Long-Term Analytical Performance of Hemostasis Field Methods as Assessed by Evaluation of the Results of an External Quality Assessment Program for Antithrombin

PIET MEIJER,1,2* MONIEK P.M. DE MAAT,1,2 CORNELIS KLUFT,1,2 FRITS HAVERKATE,1 and HANS C. VAN HOUWELINGEN1,3

**Background:** It is important for a laboratory to know the stability of performance of laboratory tests over time. The aim of this study was to adapt from the field of clinical chemistry a method to assess the long-term analytical performance of hemostasis field methods.

**Methods:** The linear regression model was used to compare the laboratory results with the consensus mean value of a survey. This model was applied to plasma antithrombin activity using the data for 82 laboratories, collected between 1996 and 1999 in the European Concerted Action on Thrombosis (ECAT) external quality assessment program. The long-term total, random, and systematic error were calculated. The variables introduced to define the long-term performance in this model were the long-term analytical CV (LCVa) and the analytical critical difference (ACD), which indicates the minimum difference necessary between two samples measured on a long-term time-scale to consider them statistically significantly different.

**Results:** The systematic error (bias) ranged from 4.5 to 103 units/L. The random error ranged from 24.4 to 242 units/L. For the majority of the laboratories, random error was the main component (>75%) of the total error. The LCVa after adjustment for the contribution of the bias, ranged from 2.8% to 48%. The ACD ranged from 78 to 1290 units/L with a median value of 190 units/L. No statistically significant differences were observed for either LCVa or ACD between the two different measurement principles for antithrombin activity based on the inhibition of either thrombin or factor Xa.

**Conclusions:** This linear regression model is useful for assessing the total error, random error, and bias for hemostasis field methods. The LCVa and ACD for measurement on a long-term time-scale appear to be useful for assessing the long-term analytical performance.

© 2002 American Association for Clinical Chemistry

In clinical chemistry, the evaluation of external quality assessment (EQA)\(^4\) results has been used for many years to assess the performance of an analytical method over time (1–3). Linear regression is frequently applied and, if linearity is assumed, provides an easy model to assess random and systematic errors (4). We adapted this approach to evaluate the long-term performance of laboratory tests involved in the diagnosis of thrombophilia based on the results of an EQA program of the European Concerted Action on Thrombosis (ECAT) Foundation and to trace possible factors involved in variations in analytical performance. Because no established standardization system exists for hemostasis (5), we compared individual laboratory test results with the consensus value, here defined as the mean value of all test results or the mean within-subject biological CV.

---

1 ECAT Foundation, 2301 CE Leiden, The Netherlands.
2 Gaubius Laboratory, TNO Prevention and Health, 2333 CK Leiden, The Netherlands.
3 Leiden University Medical Center, Department of Medical Statistics, 2300 RC Leiden, The Netherlands.
*Address correspondence to this author at: ECAT Foundation, PO Box 2215, 2301 CE Leiden, The Netherlands. Fax 31-71-5181330; e-mail P.Meijer@ecat.nl.
Received January 21, 2002; accepted April 23, 2002.

---

4 Nonstandard abbreviations: EQA, external quality assessment; ECAT, European Concerted Action on Thrombosis; LCVa, long-term analytical CV; ACD, analytical critical difference; TE, total error; RE, random error; and CVw, within-subject biological CV.
analogy with the clinical critical difference, which indicates whether the difference between two serial test results on the same individual is statistically significant \((6, 7)\). The ACMD was defined as the minimum analytical capability of a laboratory to significantly distinguish two test results.

The model was applied to the test results for plasma antithrombin activity, a routine assay frequently used in the hemostasis laboratory, over the period 1996–1999.

