Ordinary least squares linear regression was used for calculation of intercepts and slopes in the modified difference plots; Spearman rank correlation coefficients were calculated for each assay vs the mean.

Results

INSULIN ASSAYS
Serum was assayed in 7 UK and 2 US laboratories between May and July 2004 with commercial insulin assays (Table 1) described by the manufacturer as standardized against the first International Reference Preparation 66/304 (IRP 66/304). Eleven immunometric assays using various separation and detection systems were included, i.e., 4 immunoenzymometric assays, 2 immunochemiluminometric assays, 1 immunoelectrochemiluminometric assay, 3 ELISAs, and 1 RIA. Only the ELISA from Mercodia was not specific for insulin, showing 54% cross-reactivity with intact proinsulin (manufacturer’s data). Six assays were run on automated, high-throughput immunocross platforms in routine clinical laboratories, and 6 were run manually. Sample volume varied from 20 to 100 μL, not including dead volume for sampling with automated analyzers or the volume for duplicate assay if recommended by the manufacturer. Sensitivity, as stated by the manufacturers, ranged from 0.7 to 12 pmol/L. Some manufacturers provided details on the experiment used to determine sensitivity; it was usually determined by replicate analyses of a matrix material free of insulin. The mean of the zero signal plus 2 SD was used for 6 assays and mean plus 2.5 SD for another, but was not defined for the others. The upper concentration limit above which dilution was recommended ranged from 600 to 6945 pmol/L. Conversion factors, from mIU/L to pmol/L, varied from 6 to 7.46; Tosoh updated its conversion factor from 6 to 7.17 in July 2006.

ASSAY IMPRECISION
Imprecision, the mean of CVs for trilevel QC analyzed on 5 days, varied from 1.4% for 1 automated assay to 6.7% for
a manual ELISA (Table 1). The mean CV for automated assays (Table 1) was 3.0% (range 1.4%–5.6%) and 5.4% for manual methods (range 3.9%–6.7%), but this difference was not significant. Imprecision for the low QC varied from 1.1%–2.1% (Table 2) and for the high QC from 1.1%–7.1%. CVs were constant across the range for 2 automated assays (mean CVs 1.4% and 2.5%), but imprecision for the low QC in 2 laboratories using the same assay was 9.4% in one laboratory and 3.1% in the other.

Insulin values for the low and mid-range QCs varied by a factor of 3, with mean values ranging from 35.9 to 106.5 pmol/L for the low QC and 83.8 to 291.1 pmol/L for the mid-range QC. For the high QC, values varied by a factor of 4, from 167.4 pmol/L to 709.8 pmol/L. One assay (DPC Immulite 1000) consistently reported low values for QCs. The direction and magnitude of bias with reconstituted QCs was not necessarily reflected in the results for serum (Table 2). The source of insulin (bovine or recombinant human) was not stated by the manufacturer of the QC materials.

**INSULIN MEASUREMENT IN HEPARINIZED PLASMA AND SERUM**

Measured insulin concentrations were higher in serum, median (IQ range) 132 (90–205) pmol/L, minimum 38 pmol/L, and maximum 1387 pmol/L, than plasma, 121 (85–168) pmol/L, minimum 27 pmol/L, and maximum 1402 pmol/L, with differences proportional to the concentration (Fig. 1). The linear regression equations were:

\[
\text{plasma insulin} = [0.86 \times \text{(serum insulin)} - 0.75] \text{ pmol/L},
\]

or

\[
\text{serum insulin} = [1.25 \times \text{(plasma insulin)} + 7.77] \text{ pmol/L}.
\]

**INSULIN AND PROINSULIN**

Insulin results for serum from the different assays varied by a factor of 2 for median and maximum values (Table 2). The median and maximum values for the nonspecific assay ranked in 7th place (starting with the lowest value). Minimum values for insulin in human samples varied by a factor of 20 (Table 2), with values of 10 pmol/L obtained from 2 automated assays (Tosoh and Roche) and the nonspecific ELISA (Mercodia), and from 11.5 to 38 pmol/L.

