Standardization of ceruloplasmin measurements is still an issue despite the availability of a common reference material

Ilenia Infusino · Cristina Valente · Alberto Dolci · Mauro Panteghini

Received: 14 September 2009 / Revised: 15 October 2009 / Accepted: 16 October 2009 / Published online: 17 November 2009 © Springer-Verlag 2009

Abstract The purpose of measurement standardization is to achieve closer comparability of results obtained using different commercial systems. Regarding serum protein immunoassays, a reference preparation (BCR-470) was released in 1993 and adopted by manufacturers across the world to value-assign their assay calibrators for routine methods to reduce method-dependent variation. Moving from nephelometric (Beckman Immage 800) to turbidimetric determination (Roche Cobas c 501) of seven serum proteins, we preliminarily checked the comparability of results between the two systems. The study was performed according to the CLSI EP9-A protocol on 30 fresh sera, tested on each system in duplicate, and subdivided on two different days, without recalibration and using manufacturers’ control materials to validate the runs. Both manufacturers’ package inserts provide statements that kit calibrators are traceable to BCR-470. Suggested reference intervals are also the same. Although a fairly good correlation was observed ($r=0.955$), the comparison of ceruloplasmin methods produced evidence of highly significant proportional (regression slope, 0.572) and constant bias (intercept, 0.05 g/L). Absolute and percentage mean differences were $-0.11$ g/L (95% confidence interval (CI) $-0.13$ to $-0.10$ g/L) and $-39.1\%$ (CI $-43.1$ to $-35.2\%$), respectively. No other evaluated proteins showed similar problems. Lacking a ceruloplasmin reference method, it is impossible to demonstrate that one of the two assays produces true ceruloplasmin values. The problem is, however, that results coming from the two assays are clearly not comparable. This may be either due to a lack of commutability of the reference material with biological samples in the evaluated assays or to calibration problems by manufacturers in one of the stages of the calibration hierarchy.

Keywords Ceruloplasmin · Nephelometry · Immunologic techniques · Standardization · Reference materials · Biological samples

Introduction

Foremost among the clinical laboratory’s problems is the poor comparability of results that originate from different methods or platforms measuring the same analyte. This may create confusion among clinicians when interpreting results. Therefore, the standardization of measurements is of high priority in laboratory medicine, its purpose being to achieve closer comparability of results obtained using different commercial systems [1, 2]. Particularly, the promotion of result traceability to available higher-order reference measurement procedures and reference materials is the recommended approach [3].

A major prerequisite for guaranteeing comparability of results among different methods is the availability of suitable reference materials, appropriately and thoroughly
defined by a set of characteristics [4]. Reference materials can be used for calibration of routine methods, but when reference materials are intended for direct value assignment to manufacturers’ calibrators, they should be extensively investigated for commutability [5]. Commutability has been defined as the ability of a reference/control material for a given analyte to show interassay properties comparable to those of the same analyte in human serum [6].

Regarding serum protein immunoassays, a reference preparation (BCR-470) was released in 1993 and adopted by manufacturers across the world to value-assign their assay calibrators for routine methods to reduce method-dependent variation [7, 8]. Indeed, the introduction of the BCR-470 reference material has resulted in highly significant reduction of the among-laboratory variance for most proteins [9, 10]. With these premises, before switching testing of seven serum proteins from nephelometric to turbidimetric assays, we recently checked the comparability of results between the two methods to ensure that nothing could influence the clinical application of the measurements. In this paper we describe the results and, in particular, the problems associated with ceruloplasmin measurement.

Materials and methods

The evaluated serum proteins, i.e., α1-acid glycoprotein, α1-antitrypsin, ceruloplasmin, C-reactive protein, haptoglobin, transferrin, and transthyretin, were assayed by rate nephelometry on Immage 800 analyzer (Beckman Coulter, Fullerton, CA, USA) and by direct turbidimetry on the Cobas c 501 platform (Roche Diagnostics, Basel, Switzerland), a fully automated clinical chemistry analyzer, in accordance with the manufacturers’ instructions. Both manufacturers’ package inserts provide statements that kit calibrators for the seven proteins are traceable to the reference preparation BCR-470 (Institute for Reference Materials and Measurements (IRMM), Geel, Belgium). Reference intervals for the seven proteins suggested by the two manufacturers are also the same.

The study was performed according to the CLSI EP9-A protocol on a total of 30 unselected fresh sera tested on each system shortly after collection in duplicate, in two runs (15 samples per run) performed on two following days without recalibration and using the manufacturers’ control materials to validate the analytical runs [11]. Delinked and anonymous leftover patient specimens submitted for routine testing were used.

