Significant, quantifiable differences exist between IgG subclass standards WHO67/97 and ERM-DA470k and can result in different interpretation of results

Calum Wilson, Rosemary Ebling, Clara Henig, Tanya Adler, Roza Nicolaevski, Mira Barak, John Cazabon, Diane Maisin, Thibault Lepoutre, Damien Gruson, Richard G. Hughes, Antony R. Parker

Abstract

Objectives: Accurate measurement of IgG subclass (IgGSc) levels are essential to aid in the diagnosis of disease states such as primary immunodeficiencies. However, there is no single standardisation of nephelometric and turbidimetric assays for these analytes and two reference materials have been utilised. We expand on previous reports and present data from a multi-site analysis that both identifies and quantitatively defines the differences in calibration resulting from the use of different reference materials.

Design and methods: IgGSc antibodies in the serum specimens and reference materials were measured according to the manufacturers’ instructions using commercially available IgGSc assays or components.

Results: Data from four independent sites showed that in spite of the different commercial suppliers of IgGSc assays calibrating to different reference materials, ERM-DA470k and WHO67/97, the resulting calibrations were comparable for IgG1 and IgG2. However, for IgG3 and IgG4 the calibrations were significantly different. The use of assay specific normal ranges should compensate for these calibration differences, however, the two manufacturers’ assays can give differing clinical classifications. The agreement between the different manufacturers’ IgGSc assays was between 85.1% and 95.8% for all IgGSc assays, the discordance of sample classifications for IgG1 and IgG2 assays was approximately 12% and 15% respectively, whilst that for IgG3 and IgG4 was 4% and 13% respectively.

Conclusion: We discuss the similarities and differences between assays that utilise the different reference materials.

© 2013 The Authors. The Canadian Society of Clinical Chemists. Published by Elsevier Inc. All rights reserved.
(Table 1). Intra-assay precision and sample linearity of TBS and Siemens IgGSc assays.

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Function</th>
<th>IgG1</th>
<th>IgG2</th>
<th>IgG3</th>
<th>IgG4</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBS</td>
<td>% CV</td>
<td>3.3</td>
<td>1.8</td>
<td>3.6</td>
<td>3.1</td>
</tr>
<tr>
<td>mg/L</td>
<td>Linearity:</td>
<td>75%</td>
<td>-1.9</td>
<td>1.8</td>
<td>-3.7</td>
</tr>
<tr>
<td>% Deviation from expected value</td>
<td>25%</td>
<td>-1.1</td>
<td>-5.0</td>
<td>3.3</td>
<td>-7.0</td>
</tr>
<tr>
<td>Siemens</td>
<td>% CV</td>
<td>1.6</td>
<td>1.9</td>
<td>5.0</td>
<td>2.9</td>
</tr>
<tr>
<td>mg/L</td>
<td>Linearity:</td>
<td>75%</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-2.2</td>
</tr>
<tr>
<td>% Deviation from expected value</td>
<td>25%</td>
<td>16.6</td>
<td>0.3</td>
<td>4.5</td>
<td>-0.7</td>
</tr>
</tbody>
</table>

(CRM470; now superseded with ERM-DAn470k due to depletion of CRM470 stocks [7]) [9] and in the case of Sanquin and Siemens WHO67/97 [10] (later replaced by the commercial calibration material Sanquin M1590).

Bossuyt et al. [11] have previously reported the differences in calibration and thus data interpretation between TBS IgGSc assays and the Sanquin IgGSc assays. A commentary has recently been published highlighting the major difference between TBS and Siemens IgGSc assay calibration and concluded that this difference was due to the assays being standardised against two different reference materials [12].

Here, we expand on previous reports and present data from a multisite analysis that both identifies and quantitatively defines the differences in calibration resulting from the use of different reference materials. Furthermore, the subsequent effect this has on classification of patient samples is also presented.

Table 1: Intra-assay precision and sample linearity of TBS and Siemens IgGSc assays.

For intra-assay precision: twenty seven replicates were assayed for each sample in each of the manufacturer’s individual IgGSc assays and percent coefficient of variation (%CV) was calculated.