<table>
<thead>
<tr>
<th>Manufacturer/Method</th>
<th>Median QC1, pmol/L</th>
<th>Median QC2, pmol/L</th>
<th>Median QC3, pmol/L</th>
<th>Median serum, pmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbott</td>
<td>106.5</td>
<td>288.5</td>
<td>686.3</td>
<td>112.2</td>
</tr>
<tr>
<td>AxSYM Insulin</td>
<td>27</td>
<td>66</td>
<td>21</td>
<td>149.1</td>
</tr>
<tr>
<td>Bayer</td>
<td>100.4</td>
<td>262.6</td>
<td>687.1</td>
<td>149.1</td>
</tr>
<tr>
<td>ADVIA Centaur Insulin</td>
<td>7.3</td>
<td>3.8</td>
<td>5.7</td>
<td>57.6–190.2</td>
</tr>
<tr>
<td>Biosource-Invitrogen</td>
<td>83.0</td>
<td>187.8</td>
<td>463.8</td>
<td>133.0</td>
</tr>
<tr>
<td>Insulin EASIA</td>
<td>3.6</td>
<td>6.1</td>
<td>7.1</td>
<td>89.0–205.0</td>
</tr>
<tr>
<td>Dakocytomation</td>
<td>66.4</td>
<td>195.7</td>
<td>450.8</td>
<td>82.9</td>
</tr>
<tr>
<td>Insulin assay</td>
<td>12.1</td>
<td>4.1</td>
<td>3.9</td>
<td>50.4–145.5</td>
</tr>
<tr>
<td>DPC</td>
<td>35.9</td>
<td>83.8</td>
<td>167.4</td>
<td>105.0</td>
</tr>
<tr>
<td>Immulite 1000</td>
<td>6.9</td>
<td>3.9</td>
<td>3.3</td>
<td>63.0–164.4</td>
</tr>
<tr>
<td>Linco (US)</td>
<td>97.0</td>
<td>280.0</td>
<td>709.8</td>
<td>71.4</td>
</tr>
<tr>
<td>ELISA</td>
<td>5.8</td>
<td>7.7</td>
<td>5.7</td>
<td>40.8–133.2</td>
</tr>
<tr>
<td>Linco (US)</td>
<td>75.8</td>
<td>242.2</td>
<td>621.3</td>
<td>141.3</td>
</tr>
<tr>
<td>RIA</td>
<td>4.9</td>
<td>3.2</td>
<td>5.3</td>
<td>81.0–222.0</td>
</tr>
<tr>
<td>Mercodia</td>
<td>91.8</td>
<td>291.1</td>
<td>704.1</td>
<td>106.8</td>
</tr>
<tr>
<td>Iso-Insulin assay</td>
<td>2.7</td>
<td>2.1</td>
<td>6.8</td>
<td>51.6–189.9</td>
</tr>
<tr>
<td>Molecular Light Technology</td>
<td>71.4</td>
<td>207.9</td>
<td>499.9</td>
<td>83.8</td>
</tr>
<tr>
<td>Roche</td>
<td>76.9</td>
<td>227.7</td>
<td>555.8</td>
<td>46.7–152.4</td>
</tr>
<tr>
<td>Insulin E170</td>
<td>1.3</td>
<td>1.7</td>
<td>1.1</td>
<td>55.6–186.5</td>
</tr>
<tr>
<td>Tosoh</td>
<td>76.9</td>
<td>226.7</td>
<td>529.8</td>
<td>84.0</td>
</tr>
<tr>
<td>Tosoh ST AIA -Pack IRI</td>
<td>9.4</td>
<td>2.9</td>
<td>4.2</td>
<td>47.4–151.8</td>
</tr>
<tr>
<td>Tosoh (US)</td>
<td>64.5</td>
<td>220.4</td>
<td>523.8</td>
<td>78.9</td>
</tr>
<tr>
<td>As above</td>
<td>3.1</td>
<td>2.6</td>
<td>1.9</td>
<td>37.0–148.9</td>
</tr>
</tbody>
</table>

Mean for QCs over 5 days, n = 5. Mean for samples over 5 days, n = 150. Mercodia assay cross-reacts 54% with proinsulin. QC Lyphochek Immunoassay Plus Control - QC1 Level 1 Cat No 371/Lot No 40151, QC2 Level 2 Cat No 372/Lot No 40152, QC3 Level 3 Cat No 373/Lot No 40153.
pmol/L for the others. Spearman correlation coefficients ($r^2$) vs the mean ranged from 0.983 to 0.997 (Table 3). Similar results were obtained when serum was run on the same automated assay (Tosoh) in 2 laboratories (Table 2). The results fell on or around the line of identity, with a mean difference of 1.92 pmol/L (95% CI 0.12 to 3.97), $P = 0.0647$, (Fig. 2). Total proinsulin ranged from 2 to 168 pmol/L with median (IQ range) 23 (12–43) pmol/L.