Because none of the evaluated assays was considered the reference method, we used the Deming regression to calculate analytical correlations. Between-assay biases were also obtained and compared with the desirable bias for the clinical application of the measurements derived from biological variability (BV) of each protein [12]. Desirable bias was calculated as 0.25 [(intraindividual BV)² + (interindividual BV)²]^{1/2} [13].

Results

The correlation plots for the tested proteins, with the exclusion of ceruloplasmin, are shown in Fig. 1. Table 1 summarizes the regression parameter estimates of comparison data for all of the evaluated serum proteins, together with experimental and desirable biases. With the exception of ceruloplasmin, all correlation coefficients (r) exceeded 0.98. However, for some proteins there were highly significant proportional (haptoglobin) or constant (α1-antitrypsin) biases. C-Reactive protein also displayed a statistically significant slightly negative intercept, but it was compensated by the slightly positive slope. Quite surprisingly, the comparison of ceruloplasmin methods produced evidence of highly significant bias between the two evaluated assays, showing a regression slope of 0.572 and an intercept of 0.05 g/L, although a fairly good correlation was observed (r = 0.955). Absolute and percentage mean differences for ceruloplasmin values were −0.11 g/L (95% confidence interval (CI) −0.13 to −0.10 g/L) and −39.1% (95% CI −43.1 to −35.2%), respectively (Fig. 2). No other evaluated proteins showed so huge comparability problems, even if three of them (transferrin, transthyretin, and α1-antitrypsin) displayed a mean experimental bias that was higher than the desirable target (Table 1). As can be seen from the figures, no data points that might represent “outliers” when measured by the two methods were present. Considering the desirable bias, optimal comparability was only obtained for α1-acid glycoprotein and C-reactive protein.

Discussion

The use of protein reference material BCR-470 is intended to lead to reduced method-dependent variation in specific protein analyses. Observations from the College of American Pathologists surveys and several External Quality Assessment Schemes (EQAS) for specific proteins in Europe have indicated that this is true for most proteins [9, 10]. However, based on our correlation data (slope, intercept, mean bias, or more than one of these parameters), result comparability for the majority of the proteins tested in this study is not as good as required for their clinical application, suggesting that further improvement in method standardization is still beneficial. Particularly, results from Beckman Immage nephelometer and Roche Cobas platform...
for α₁-antitrypsin and ceruloplasmin do not resemble each other, and the use of different method-specific reference intervals or decision limits would be advisable for their interpretation.

As previously noted, free and complexed α₁-antitrypsin may react differently with different antisera used in different assay methods [9]. Differences in antibody specificity and epitope recognition between the evaluated assays may indeed explain the significant constant bias (−0.25 g/L) recorded in our study. Further efforts at standardization of α₁-antitrypsin measurement are, therefore, needed if universal decision limits are to be applied for laboratory screening of inherited α₁-antitrypsin deficiency [14].

Measurement of ceruloplasmin is an important part of the initial screening procedure for Wilson disease, an

### Table 1

<table>
<thead>
<tr>
<th>Protein</th>
<th>Slope (95% CI)</th>
<th>Intercept (95% CI)</th>
<th>Correlation coefficient (95% CI)</th>
<th>Mean bias (95% CI)</th>
<th>Desirable bias (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α₁-Acid glycoprotein</td>
<td>0.995 (0.923 to 1.066)</td>
<td>0.01 g/L (−0.06 to 0.09)</td>
<td>0.989 (0.977 to 0.995)</td>
<td>0.4% (−1.7 to 2.6)</td>
<td>±6.8</td>
</tr>
<tr>
<td>α₁-Antitrypsin</td>
<td>1.003 (0.960 to 1.045)</td>
<td>−0.25 g/L (−0.34 to −0.17)</td>
<td>0.989 (0.976 to 0.995)</td>
<td>−16.8% (−19.3 to −14.3)</td>
<td>±4.3</td>
</tr>
<tr>
<td>Ceruloplasmin</td>
<td>0.572 (0.501 to 0.643)</td>
<td>0.05 g/L (0.03 to 0.08)</td>
<td>0.955 (0.906 to 0.979)</td>
<td>−39.1% (−43.1 to −35.2)</td>
<td>±3.1</td>
</tr>
<tr>
<td>C-reactive protein*</td>
<td>1.072 (1.006 to 1.138)</td>
<td>−0.6 mg/L (−1.0 to −0.1)</td>
<td>0.985 (0.967 to 0.994)</td>
<td>−8.1% (−16.0 to −0.2)</td>
<td>±21.8</td>
</tr>
<tr>
<td>Haptoglobin</td>
<td>1.169 (1.069 to 1.269)</td>
<td>−0.09 g/L (−0.22 to −0.05)</td>
<td>0.984 (0.967 to 0.993)</td>
<td>10.1% (7.3 to 12.9)</td>
<td>±10.4</td>
</tr>
<tr>
<td>Transferrin</td>
<td>0.985 (0.918 to 1.052)</td>
<td>−0.07 g/L (−0.23 to −0.09)</td>
<td>0.990 (0.977 to 0.996)</td>
<td>−5.1% (−7.0 to −3.2)</td>
<td>±1.3</td>
</tr>
<tr>
<td>Transthyretin</td>
<td>1.024 (0.954 to 1.095)</td>
<td>0.02 g/L (0.00 to 0.03)</td>
<td>0.987 (0.972 to 0.994)</td>
<td>9.5% (7.4 to 11.5)</td>
<td>±5.5</td>
</tr>
</tbody>
</table>