For linearity: the samples were diluted to 75%, 50% and 25% of its original concentration and the values of the dilution recorded. The percentage deviation from the expected value was determined as described in materials and methods. All individual dilutions were assayed in triplicate.

Materials and methods

Assay method

IgGSc antibodies in the serum specimens and reference materials were measured according to the manufacturers’ instructions using commercially available IgGSc assays: SPaPLUS IgGSc assays (IgG1–IgG4; NK006.S, NK007.S, LK008.S, LK009.S; The Binding Site, UK). Siemens BNII IgGSc assays were performed with the following components: N AS IgG1 (OQXI092), N AS IgG2 (OQXK092), N Latex IgG3 (OPAV032), and N Latex IgG4 (OPAU032), N-supplementary reagent (QQTD115), Siemens Cleaner SCS (QQU195), and the N protein standard SL (OQIM135) (Siemens Healthcare Diagnostic Products, Germany).

The adult normal ranges stated for the TBS assays are IgG1: 3.82–9.29 g/L, IgG2: 2.42–7.00 g/L, IgG3: 0.22–1.76 g/L, IgG4: 0.04–0.87 g/L, and for the Siemens assays IgG1: 4.1–10.1 g/L, IgG2: 1.7–7.9 g/L, IgG3: 0.11–0.85 g/L, IgG4: 0.03–2.0 g/L.

Precision

The precision of each assay was compared by running twenty seven replicates of the same sample on each assay on both the Siemens BNII and the TBS SPaPLUS IgGSc assays. The sample consisted of pooled human serum with IgGSc levels within the standard measuring range for each assay on both manufacturers’ assays.

Linearity

A serum sample was identified that gave a readable concentration towards the upper value of the measuring range for each IgGSc on both manufacturer’s assays. Dilutions of the samples were prepared at 75%, 50% and 25% concentration of the original fluid. The linearity of the IgGSc assays was assessed by running each dilution in triplicate and comparing the mean result to the expected results. The
expected concentrations were calculated as a percentage of the concentration obtained at 100% (no dilution). The % deviation from linearity was calculated as ((concentration obtained/concentration expected)*100).

**Method comparison**

Randomly selected human serum samples were collected at each of the four independent sites and measured on each of the manufacturers' IgGSc assays.

**Statistical analysis**

Pearson correlation, Altman–Bland analysis, and Deming regression analysis were all performed using Analyse-It software. Percent coefficient of variation was calculated in Excel using the following formula: ((Standard deviation/mean)*100).

**Table 2**

IgG subclass comparison characteristics between TBS and Siemens IgGSc assays.

Human serum samples were assayed in all four subclass assays from both manufacturers at the four independent sites. The data were analysed to quantitatively measure the correlation and agreement between the appropriate subclass assays.

<table>
<thead>
<tr>
<th>Component</th>
<th>TBS IgG1 (n=30)</th>
<th>Siemens IgG1 (n=30)</th>
<th>TBS IgG2 (n=25)</th>
<th>Siemens IgG2 (n=30)</th>
<th>TBS IgG3 (n=40)</th>
<th>Siemens IgG3 (n=34)</th>
<th>TBS IgG4 (n=30)</th>
<th>Siemens IgG4 (n=34)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson correlation</td>
<td>0.97</td>
<td>0.96</td>
<td>0.92</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>0.94-0.99</td>
<td>0.93-1.00</td>
<td>0.81-0.96</td>
<td>0.92-0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altman–Bland: bias</td>
<td>-1699.3</td>
<td>537.9</td>
<td>588.2</td>
<td>-136.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>-1935.2 to -1463.5</td>
<td>444.80 to 631.1</td>
<td>460.35 to 712.6</td>
<td>-237.4 to -15.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deming regression: intercept</td>
<td>25.08</td>
<td>110.82</td>
<td>121.4</td>
<td>104.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>-965.73 to 1025.09</td>
<td>-75.7 to -297.35</td>
<td>-351.05 to 108.25</td>
<td>-102.01 to 311.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deming regression: slope</td>
<td>0.81</td>
<td>1.13</td>
<td>2.18</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>0.70-0.93</td>
<td>1.07-1.19</td>
<td>1.70-2.63</td>
<td>0.2-1.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Results**

**Measurement of assay performance and assay comparison**

Intra-assy precision and sample linearity were calculated for each assay (Table 1) with one sample assayed per IgGsC. The intra-assay precision for all IgGSc assays were ≤5% with IgG3 subclass assays the least precise for each manufacturer. The linearity study showed that deviation from the expected values were <10% for all assays with the exception of the sample diluted to 25% of its original concentration in the Siemens IgG1 assay which showed a deviation of 16.6% from the expected value.

