A model for the harmonisation of test results of different quantitative D-dimer methods

Piet Meijer1,2, Frits Haverkate1, Cornelis Kluft2, Philippe de Moerloose3, Bert Verbruggen4, Michael Spannagl5,6

1ECAT Foundation, Leiden, The Netherlands
2TNO Prevention and Health, Leiden, The Netherlands
3University Hospital, Geneva, Switzerland
4Academic Medical Center St. Radboud, Nijmegen, The Netherlands
5University Hospital, Munich, Germany
6Instand, Düsseldorf, Germany

Summary

The numerical test results of different D-dimer assays vary widely. Because of the complexity of the analyte of target as well as the variability in specificity of different D-dimer assays, only harmonisation of the test results seems to be feasible. The use of a single conversion factor does not take into account for several methods the lack of commutability between test results and consensus values at different D-dimer levels. This is probably related to the mutually different response of methods to high and low levels. We therefore designed a harmonisation model based on the transformation of a method-specific regression line to a reference regression line. We used the data for the measurement of a set of plasma samples with different D-dimer levels by 353 different laboratories using 7 of the most frequently used quantitative D-dimer methods. For each method we calculated the method-specific consensus value for each sample. The overall median value was also estimated. Per method linear regression was applied throughout the method-specific consensus values using the amount of patient pooled plasma added to the different plasma samples as the independent variable. The line through the overall median values of all 7 methods was used as the reference line. Harmonisation between the methods was obtained by transformation of the method-specific regression line to the reference line. This harmonisation resulted in a reduction of the variability between the method-specific consensus values from about 75% to about 5.5%. Clinical validation of this concept had shown significant improvement of the test result comparability. We conclude that this model is a feasible approach in the harmonisation of D-dimer methods. If the harmonisation procedure is included in the calibration procedure by the manufacturers, customers will automatically obtain harmonised test results.

Keywords

D-dimer, harmonisation, linear regression, commutability

Introduction

The measurement of D-dimer, a degradation product of cross-linked fibrin, is mainly applied for the exclusion of deep venous thrombosis (DVT) and pulmonary embolism (PE) in symptomatic out-patients. A variety of different quantitative immunological laboratory methods are available nowadays. However, the numerical test results of different D-dimer assays vary considerably (1). This is mainly caused by the heterogeneity of fibrin degradation products in patient samples as well as the specificity of the different antibodies used in the commercially available test kits for D-dimer (2). This has been clearly shown in an evaluation of 23 quantitative D-dimer assays (3). D-dimer assays show differences in reactivity to different kinds of fibrin derivates, such as high or low molecular weight fibrin derivates, or cross-reactivity with non-cross-linked fibrinogen- and fibrin degradation products. This variety in D-dimer test behaviour is also shown by the results of an external quality assessment programme (4). Another issue which plays a role in the lack of standardisation of D-dimer assays is the different type of calibrators used by the manufacturers. Within the framework of the Subcommittee on Fibrinogen of the Scientific and Standardisation Committee

Thromb Haemost 2006; 95: 567–72

Correspondence to:
Dr. P. Meijer
ECAT Foundation
P.O. Box 2215
2301 CE Leiden
The Netherlands
Tel.: +31 71 5181496, Fax: +31 71 5181330
E-mail: PMeijer@ecat.nl

Received January 20, 2005
Accepted after revision January 4, 2006
Preprinted online February 10, 2006 DOI: 10.1160/TH05–01–0042

© 2006 Schattauer GmbH, Stuttgart
(SSC) of the International Society on Thrombosis and Haemostasis (ISTH) a project was started in 1993 aiming at method standardisation by the use of purified D-dimer as an international standard. However, this approach failed. A second approach, using a whole blood lysate (5), failed too and the use of X-oligomers as an international standard was also not feasible (6). In a second study performed by the Subcommittee on Fibrinogen it was concluded that standardisation of D-dimer assays is impossible. However, harmonisation of D-dimer assays seems to be feasible with the use of a pool of patient plasmas as an international reference preparation (7). The ratio between the consensus value of all methods included and the value of a particular method for a pooled plasma was used as a conversion factor to harmonise the test results of individual patient samples of different methods (7). The feasibility of this approach was confirmed in the Fibrin Assay Comparison Trial (3). However, using this approach, the lack of commutability between test results of some methods and the consensus values at different concentration levels is insufficiently taken into account, as shown by the considerable variation in the ratio at different D-dimer levels for several methods (7).