*a Two C-reactive protein values above the assay linearity were eliminated from calculation
95% CI 95% confidence interval
autosomal recessive disorder of copper metabolism [15]. Diagnosis relies on reduced serum ceruloplasmin concentrations: Particularly, a serum ceruloplasmin concentration below 0.20 g/L is conventionally considered as one of the major diagnostic criteria [16]. Our results, however, show that despite the use of the same reference material for assay calibration, the method-dependent variation of ceruloplasmin measurement is still huge. Beetham et al. [17] first observed that turbidimetric methods gave significantly higher values than some nephelometric methods in the UK EQAS evaluation, even after the introduction of the BCR-470 for homogeneous calibration of diagnostic systems. They speculated that this may be caused by an alteration in ceruloplasmin structure in the lyophilized reference material, resulting in changes in antigen-antibody reaction when compared with protein in fresh human samples. Indeed, ceruloplasmin is an extremely labile protein, so that the probability that the BCR-470 may contain a protein, which is different in structure and antigenicity, is high [18]. Unfortunately, no studies specifically devoted to evaluate commutability of BCR-470 for ceruloplasmin assays has been performed, as well as no investigations on copper loss from the protein and possible steric changes and proteolytic fragmentation, before to recommend its use for calibration of commercial methods. Lacking a ceruloplasmin reference method, it is currently impossible to demonstrate that one of the two evaluated assays produces true ceruloplasmin values. The problem is, however, that results coming from the two assays are clearly not comparable. Consequently, in agreement with our experimental results, clinicians should use method-dependent decision limits for the diagnosis of Wilson disease instead of taking 0.20 g/L as a universal cut-off [19]. For instance, if a cut-off of 0.20 g/L should be used when ceruloplasmin is measured by the Beckman Immage system, to obtain the same diagnostic performance a decision limit of approximately 0.16 g/L should be used when measured by Roche analyzers.

The problems observed for ceruloplasmin may be either due to a lack of commutability of the reference material with biological samples in the evaluated assays or to calibration problems by the assay manufacturers in one of the stages of the calibration hierarchy. Even though assay manufacturers claim traceability to a higher-order reference material, there is often no information on how traceability is achieved. For instance, there would be a problem with using BCR-470 to standardize clinical assays because this reference material only has a single concentration level for each protein analyte. In order to use BCR-470 to value-assign assay calibrators, assay manufacturers would need to obtain materials with other values of analyte concentration, such as by dilution of BCR-470 [20]. It has already been highlighted in the literature that if BCR-470 is not diluted properly, there are commutability issues for C-reactive protein measurements in the low-concentration range of assays [21].

Although lack of commutability is a possible explanation, we were unable to provide direct evidence that it is the cause of the highlighted problems, because samples of the BCR-470 material are no longer available to perform a commutability study. Indeed, according to the Joint Committee on Traceability in Laboratory Medicine listing, no reference material is currently available for ceruloplasmin [22]. A new batch of serum protein reference material (identified as ERM-DA470k) has recently been released by the IRMM, but ceruloplasmin concentrations were not finally included in the certificate [23]. Without any reference currently available, it is quite difficult to make practical recommendations on how to overcome the issue of non comparability of ceruloplasmin results. IRMM is working with the Committee on Plasma Proteins of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) to identify a suitable material to be able to maintain the assay traceability to the US National Reference Preparation no. 12-0575C, representing the highest level of the ceruloplasmin traceability chain [7, 8].
However, the extreme sensitivity of the protein to degradation, with consequent changes in its immunoreactivity, could make it difficult to finalize the project.

Acknowledgment  The authors gratefully acknowledge the skilful technical assistance of Simona Borille.

References