**Materials and Methods**

**ANTITHROMBIN SAMPLES**

To mimic real clinical laboratory practice as much as possible, we used only lyophilized plasma from healthy volunteers or patients with antithrombin activity in the range of 400-1200 units/L for distribution within the surveys of the ECAT Foundation. Lyophilized normal citrated pooled plasma samples were obtained from Chromogenix and Biopool. Citrated patient plasma samples were collected at different hospitals, according to the local ethical and technical procedures, and were lyophilized at the Institute for Public Health (Brussels, Belgium). Antithrombin activity is expressed as the percentage of activity of a normal pooled plasma: 100% corresponds to 1000 units/L.

**LABORATORY METHODS**

Two enzymatic assay principles exist for the determination of antithrombin activity, both using a chromogenic substrate. One is based on the inhibition of thrombin (IIa) by antithrombin; the other is based on the inhibition of activated clotting factor X (Xa) by antithrombin.

**DESIGN**

One plasma sample with instructions for sample handling was distributed four times a year to participating laboratories. The participants were asked to handle the sample as if it was a typical clinical laboratory sample, using their routine method for the determination of antithrombin activity. The consensus value for antithrombin activity was estimated after outliers (>3 SD) were excluded \((8)\). The final mean value was calculated and considered the consensus value for that particular survey. The data set was checked for gaussian distribution using the Kolmogorov–Smirnov test.

Comparison of the consensus values for both assay types was performed by the Student t-test.

Consensus values obtained for plasmas used in more than one survey were compared by ANOVA for repeated measures, adjusting for missing values, including the possible effect of the assay type. Because some laboratories changed to another method based on a different assay principle within the period for the long-term evaluation, they were categorized as a heterogeneous assay group.

The correlation of interlaboratory variation with the consensus value was evaluated by the Spearman \(\rho\) correlation.

---

**LONG-TERM DATA EVALUATION**

The long-term laboratory performance for plasma antithrombin activity for the time period 1996–1999 was evaluated. Only those laboratories with at least eight test results in that period were included in the evaluation. The total number of laboratories included was 82.

Linear regression based on the least-squares method was applied using the consensus value \((x)\) as the independent and the laboratory value \((y)\) as the dependent variable. The slope \((b)\) and the variability \((S_{y|x})\) of each regression line were calculated. Goodness-of-fit was evaluated by the addition of a quadratic term in the regression model \((4)\), and the mean values for \(x\) (\(\bar{x}\)) and \(y\) (\(\bar{y}\)) as well as the standard error of \(x\) (\(S_x\)) were calculated. The number of laboratory results included is expressed by \(n\).

The long-term total error (TE) is expressed by the sum of the differences between the laboratory value and the consensus value for each set of these values for a laboratory included in the evaluation. The long-term TE can be calculated by the formula:

\[
TE = \sqrt{\frac{1}{n} \sum (y - x)^2} \tag{1a}
\]

The long-term TE is determined by two components, the total bias \((B)\), which is the deviation of the ideal regression line \(y = x\), and the random analytical error \((RE)\), which is the variability of data points around the regression line. Both components can be derived from Eq. 1a as follows:

\[
\frac{1}{n} \sum (y - x)^2 = \frac{1}{n} [(b - 1)^2 \cdot \sum (x - \bar{x})^2 + n(\bar{y} - \bar{x})^2 + \sum (y - \bar{y})^2]
\]

\[
= (b - 1)^2 \cdot \frac{n - 1}{n} S_x^2 + (\bar{y} - \bar{x})^2 + \frac{n - 2}{n} S_{y|x}^2 \tag{1b}
\]

The factors \(\frac{n-1}{n}\) and \(\frac{n-2}{n}\) in the formulas are attributable to the difference in corrections for degrees of freedom in the definition of TE, \(S_{y|x}\), and \(S_x\). These terms are needed to obtain strict equality in Eq. 1b. A slightly looser definition would dispose of all these terms and would define TE as:

\[
TE = \sqrt{(b - 1)^2 \cdot S_x^2 + (\bar{y} - \bar{x})^2} \tag{1c}
\]

The derivative of Eq. 1b can be used to calculate the long-term total bias \((B)\) by the formula:

\[
B = \sqrt{\frac{n - 1}{n} \cdot (b - 1)^2 \cdot S_x^2 + (\bar{y} - \bar{x})^2} \tag{2}
\]