On the basis of the observation that the use of patient pooled plasma seems to be feasible in the harmonisation of D-dimer methods we set up a project using a set of plasma samples with different concentrations of D-dimer by adding several amounts of a large pool of patient plasmas with elevated D-dimer levels to normal pooled plasma. The samples were distributed to a large number of laboratories, which measured these samples using their regularly used D-dimer method. The results of this study were used to develop a harmonisation model based on the transformation for each method of the linear regression line through the method-specific consensus values to those through the overall-median values. This eventually provided for the harmonisation of D-dimer values over the entire range expected in patients.

Materials and methods

Plasma samples and study set-up
Normal pooled plasma was purchased from Universal Reagents (Indianapolis, United States). A patient pooled plasma was made from about 50 single patient plasmas, routinely collected in the University Hospital Munich from patients with high D-dimer levels, including patients with venous thromboembolism, acute arterial occlusive disease, disseminated intravascular coagulation due to severe liver failure, multi-organ failure or infections. To avoid a significant contribution of degradation products of extravascular origin, post-operative patients were excluded from this plasma pool. The D-dimer level in the patient pool was 12,500 µg/l, measured using the Vidas D-Dimer method (bioMérieux, France).

A set of 7 plasma samples was prepared. The first sample was a 1:1 diluted normal pooled plasma sample with phosphate-buffered saline (0.9%). This sample was not included in the development of the harmonisation model. The second sample was a normal pooled plasma. To prepare the samples A – E an increasing amount of patient pooled plasma was added to normal pooled plasma. The amount of patient pooled plasma added to sample A (15 ml patient pooled plasma to a final volume of 650 ml) was stated as one part. The amount of patient pool in the samples B – E was 2, 3, 8 and 14.7 parts, respectively. The plasma samples were lyophilised at the Institute for Public Health, Brussels, Belgium. The coded samples were, together with instructions for reconstruction, distributed to 502 laboratories participating in the external quality assessment programmes of Instand (national programme in Germany) and the ECAT Foundation (international programme organised in The Netherlands) by postal service. Participating laboratories were asked to measure the D-dimer concentration in the samples with their routinely used D-dimer method. Four hundred and forty-nine laboratories returned their results (89.4%). Because of the lack of sufficient data from 26 laboratories we were able to use the results of 423 laboratories for data evaluation (84.3%). The descriptive results of this study are described elsewhere (8).

D-dimer methods
The participating laboratories used in total 20 different D-dimer methods (8). We included the results of the 8 most frequently used D-dimer methods in the set-up of the harmonisation model, covering 89% of the assessable results. The following D-dimer methods were included in the evaluation: Vidas D-Dimer and Latex D-Dimer (bioMérieux, Marcy L’Etoile, France), D-Dimer Plus and Turbiquant D-Dimer (Dade Behring, Marburg, Germany), Liestest D-Dimer (Diagnostica Stago, Asnières, France), D-Dimer test (Instrumentation Laboratory, Milan, Italy), NycoCard D-Dimer (Nycomed, Oslo, Norway), Tinaquant D-Dimer (Roche Diagnostics, Mannheim, Germany).

Validation of harmonisation procedure
The model for the harmonisation of test results of different quantitative D-dimer methods as described below is validated using samples of consecutive patients with the suspicion of venous thromboembolism (VTE) entering the academic medical centre St. Radboud, Nijmegen, The Netherlands. From all these patients citrated blood was collected after they had given informed consent. In the citrated plasma samples D-dimer was measured using the Tinaquant D-Dimer (Roche, Mannheim, Germany) and the STA Liatest D-Dimer (Diagnostica Stago, Asnières, France).

Data evaluation
Excel 97 was used for the calculations necessary for the harmonisation of D-dimer test results. Statistical analyses were performed using SPSS 10.0.