Results from the comparison studies using both manufacturers' assays are shown in Fig. 1 and Table 2. The correlation between the IgG1 and IgG2 values were good (Deming regression slope IgG1; 0.81–0.98 and IgG2; 0.98–1.13). Correlation of IgG3 and IgG4 however suggested significant differences existed between the values obtained on the different manufacturers’ assays (Deming regression slope IgG3; 1.53–2.18 and IgG4; 0.50–0.63).

**Quantitation of assay comparison**

The values obtained for both the IgG1 and IgG2 in ERM-DA470k were within 15% of each other when assayed with either TBS IgGSc or Siemens IgGSc assays (Fig. 2). The difference between the values obtained for IgG3 and IgG4 with the different manufacturers’ assays, however, was much greater and further supported the observation of substantial calibration differences between the two manufacturers’ assays. For IgG3 the percentage difference in values obtained was >40% (304 mg/L in the Siemens assay vs. 559 mg/L in the TBS assay) and >60% in the IgG4 assay (607 mg/L in the Siemens assay vs. 374 mg/L in the TBS assay).

**Effect of different calibrations on sample classification**

We investigated whether there were any differences in sample classification due to the calibration differences between the TBS and Siemens IgGSc assays, particularly the IgG3 and IgG4, for a range of samples measured at the four sites (Table 3).

The overall agreement between the different IgGSc assays, that is the percentage of samples that fell into the same classification on both assays, was >85% in all four IgGSc assays (85.1%–95.8%). Even though the linear regression analysis of the results from the IgG1 and IgG2 assays gave slopes suggesting reasonable agreement and certainly far closer agreement than between the IgG3 and IgG4 assays, the discordance of sample classification for both the IgG1 and IgG2 assays was

![Fig. 2. Plot of mean values for ERM-DA470k standard obtained on TBS IgGSc assays and Siemens IgGSc assays. There is a good agreement between the IgG1 and IgG2 assays. For IgG3 and IgG4 a significant difference between the manufacturers' assays is observed. ERM-DA470k was diluted according to the manufacturer's instructions and run three times in triplicate in all assays. The error bars represent three standard deviations.](image-url)
Discussion

The accurate and reproducible measurement of IgGSc is essential to aid the diagnosis of both hypo- and hyper-gammaglobulinaemias particularly in the presence of normal serum IgG levels. The concentration of IgGSc can vary with age, gender, race, laboratory to laboratory due to method differences as well as reference material and calibration differences. Bossuyt and colleagues previously noted that disparity exists between IgG3 and IgG4 assays calibrated with different materials [11]. They concluded that this was the direct result of standardising the assays with incongruent calibration reference materials. Our recent publication highlighted the differences in the paediatric reference intervals between both types of standardised assays and the use of different reference materials, WHO 67/97 and ERM-DA470k. We further concluded that the appropriate reference intervals must be applied to the appropriate assay in accordance with the individual manufacturer’s recommendations [12].

Deming regression analysis and quantitation of ERM-DA470k have suggested that although the calibrations were acceptable and comparable between the IgG1 and IgG2 assays but not between the IgG3 and IgG4 assays, all four paired IgGSc assays from the different manufacturers were statistically different. Calibration is not the only cause of differences between the manufacturers’ assays, the instrumentation used or the antisera may have an influence. We compared the Siemens data with data obtained using the Sanquin IgGSc assays for the IMMAGE, both assays utilise WHO67/97 as a reference material. Whilst the data obtained for IgG3 and IgG4 were more comparable than found with the other two manufacturers’ assays, the instrumentation used or the antiserum may have an influence. We compared the Siemens data with data obtained using the Sanquin IgGSc assays for the IMMAGE, both assays utilise WHO67/97 as a reference material. Whilst the data obtained for IgG3 and IgG4 were more comparable than found with the other two manufacturers’ assays, the instrumentation used or the antiserum may have an influence.