The total bias consists of two components, the proportional bias \((PB)\) and the constant bias \((CB)\). The proportional bias is caused by the deviation of the slope and depends on the variability of the consensus value:
The constant bias reflects the deviation from the consensus value:

\[ \text{CB} = \sqrt{\left( \bar{y} - \bar{x} \right)^2} \]  
(4)

The ratio of the proportional and constant bias indicates whether the bias is caused mainly by calibration errors (proportional bias) or by other factors, such as matrix effects (constant bias).

The long-term random analytical error (RE) can be calculated by the formula:

\[ \text{RE} = \sqrt{\frac{n - 2}{n} \cdot S_{y|x}^2} \]  
(5)

The LCVa is based on the variability of the regression line (S_{y|x}) and the mean value of all consensus values (\(\bar{x}\)). To allow comparison of the LCVa among laboratories, it should be calculated after adjustment for the bias. Therefore, the LCVa is now calculated using the formula:

\[ \text{LCVa} = \frac{S_{y|x}}{\bar{x}} \cdot 100\% \]  
(6)

ACD

To evaluate the practical consequence of the long-term data analysis on laboratory performance, we calculated the ACD. The ACD, which reflects the minimum difference necessary on the consensus scale to consider two samples measured on a long-term time-scale as significantly different, was calculated by the following formula:

\[ \text{ACD} = \frac{S_{y|x}}{b} \cdot \sqrt{2} \cdot t \]  
(7)

where \(t\) is the \(t\)-statistic for \((n - 2)\) degrees of freedom with a level of significance of 95%. The value of \(t\) depends on the number of test results included in the evaluation.

For both the LCVa and the ACD, the effect of the different assay types was evaluated by Kruskal–Wallis one-way ANOVA. Twenty-six laboratories consistently used a method based on the inhibition of thrombin, whereas 36 laboratories consistently used a method based on the inhibition of activated clotting factor X. Twenty laboratories belonged to the heterogeneous assay group. Because of the small numbers, no further discrimination within the homogeneous assay group for the method used was made.

All statistical analyses were performed with the computer program SPSS, Ver. 9.0

### Results

**APPLICATION OF LINEAR REGRESSION TO ANTITHROMBIN**

**Consensus values.** The consensus antithrombin activity values (units/L), the interlaboratory CV, and the number of outliers for each survey are shown in Table 1. For each survey, the total mean value of all methods was used as the consensus value because no statistically significant differences were observed between the results of the two assay principles for the measurement of antithrombin activity. For all surveys, the data set showed a gaussian distribution. With the exception of survey 99/04, the interlaboratory CVs were comparable for all surveys (6.0–9.5%). There were slightly higher interlaboratory CVs for samples with a lower antithrombin value (Spearman correlation, \(P = 0.049\)). Although sample ECAT-17 had a rather low antithrombin activity, we have no explanation for the very high interlaboratory CV observed with this plasma in survey 99/04.

The plasmas ECAT-1, -6, -7, and -14 were used in more than one survey. For these plasmas, no statistically significant differences between consensus values for the different surveys were observed, both for the total group and the different assay groups. These findings supported the use of the consensus value based on all test results in a survey.

**Application of the linear regression model.** From the data for the slope (\(b\)) and the variability (\(S_{y|x}\)), it is clear that there was a wide variation in the characteristics of the different regression curves (Table 2). The minimum and maximum slopes differed by a factor of 2, and the \(S_{y|x}\) by a factor of 10. Verification of the goodness-of-fit showed that in \(\sim90\%\) of the laboratories, the linear regression model was

### Table 1. Consensus values, interlaboratory CV, and number of outliers of the 16 surveys included in the long-term data evaluation of plasma antithrombin activity.