DIFFERENCE PLOTS
The modified difference plots (Fig. 3) and usual difference plots (see Figs. in the online Data Supplement) showed constant differences or differences that were proportional to the concentration. The difference plots for one assay vs another, available electronically, validate the use of the mean value for a common x-axis and will be of interest to laboratories running a particular assay/s.

The slope and intercept for differences vs the mean are presented in Table 3 as linear regression equations A, and for insulin values vs mean (scattergrams not shown) in Table 3 as B. Marked differences from the means were found for manual assays from Linco ELISA and Dako, with slopes of −0.32 and −0.20 (Table 3A), with the Linco RIA giving higher results, slope 0.19. The Biosource-Invitrogen assay that did not report results <38 pmol/L showed high concentration-dependent differences, as did the Mercodia ELISA that was reported to cross-react with total proinsulin, whereas the other 2 ELISAs had negative concentration-dependent differences. The MLT assay showed an offset with an intercept of −17 pmol/L. Of the automated assays, Abbott AxSYM and Bayer Advia Centaur showed positive, concentration-dependent differences (N.B., conversion factors >6.0), the DPC Immulite 1000 showed negative differences, and the Roche Elecsys E170 and Tosoh ST AIA had results closest to the mean of the assays.

Discussion
Since the ADA report on insulin standardization in 1996 (1), specific insulin assays have been introduced that incorporate nonradioisotopic labels and are suitable for automated platforms. In this comparison of 11 insulin assays, values measured for serum varied by a factor of 2, as reported earlier (1). However, progress with an international reference method and preparation has been slow, although recently mass spectrometric methods for insulin have been described (23).

Specificity of assays, calibration procedures, specimen type, assay performance, and conversion factors may all contribute to interassay variation. In this study, total proinsulin concentrations were in accord with previously published reports of 10% (proinsulin:insulin) for samples from persons without diabetes and 20% for patients with type 2 diabetes (24), with no samples displaying inappropriate hyperproinsulinemia (25). All assays except 1 were reported by manufacturers to be specific for insulin, and 1 was a specific RIA (Linco) that was included in the previous ADA comparison. The TOSOH assay, run in the US and UK for this study, produced similar results in the different laboratories, in contrast to the previous ADA comparison. In US epidemiological studies this assay was chosen recently to replace the nonspecific Pharmacia RIA100, which had been withdrawn because antibody stocks were exhausted. The nonspecific assay (Mercodia Iso-Insulin) ranked in the middle of the distribution of the assays in this study. This manufacturer, Mercodia, also produces a specific insulin assay with no detectable cross-reactivity with proinsulin and does 31–21 and 32–33 split proinsulins.

All manufacturers reported that their assays were calibrated against the same insulin reference preparation (IRP 66/304) prepared from human pancreas with a stated number of units per ampule. Details of manufacturers’ calibration procedures and calibrants (26) were not
Table 3. Mean differences for insulin for each assay from the mean of all 12 assays for 150 serum samples, linear
regression equations $A$ differences $A_x$ from mean values vs mean insulin and $B$ actual value $y_x$ vs mean insulin, and
correlation coefficients