Results
Consensus values
The method-specific D-dimer consensus value (MSCV) for each sample, defined as the mean value of all results of that particular method, was estimated after the exclusion of outliers ( > 3 * standard deviation) (Table 1). In most cases there were only a few outliers. For the D-Dimer Plus method (Dade Behring) and the Tinaquant method (Roche) we completely excluded the results of 3 and 4 participants, respectively. The results of these participants differed significantly from the rest of the users of these
methods, most probably through the use of other units. The MSCVs of the samples A–E were corrected for the contribution of the D-dimer of the normal pooled plasma. This means that the MSCVs represent only the responsiveness to the patient pooled plasma.

To investigate whether the use of the mean value as method-specific consensus value (MSCV) is permitted, we tested for each method the data set of each plasma sample for normal distribution using the Kolmogorov-Smirnov test. All samples, except the samples B and D of the Latex D-Dimer method of bioMérieux (p=0.037 and p=0.001, respectively), showed a normal distribution. On the basis of this evaluation we concluded that the use of the method-specific consensus values is justified with the exception of the Latex D-Dimer method of bioMérieux. This method was excluded from the development of the harmonisation model.

The overall median value per sample was estimated including all results of the 7 included methods after the exclusion of outliers (see above).

We also investigated whether there is a linear relationship between the method-specific consensus values and the amount of patient pooled plasma added to the different samples. Table 2 shows the slope, intercept and regression coefficient of these regression lines for each of the methods included, as well as the variables for the regression line through the overall median values. All methods showed a high regression coefficient (r > 0.995). We concluded that for all methods included there is good linear relationship between the amount of patient pooled plasma added to the different samples and the method-specific consensus values (MSCVs).

Commutability
To establish whether there is a constant relationship between MSCVs and the overall median values over the whole concentration range under investigation we calculated per method for each of the plasma samples the ratio between those values. These ratios are presented in Table 3. Within a method the coefficient of variation of the ratio varied between 8.3% for the Vidas D-Dimer and 70.9% for the Turbiquant D-Dimer. From Table 3 it is clear that most of the methods showed for the concentration range covered by the samples an inconsistency in the ratio. Inconsistency of the ratio implies a lack of commutability between the MSCVs and the overall median values. This means that for most of the methods the use of a single factor for the harmonisation of test results between methods is not applicable for the whole concentration range. We therefore applied a harmonisation procedure by transforming the linear regression line through the method-specific consensus values of each method to the linear regression line through the overall median values.

Harmonisation procedure
The principle of the harmonisation procedure is shown in Figure 1. To transform the linear regression line through the method-specific consensus values of each method to the linear regression line through the overall median values, the differences of the slopes and intercepts, as given in Table 2, between the method-specific regression lines and the overall median regression line were calculated. The residual slopes and intercepts are given in

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of results</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>bioMérieux Vidas D-Dimer</td>
<td>52</td>
<td>321</td>
<td>655</td>
<td>942</td>
<td>2409</td>
<td>4118</td>
</tr>
<tr>
<td>bioMérieux Latex</td>
<td>25</td>
<td>64</td>
<td>212</td>
<td>251</td>
<td>978</td>
<td>1235</td>
</tr>
<tr>
<td>Dade Behring D-Dimer Plus</td>
<td>75</td>
<td>28</td>
<td>56</td>
<td>88</td>
<td>230</td>
<td>421</td>
</tr>
<tr>
<td>Dade Behring Turbiquant</td>
<td>16</td>
<td>11</td>
<td>35</td>
<td>89</td>
<td>328</td>
<td>638</td>
</tr>
<tr>
<td>Diagnostica Stago Liatest</td>
<td>50</td>
<td>402</td>
<td>945</td>
<td>1336</td>
<td>2863</td>
<td>3551</td>
</tr>
<tr>
<td>IL D-Dimer</td>
<td>35</td>
<td>118</td>
<td>269</td>
<td>413</td>
<td>1197</td>
<td>2159</td>
</tr>
<tr>
<td>Nycomed Nycocard D-Dimer</td>
<td>60</td>
<td>121</td>
<td>246</td>
<td>454</td>
<td>1089</td>
<td>2004</td>
</tr>
<tr>
<td>Roche Tinaquant</td>
<td>65</td>
<td>522</td>
<td>1155</td>
<td>1677</td>
<td>4282</td>
<td>7480</td>
</tr>
<tr>
<td>Overall median value</td>
<td>252</td>
<td>425</td>
<td>736</td>
<td>1733</td>
<td>2816</td>
<td></td>
</tr>
</tbody>
</table>

Note: The bioMérieux Latex D-Dimer method was excluded from the calculation of the overall median value.