<table>
<thead>
<tr>
<th>Survey code</th>
<th>No. of participants</th>
<th>Plasma code</th>
<th>Consensus value, units/L</th>
<th>CV, %</th>
<th>No. of outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>96/02</td>
<td>61</td>
<td>ECAT-1</td>
<td>948</td>
<td>6.1</td>
<td>2</td>
</tr>
<tr>
<td>96/03</td>
<td>64</td>
<td>ECAT-1</td>
<td>946</td>
<td>7.2</td>
<td>1</td>
</tr>
<tr>
<td>96/04</td>
<td>64</td>
<td>ECAT-1</td>
<td>957</td>
<td>7.1</td>
<td>1</td>
</tr>
<tr>
<td>97/03</td>
<td>77</td>
<td>ECAT-1</td>
<td>932</td>
<td>6.8</td>
<td>2</td>
</tr>
<tr>
<td>97/04</td>
<td>78</td>
<td>ECAT-1</td>
<td>950</td>
<td>7.1</td>
<td>1</td>
</tr>
<tr>
<td>98/01</td>
<td>75</td>
<td>ECAT-1</td>
<td>928</td>
<td>7.4</td>
<td>0</td>
</tr>
<tr>
<td>96/01</td>
<td>59</td>
<td>ECAT-2</td>
<td>456</td>
<td>9.2</td>
<td>2</td>
</tr>
<tr>
<td>97/02</td>
<td>65</td>
<td>ECAT-6</td>
<td>872</td>
<td>9.5</td>
<td>1</td>
</tr>
<tr>
<td>97/02</td>
<td>69</td>
<td>ECAT-6</td>
<td>874</td>
<td>6.0</td>
<td>2</td>
</tr>
<tr>
<td>98/03</td>
<td>82</td>
<td>ECAT-6</td>
<td>864</td>
<td>8.5</td>
<td>1</td>
</tr>
<tr>
<td>99/01</td>
<td>102</td>
<td>ECAT-6</td>
<td>852</td>
<td>7.9</td>
<td>3</td>
</tr>
<tr>
<td>98/02</td>
<td>82</td>
<td>ECAT-7</td>
<td>665</td>
<td>8.9</td>
<td>1</td>
</tr>
<tr>
<td>98/04</td>
<td>87</td>
<td>ECAT-7</td>
<td>649</td>
<td>8.8</td>
<td>2</td>
</tr>
<tr>
<td>99/02</td>
<td>107</td>
<td>ECAT-14</td>
<td>1152</td>
<td>8.1</td>
<td>1</td>
</tr>
<tr>
<td>99/03</td>
<td>105</td>
<td>ECAT-14</td>
<td>1154</td>
<td>7.8</td>
<td>1</td>
</tr>
<tr>
<td>99/04</td>
<td>116</td>
<td>ECAT-17</td>
<td>467</td>
<td>27</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 2. Overall statistical characteristics of the linear regression curves of the 82 laboratories based on the long-term evaluation of EQA results of plasma antithrombin activity.

<table>
<thead>
<tr>
<th>Curve characteristics</th>
<th>Calculated TE, bias, and RE, units/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope (b)</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>70.4</td>
</tr>
<tr>
<td>0.13</td>
<td>36.3</td>
</tr>
<tr>
<td>1.02</td>
<td>61.9</td>
</tr>
<tr>
<td>0.63</td>
<td>26.2</td>
</tr>
<tr>
<td>1.28</td>
<td>260.9</td>
</tr>
</tbody>
</table>

*For equations, see Materials and Methods.*

applicable to this verification. Visual inspection of the curves from the other laboratories did not show any essential difference from curves from laboratories with a valid linear regression curve. On the basis of these observations, it was decided that application of the linear regression model for laboratories was permitted.