In summary we have provided detailed evidence that the calibration of IgG3 and IgG4 in this study but there were surprisingly large differences between sample classifications when measuring IgG1 and IgG2, harmonisation of reference ranges would also be required to minimise inter-laboratory differences.

Standardisation of immunoassays is necessary to minimise inter-laboratory variation and the implementation of CRM 470 for the major serum assays illustrates how this approach can be successful. A common reference material is required for standardisation and therefore ideally, a single internationally accepted reference material should be adopted for the IgGSc assays; in the absence of this an appropriate conversion factor between the different materials may be beneficial. However, manufacturer normal ranges are supplied with each assay which should in theory compensate for the impact of using different reference materials for their respective calibration. This was indeed the case for IgG3 and IgG4 in this study but there were surprisingly large differences between sample classifications when measuring IgG1 and IgG2, harmonisation of reference ranges would also be required to minimise inter-laboratory differences.

In summary we have provided detailed evidence that the calibration of IgG1 and IgG2 Siemens and Binding Site assays are similar but that the calibration of IgG3 and IgG4 assays are significantly different and we have provided a plausible explanation for this difference. For all IgGSc assays the reference ranges quoted in the manufacturers’ inserts in the early reference material WHO 67/97 and its suitability for nephelometric or turbidimetric assays. Whicher reported the precipitation of residual proteins from long term storage of WHO material [16] which has been shown to decrease the precision of nephelometric and turbidimetric assays. Klein and colleagues reported that the levels of IgC3 were unstable in WHO 67/97 and may be the reason for the laboratory to laboratory variation observed in the measurement of WHO 67/97 IgG3 concentrations obtained and the under reading of IgG3 levels [10]. It appears likely that the lower calibration of IgG3 adopted in the Siemens assay may be due to the degradation of IgG3 in the early reference preparation WHO 67/97 and this view is further supported by the many publications that have investigated the levels of the IgGSc (Table 4). With the exception of one study the results clearly indicate with considerable consistency that the concentration of IgG3 > IgG4 in human serum. The relative concentrations of IgGSc in CRM470 have been determined nephelometrically to be IgG1 > IgG2 > IgG3 > IgG4 [7] and are in agreement with published literature in Table 4. CRM470 (replaced by ERM-DA470k due to stock exhaustion) was accepted as the International Reference Preparation for proteins in human serum in 1992 and the long term stability of the analytes including IgG has been shown. However, although the IgG1 and IgG2 assays have similar calibrations and give comparable values on patient samples, the normal ranges given within the manufacturers’ inserts differ significantly. When these are applied this leads to classification discordance of up to 15% between the patient samples assessed in this study and thus emphasises the requirement to use manufacturer specific normal ranges and the lack of assay interchangability.

Table 3

Sample classification in the TBS and Siemens IgGSc assays. Randomly selected human serum samples were run at four independent locations (IgG1, n = 121; IgG2 n = 167; IgG3, n = 141; IgG4, n = 121) on all four IgGSc assay systems. The resulting sample values were classified using the normal ranges in each manufacturer’s insert. (1) low, below the stated normal range; (2) normal, within the stated normal range and (3) high, higher than the stated normal range. The % agreement between each subclass assay from the two manufacturer’s was calculated by dividing the number of samples that showed agreement in classification by the total number of samples assayed (shaded boxes).

Table 4

Publication search of IgGSc percentage composition in the absence of external reference material. Publications in which the relative levels of IgGSc have been investigated. Column 1 lists the references, column 2 states the subclass method used, column 3 states the sample types and column 4 summarises the relative proportions of each IgGSc as a percentage.
are different and this leads to differences in sample classification particularly for IgG1, IgG2 and IgG4 the assays and reference ranges should not be used interchangeably.

References