<table>
<thead>
<tr>
<th>Method</th>
<th>Slope (ng/ml)</th>
<th>Intercept (ng/ml)</th>
<th>Regression coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>bioMérieux Vidas D-Dimer</td>
<td>276.6</td>
<td>100.8</td>
<td>0.999</td>
</tr>
<tr>
<td>Dade Behring D-Dimer Plus</td>
<td>28.7</td>
<td>0.0</td>
<td>1.000</td>
</tr>
<tr>
<td>Dade Behring Turbiquant</td>
<td>46.6</td>
<td>-47.4</td>
<td>0.999</td>
</tr>
<tr>
<td>Diagnostica Stago Liatest D-Dimer</td>
<td>364.2</td>
<td>129.0</td>
<td>0.996</td>
</tr>
<tr>
<td>IL D-Dimer</td>
<td>149.6</td>
<td>-27.6</td>
<td>1.000</td>
</tr>
<tr>
<td>Nycomed Nycocard D-Dimer</td>
<td>136.7</td>
<td>-1.7</td>
<td>0.999</td>
</tr>
<tr>
<td>Roche Tinaquant</td>
<td>504.9</td>
<td>125.1</td>
<td>0.999</td>
</tr>
<tr>
<td>Overall median line</td>
<td>188.2</td>
<td>112.0</td>
<td>0.994</td>
</tr>
</tbody>
</table>

Table 1: The method-specific consensus values (ng/ml) of 8 different D-dimer methods (between brackets: standard error of the mean) and the overall median value of 5 different plasma samples.

<table>
<thead>
<tr>
<th>Method</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>bioMérieux Vidas D-Dimer</td>
<td>0.79</td>
<td>0.65</td>
<td>0.78</td>
<td>0.72</td>
<td>0.68</td>
<td>8.3</td>
</tr>
<tr>
<td>Dade Behring D-Dimer Plus</td>
<td>9.00</td>
<td>7.59</td>
<td>8.36</td>
<td>7.53</td>
<td>6.69</td>
<td>11.2</td>
</tr>
<tr>
<td>Dade Behring Turbiquant</td>
<td>22.9</td>
<td>12.1</td>
<td>8.27</td>
<td>5.28</td>
<td>4.41</td>
<td>70.9</td>
</tr>
<tr>
<td>Diagnostica Stago Liatest D-Dimer</td>
<td>0.63</td>
<td>0.45</td>
<td>0.55</td>
<td>0.61</td>
<td>0.51</td>
<td>13.2</td>
</tr>
<tr>
<td>IL D-Dimer</td>
<td>2.14</td>
<td>1.58</td>
<td>1.78</td>
<td>1.45</td>
<td>1.30</td>
<td>19.6</td>
</tr>
<tr>
<td>Nycomed Nycocard D-Dimer</td>
<td>2.09</td>
<td>1.73</td>
<td>1.62</td>
<td>1.59</td>
<td>1.41</td>
<td>14.9</td>
</tr>
<tr>
<td>Roche Tinaquant</td>
<td>0.48</td>
<td>0.37</td>
<td>0.44</td>
<td>0.41</td>
<td>0.38</td>
<td>11.4</td>
</tr>
</tbody>
</table>