<table>
<thead>
<tr>
<th>Manufacturer/ Mean Method difference (95% limits) pmol/L</th>
<th>Linear regression equations $r^2$</th>
<th>$A$</th>
<th>$B$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slope</td>
<td>Intercept</td>
</tr>
<tr>
<td>Abbott AxSYM Insulin</td>
<td></td>
<td>$+23\ (-64\ to\ +110) +0.22^c$</td>
<td>$-12^c$</td>
</tr>
<tr>
<td>Bayer ADVIA Centaur Insulin</td>
<td></td>
<td>$+21\ (-58\ to\ +99) +0.19^c$</td>
<td>$+8.2^c$</td>
</tr>
<tr>
<td>Biosource-Invitrogen Insulin EASIa</td>
<td></td>
<td>$+37\ (-23\ to\ +97) +0.11^c$</td>
<td>$+20.4^c$</td>
</tr>
<tr>
<td>Dakocytomation Insulin assay</td>
<td></td>
<td>$-27\ (-112\ to\ +57) -0.20^c$</td>
<td>$+2.9^{NS}$</td>
</tr>
<tr>
<td>DPC Immulite 1000</td>
<td></td>
<td>$-10\ (-64\ to\ +45) -0.11^c$</td>
<td>$-4.5^c$</td>
</tr>
<tr>
<td>Linco (US) ELISA</td>
<td></td>
<td>$-47\ (-165\ to\ +71) -0.32^c$</td>
<td>$-2.6^{NS}$</td>
</tr>
<tr>
<td>Linco (US) RIA</td>
<td></td>
<td>$+47\ (-36\ to\ +131) +0.19^c$</td>
<td>$+18.1^c$</td>
</tr>
<tr>
<td>Mercodia Iso-Insulin assay</td>
<td></td>
<td>$+10\ (-48\ to\ +67) +0.07^c$</td>
<td>$-1.2^{NS}$</td>
</tr>
<tr>
<td>Molecular Light Technology</td>
<td></td>
<td>$-17\ (-49\ to\ +15) 0^{NS}$</td>
<td>$-17.0^c$</td>
</tr>
<tr>
<td>Roche Insulin E170</td>
<td></td>
<td>$+9\ (-25\ to\ +42) +0.07^c$</td>
<td>$-1.6^{NS}$</td>
</tr>
<tr>
<td>Tosoh Tosoh ST AIA Pack IRI</td>
<td></td>
<td>$-22\ (-72\ to\ +28) -0.13^c$</td>
<td>$-1.9^{NS}$</td>
</tr>
<tr>
<td>Tosoh (US) as above</td>
<td></td>
<td>$-24\ (-62\ to\ +15) -0.09^c$</td>
<td>$-10.5^c$</td>
</tr>
</tbody>
</table>

$^{NS}$ not significant $^a P < 0.05$; $^b P < 0.01$ $^c P < 0.001$.  

requested by the authors. Mean (range) recovery of IRP 66/304 in 14 methods has been reported to be 89% (75% to
127%), (27). Some assays showed overrecovery (Abbott AxSYM 127%), some underrecovery (DPC Immulite 75%),
and some had recoveries within 10% (Bayer Centaur 100%, Dako 91%, Mercodia ELISA 99%, MLT 98%, and
Roche Elecsys 103%). It remains to be seen whether the use of a human recombinant insulin preparation, cali-
brated in molar and mass terms, will reduce the interas-
say variability.  

All participating laboratories measured insulin in se-
rum, although heparinized plasma was also available. For il-
lustrative purposes, both sample types were compared in 1 assay, and in that case higher insulin values were
obtained in serum. Because these differences may be assay specific, individual laboratories should confirm
with the assay manufacturer the procedures recom-
meded for specimen collection and compare results for
plasma and serum as required. In this study, however, the
conversion factors used by the manufacturers were found
to vary considerably, with the potential to influence insu-
nol values by up to 15%. For 3 of 4 specific assays
using conversion factors $>6$, reported mean values were
in the top half of the distribution, and the other value was
in the middle. Assays using a factor of 6 ranked mainly in
the bottom half, except for the specific RIA (Linco) that gave

Five assays were run on automated analyzers, and no
significant difference in imprecision was observed be-
tween automated and manual assays. Bias and imprec-
sion for lyophilized QCs in the individual assays did not
always mirror those seen in serum (Table 2), an important
point for external quality assessment schemes. The use of
nenhuman insulin in QCs and matrix effects may explain
these differences. Although the security of baseline or low
concentrations is important (28), sensitivity, as reported
by the manufacturers, was determined in a number of
ways that were not comparable. Use of the minimum
detectable concentration from precision profiles may yield
more reliable estimates.

A conversion factor of 6 for recombinant human insu-
lin has been reported (1, 29). In this study, however, the
conversion factors used by the manufacturers were found
to vary considerably, with the potential to influence insulin values by up to 15%. For 3 of 4 specific assays
using conversion factors $>6$, reported mean values were
in the top half of the distribution, and the other value was
in the middle. Assays using a factor of 6 ranked mainly in
the bottom half, except for the specific RIA (Linco) that gave
the highest results. Interestingly, an ELISA from the same company, also using 6 as a factor, gave the lowest results.

