Introduction

Comparing against peers has been a common practice in proficiency testing (PT) for clinical field laboratories in Mexico, specifically by using the participants’ consensus as the assigned value. However this practice introduces unbounded risks and consequences for the final users of these results. In order to show the benefits of using an assigned value with a higher metrological level a PT for clinical laboratories using certified reference materials (CRM) was conducted by CENAM between 2002 and 2006 [1,2,3]. The results of this program are reviewed in this presentation.

Method

The program’s original scope was to analyze five clinical key markers (calcium, glucose, cholesterol, creatinine, uric acid) in human serum. All participants received one sample and were asked to obtain 10 independent measurements during 2002 and 2004 independent measurements thereafter. Participants were not asked to estimate measurement uncertainty, uncertainty due to repeatability conditions may be underestimated by the sample standard deviation. Only three measurands were selected for this study, the measurands’ certified concentration value varied in the following ranges: glucose: 82.9 – 90.82 mg/dl, cholesterol: 156.1 – 161.18 mg/dl, creatinine: 0.74 – 0.75 mg/dl. All reference measurements were produced and certified by CENAM. CENAM used Isotopic Dilution method considered a potential primary method. The relative standard uncertainty [4] of the certified values ranges from 0.5% to 4%. The participants’ consensus was obtained by using non-robust (mean and standard deviation) and robust statistics [5] (median and scaled mad). The robust estimate avoids any criticism about declaring some participants as outliers. The assigned value obtained by the participants’ consensus was compared against the certified reference value and the zeta (ζ) score was obtained for each one in both cases. The results based on robust statistics are shown in Figure 1, the results based on non-robust statistics are shown in Figure 2.

Simulation data was fitted with a one way random effects model in order to estimate the between PT exercises variance for each simulated participant. This component of uncertainty σ2mpt accounts for intermediate measurement precision [4,8,10] and is included in the uncertainty budget for each participant. The distributions of the simulations results are tested under two conditions: (a) using all the simulated data and (b) using the simulated data filtered for large variability (>20%) and bias.

Results

Modified target plot [11] with the simulation results were prepared for each measurand. Target plots were modified in two aspects:

i. The plotted quantities are relative to the reference value, in this order we use the percentage differences [7],

\[ D_n = 100(x - X_{CRM})/X_{CRM} \]

ii. The acceptance region is in arc-shape. This arc-shaped region is approximately the region with higher probability for the true joint distribution of both parameters under normality and independence assumptions [11]. The diagonal lines satisfy the equation

\[ D_1 = \Phi^{-1}(0.975) \cdot \text{bias} \]

iii. The uncertainty of the percentage difference was estimated as

\[ \sigma_{D_1} = \sqrt{\frac{\text{bias}^2}{2} + \frac{100 \cdot \text{bias}^2}{X_{CRM}^2}} \]

Figure 3 shows the modified target plots for glucose and creatinine.

Discussion

For glucose about 8% of the population of laboratories appears to have biased results, for cholesterol about 2% of the population of laboratories and for creatinine about 44%. Table 1 summarizes the distribution of the uncertainty of the unbiased results for each measurand.

Table 1. Summary of the distribution of estimated uncertainty excluding simulation results suspected as biased. The use of CRM warrants the contribution to these uncertainties is less than 4%.

<table>
<thead>
<tr>
<th>Measurand</th>
<th>Uncertainty of the percentage difference</th>
<th>% results Glucose</th>
<th>% results Cholesterol</th>
<th>% results Creatinine</th>
</tr>
</thead>
<tbody>
<tr>
<td>glucose</td>
<td>20 ≤ σ&lt;sub&gt;D_1&lt;/sub&gt; ≤ 50</td>
<td>20</td>
<td>5</td>
<td>38</td>
</tr>
<tr>
<td>cholesterol</td>
<td>10 ≤ σ&lt;sub&gt;D_1&lt;/sub&gt; ≤ 15</td>
<td>10</td>
<td>15</td>
<td>76</td>
</tr>
<tr>
<td>creatinine</td>
<td>0 ≤ σ&lt;sub&gt;D_1&lt;/sub&gt; ≤ 10</td>
<td>64</td>
<td>32</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 4 shows a normal quantile-quantile plot of the area (Q) scores for glucose of the simulation study under condition (b). Remarkably there is no evidence of non-normality although this distributional property had not been induced by the applied non-parametric resampling techniques.

Conclusions

About 60% of the assigned values by the participants’ consensus appear to be significantly biased. The observed bias is confounded between the field laboratory method and the commercially available IVD kits used for the measurements hence the estimated bias may be method dependent. Getting unbiased creatinine measurements is a clear opportunity for improvement with straight implications for the population’s health, rather for diagnostics or treatment. The uncertainty of the participants cannot be explained by the argument of biological variability [9]. Although this improvement opportunity was intended for the field laboratories it may be fixed by the IVD kit producers. The use of CRM is justified for PT schemes and traceability recommendations of JCTLM to IVD kit producers in order to warrant traceability of their calibrants.

Acknowledgements

Thanks to the technical staff of the Division of Materials Metrology, who experiments and professionalism made possible this data: Esther Castro Gabino, Mayra Beldawas Escamilla, Laura Regalado Contreras, Judith Rivera Mellado, Gabriela Salazar Briones, Mauricio Malدونado Torres, Marco Antonio Ávila Calderón, Melina Pérez Usoiqui. Special thanks to the referee, whom critical feedback improved this work, Dr. Yoshito Mitani and Dr. Mariana Arce-Osuna for their enthusiastic support and critical feedback.

References