Table 3: The commutability reflected by the ratio between the overall median value and the method-specific consensus values for each of the 5 plasmas and the coefficient of variation (CV, %) of these ratios for each D-dimer assay included in the evaluation.
Table 4. (As an example: Slope of Vidas regression line = 276.6, slope of the overall median regression line = 188.2, residual slope = 88.5). These residual slopes and intercepts were used to recalculate for each method the consensus values for all plasma samples, resulting in the so-called harmonised consensus values, using the formula: [method-specific consensus value] – [slope * amount of patient pooled plasma + intercept] = harmonised method-specific consensus value. As an example using sample A of the bioMérieux Vidas method: slope = 88.5, intercept = –11.1, consensus value = 321 ng/ml, harmonised consensus value: 321 – (88.5 * 1 – 11.1) = 243.6 ng/ml. If this analysis is performed per method for all the plasma samples the comparability between the consensus values of all methods improves significantly, as indicated by the coefficient of variation (Table 5). The residual variation between the harmonised consensus values is the result of the variation of the method-specific consensus values around their particular regression lines. If the method-specific consensus values are firstly transformed to their exact position on their particular regression lines before the application of the harmonisation procedure, the harmonised consensus values becomes highly comparable.

Validation of harmonisation procedure
The harmonisation procedure was validated on a set of clinical samples from patients suspected from VTE.

The measured test results for the Tinaquant D-Dimer and the Liatest D-Dimer were harmonised as followed: transformation per method of the measured D-dimer result ($Y_M$) to the harmonised D-dimer concentration ($Y_H$) was performed by using the slope (a) and intercept (b) of the method-specific regression line (M) and the reference line (H). This results in the following equation:

$$Y_H = a_M \cdot \left( \frac{Y_M - b_M}{a_M} \right) + b_H$$

For this transformation the slopes and intercepts as given in Table 2 were used. Because we would like to validate if test results of different methods agreed after the application of the harmonisation procedure in the case of exclusion of VTE we only include test results with a D-dimer level < 1500 ng/ml. For this analysis we include the test results of 94 different patient samples. Before and after the application of the harmonisation procedure the difference in test results between the Tinaquant and the Liatest were estimated for all 94 patient samples. The mean difference (between brackets: standard error of the mean) before harmonisation was 517 (82) ng/ml, while after harmonisation this was –62 (32) ng/ml.

Discussion
We devised a method to harmonise the numerical test results of the different D-dimer methods. This initiative was provoked by the fact that the numerical test results of the different D-dimer methods varied considerably. This hampers the comparability of test results using different methods on the same patient sample or the exchange of test results between hospitals using different D-dimer methods. One of the major reasons for these differences is the heterogeneity of fibrin degradation products in patient samples and the difference in the reactivity of antibodies used in the different methods to these degradation products. Studies in vitro have shown that the degradation of cross-linked fibrin by plasmin results in a variety of degradation products containing the dimeric fragment D (9). This variety, which also exists in patient samples, may lead to different test results because of the difference in reactivity of D-dimer methods (3). It was shown in several comparative studies of different D-dimer methods that
this may also have some clinical impact (10–17).

The concordance between the different methods, which was indicated by the kappa coefficients, varied between poor and very good. Because of this complex situation of the heterogeneity of the analyte of target as well as the specificity of the different antibodies used in the commercially available test kits, strict standardisation of D-dimer is in principle impossible. The only option is to harmonise the numerical test results of different D-dimer methods using a pool of patient plasmas as a reference preparation (7). To that end the use of a ratio of a consensus value of a pooled plasma over a reported value for a particular method was suggested (7). However, careful evaluation of the data published by Nieuwenhuizen (7) showed that within a method the ratios obtained for pooled plasmas with different D-dimer levels varied (coefficient of variation: 3.4% to 13.5%, depending on the method used). We confirmed this variation in ratio in our study for several of the methods included in our evaluation. This observation indicates that methods mutually differ in efficacy of response to high and low values. This additional difference between methods is another cause of the lack of concordance expressed in the observation of different ratios between methods at high and low concentrations. On the basis of this observation we conclude that the use of a single factor for each method to harmonise test results of different D-dimer methods does not take into account a possible lack of concordance over the whole concentration range. This requires a solution. Therefore, in our study we used a harmonisation model based on the regression lines through the method-specific consensus values of a set of plasma samples with different D-dimer levels. Because no reference method for D-dimer is available we used as an independent variable the amount of patient pooled plasma added to normal pooled plasma.