Because of potential assay interference, our study has limitations that prohibit the extrapolation of results to patients with insulin antibodies. The results are also not applicable to patients treated with insulin analogs, because these cannot be detected with some specific insulin assays (11–14). A limited number of batches of reagent and calibrators were used in this study. Because there is no reference method for insulin or consensus on standard statistical techniques for such multiassay analyses (30), the mean insulin value for each sample was used as a common comparator. This approach may not be appropriate for a smaller number of assays.

Various techniques have been used over the years for comparing measurements from different studies and for reducing assay variation. For major research studies, insulin has been measured in central laboratories (3). In some studies data have been divided into ranks, but this method may not be reliable because populations differ (31), and it is not helpful when considering individuals. Results for other analytes such as hemoglobin A1c, have been aligned to those reported by major clinical trials while the international reference method was being evaluated (32). In the case of human chorionic gonadotropin, clearly defined calibration procedures for the international standard and careful selection of antibodies reduced assay variation (33). Intermethod differences of 2- to 5-fold in prostate-specific antigen measurement results (34) were eliminated by the introduction of standard calibrators and guidelines, but differences in the recovery of international standards on automated platforms still varied by −5% to +22%. These results may be related to data transfer between the primary standards and master calibrator, and between master calibrator and individual reagent set calibrators, a potential issue for all analytes.

International diabetes and clinical chemistry organizations should be aware of the need for comparability of insulin values for patient care, clinical research studies, and any resulting metaanalyses or clinical research networks (35). Clinicians and researchers should be made aware of the fact that serum insulin measurements from the various assays now available can differ by a factor of 2.

Grant/funding support:
Financial disclosures: Diagnostic companies for insulin kits and some despatch costs: Abbott (C. Thompson); Bayer (A. Thompson); DakoCytomation (H. Masters, C. Smellie); DPC (C. Field, J. Hayes); Linco US (R. Gingerich, J. Mistry, A. Tunnicliffe); Biosource-Invitrogen, previously Medgenix (M. Salmon); Mercodia (P. Atkinson); M.L.T. (H Jeffreys, S Woodhead); Roche (A D'Souza); Tosoh UK and US (S. Marivoet).
Sponsor for centers’ meeting – Mr. M Fineman, Amylin Pharmaceuticals, San Diego, California.
Dr. P.M. Clark is a member of the Specialist Advisory Group for Immunoassay in UKNEQAS for which she receives no remuneration.

Acknowledgements: We would like to thank Diabetes UK for their grant, Professor Doug Altman, University of Oxford, for advising on statistical issues, Dr Gwen Wark, UK NEQAS Scheme Organiser, for EQA data supplied with permission from the EQA scheme, Mrs Rachel Round for administration, and Professor Stephen Gough, Miss Janet Smith, Dr Malcolm Nattrass from University Hospital Birmingham NHS Foundation Trust and Dr Tim Barrett from Birmingham Children’s Hospital NHS Trust for their helpful comments and support.
Fig. 3. Difference plots for human insulin in 150 serum samples measured by 12 assays in UK or US, on left 0–150 pmol/L and right 0–1500 pmol/L.

Difference plot (solid line) zero line, (dashed line) mean difference, 2 SD lines.
Fig. 3. Continued
Fig. 3. Continued
Also the following in the UK: A Powell, P Griffiths, (Birmingham Children’s Hospital); R Round; M Palma, M Peyton, J Babrah (Clinical Biochemistry, University Hospital Birmingham); R Williams, (Derriﬀord Hospital, Plymouth); G Dunseath, A Clark (Diabetes Research Unit, Cardiff University); G Wark (Royal Surrey County Hospital, Guildford); R D’Souza, R Elms, (St George’s Hospital, Tooting); S Fell, B Starling (Walsall Manor Hospital): in the US: H-M Wiedmeyer, A Tennill (Diabetes Diagnostic Laboratory, Columbia, University of Missouri Health Care); J Mistry (Linco Research Inc, St Charles, Missouri).

References
27. Wark, G. UKNEQAS Guildford Peptide Hormones. Annual Review 2005. Available from UK NEQAS Guildford Peptide Hormones Scheme, Clinical Laboratory, Royal Surrey County Hospital, Guildford, Surrey GU2 7XX, UK.