Because we also distributed to the participating laboratories the normal pooled plasma as a separate sample, we were able to correct for the contribution of the normal pooled plasma to the result of the samples containing patient pooled plasma. Therefore, the method-specific consensus values used in our harmonisation model reflect only the response of the different methods to the kinds of D-dimer available in this patient pooled plasma. We used a large patient pool including only patients with elevated D-dimer levels from intra-vascular origin. This was chosen because D-dimer testing is nowadays mostly applied for intra-vascular indications, like DVT, PE or DIC. It is known from previous studies that the use of a patient pooled plasma seems to be the most reliable approach for method harmonisation (3, 7). The advantage of this approach is the fact that all different kinds of fibrin degradation products will be available in the patient pooled plasma. A limitation of this approach may be the fact that an individual patient or specific patient groups (e.g. DVT or DIC) may behave differently.

Furthermore, due to the large number of laboratories participating in this study we believe that the estimated method-specific consensus values are a reliable approach to the so-called “analytical true values” for these samples.

Due to the lack of a reference method it is impossible to compare the results of a particular method with those of the reference method. To overcome this problem we use the regression line through the overall median values as a reference as based on the test results of the 7 included methods in our evaluation. Although this approach is arbitrary we believe that it is reliable because the 7 included methods represent about 84% of all test results obtained in this study, which includes a large number of different laboratories from all over the world. The remaining 16% of the test results were obtained by the use of 13 other D-dimer methods. It is clear that these remaining methods are only used by a minor part of the participating laboratories.

An important condition for the use of our harmonisation approach is a linear relationship between the amount of patient pooled plasma in the different samples and the response to these amounts by the different methods. We observed for almost all methods a satisfactory regression coefficient for the concentration range used in our study, which indicates that there is indeed a linear relationship. It allows us to develop a harmonisation approach based on linear regression analysis. Our harmonisation approach is based on the transformation of the regression line through the method-specific consensus values to the regression line through the overall median values. The difference between the regression lines through the method-consensus values and the overall median values resulted in a residual slope and intercept. These parameters were used to calculate the so-called harmonised consensus values. The effect of this harmonisation was reflected by a significant reduction of the variation between the method-consensus values before and after harmonisation. The variation left reflects the variability of the method-specific consensus values around the regression lines. If the method-specific consensus values were firstly exactly transformed to the regression lines no difference between the harmonised consensus values existed anymore.

To validate the proposed harmonisation model it was applied to a set of clinical samples of patients suspected of VTE on which both the Tinaquant and STA Liatest D-dimer tests were applied. From the mean difference value before and after harmonisation it is clear that the difference between both methods was reduced significantly. The negative mean difference value after harmonisation is probably caused by the fact that the method-specific consensus values are established based on a large field study. It is expected when these consensus values are established using expert laboratories the difference between test results of different methods will be negligible. From the residual standard error of the mean it is clear that the harmonisation procedure will not result in complete negligible differences in test results between different methods. This is related to the fact that particular patient samples behave differently in different methods which is shown by the kappa coefficients in different comparative studies (11, 14). This phenomenon is also known from the harmonisation of the prothrombin time by the calculation of the International Normalized Ratio (INR) (18). This concept is well accepted nowadays. Although the harmonisation of D-dimer results will probably not be completely possible for particular clinical samples, in general it is expected that the application of the harmonisation procedure described in this paper will significantly improve test result comparability.

We conclude that harmonisation of the test results of different D-dimer assays, on the basis of the transformation of a method-specific regression line to a reference line, seems to be feasible. The limitation of the harmonisation procedure dis-
discussed here is its dependence on the methods included. Because we only included the data of the most frequently used D-dimer methods performed by a large number of laboratories, we believe that we were able to estimate a reliable reference line based on the overall median values for the different samples used. In future this reference line can be used as a “golden standard” to harmonise each D-dimer method.

To make this harmonisation approach applicable in clinical practice, manufacturers of D-dimer assays should apply this harmonisation procedure to the calibrators they used. If the manufacturers assign their calibrators a harmonised value, customers of a particular method will obtain automatically harmonised test results. The effect of harmonisation on the test outcome of different D-dimer methods should be proven by the measurement of a set of clinical samples.

Because this harmonisation procedure is in fact only a mathematical procedure it is expected that cut-off levels used in clinical practice could also be harmonised in the same way. However this needs to be subjected to further study.

References