**TE, RE, and bias of the antithrombin assay.** Data on the long-term TE, RE, and bias showed a great difference in analytical performance among the participating laboratories (Table 2). The imprecision of the TE in 46 laboratories (56% of laboratories) was >75%, whereas the bias of the TE was >75% in 3 laboratories (4%). The contribution of constant bias to the total bias was evaluated by the ratio of the constant bias (Eq. 4) and the total bias (Eq. 2). In 24 laboratories (29%), the bias was caused mainly by a proportional bias (ratio <25%). In 26 laboratories (32%), the bias was mainly caused by a constant bias (ratio >75%). All other laboratories showed a mixture of both types of bias (25% < ratio < 75%).

**Long-term Analytical Performance of an Individual Laboratory**

**LCV<sub>a</sub> of the antithrombin assay.** The LCV<sub>a</sub> was calculated for each laboratory on the basis of the standard error of the regression curve. The wide variation in LCV<sub>a</sub> among laboratories is clearly shown in Fig. 1. The median LCV<sub>a</sub> was 7.2% (range, 2.8–48%). Fifty percent of the laboratories had an LCV<sub>a</sub> <7.2%, whereas 90% of the laboratories had an LCV<sub>a</sub> <13%. There was no statistically significant difference between the assay principles, including the heterogeneous assay type group, for the LCV<sub>a</sub>.

**Long-term ACD of the antithrombin assay.** We calculated the long-term ACD as a practical indicator of the analytical laboratory performance. The ACD reflects the minimum difference necessary on the consensus scale to consider two samples measured on a long-term time-scale as significantly different. In our survey, the longitudinal period was that of the survey frequency and was divided into 3-month periods. The median ACD was 190 units/L (range, 78–1287 units/L), whereas the 95% content interval was 79–647 units/L.

On the basis of these results, for the best-performing laboratory two samples measured on a long-term time-scale could be considered statistically different if the difference in antithrombin activity between the samples was 78 units/L, whereas for the worst-performing laboratory, a difference of ~1290 units/L was necessary.

A comparison of laboratories with a low or high ACD showed that the size of the ACD was not related to the method used. However, because of the small number of laboratories consistently using the same method over the whole time period, no further statistical evaluation of laboratory performance in relation to the method used could be made.

**Discussion**

To support laboratories performing hemostasis assays in monitoring the long-term analytical performance of their field methods, we adapted a linear regression model from the field of clinical chemistry for evaluation of EQA results over a prolonged period of time (4).

We showed in this study that the consensus values were stable over time and independent of the number of laboratories participating in the survey. Furthermore, in our study, the consensus values were independent of the assay principle. This indicates that the consensus values obtained in these surveys are reliable estimates, which makes the use of the consensus value suitable for the assessment of the long-term laboratory performance for the analyte evaluated: antithrombin activity.

Although the interlaboratory CVs for antithrombin in the majority of the surveys were quite reasonable for this type of bioanalytical laboratory method (Table 1), we observed considerable variability in the long-term performance of each laboratory. This was related to large variations in the bias and imprecision. Our results provide evidence that the source of bias differs among
laboratories, but that for the majority of laboratories, RE plays a more important role in TE than does bias. This conclusion is supported by the wide range of the LCVs (Fig. 1).

To make the impact of long-term laboratory performance clearer for the daily laboratory routine, we introduced the long-term ACD. On average, the ACD was ~230 units/L, which means that the average laboratory could distinguish two samples measured on a long-term time-scale as significantly different only if there was a difference of ~230 units/L in antithrombin activity. The ACD applies to the time period covered by this evaluation. A change in the regression relationship implies a change of the ACD. The ACD therefore can be used to monitor changes in laboratory performance in the future.

We evaluated the number of laboratories for which the LCVa was less than one-half the within-subject biological CV (CVi) for monitoring and 0.58 times the total biological variation for diagnostic testing (9, 10). The reported CVi for antithrombin varies from 1.1% (11) to 3.1% (6). On the basis of the reported median CVi, none of the participating laboratories fulfills the criterion of CVa <0.5 CVi for monitoring. On the basis of the data for the within- (1.1%) and between-subject (10.4%) biological variation reported by Chambless et al. (11), we calculated a total biological variation for antithrombin of 10.5%. If we use these data in our model, only 20% of the participating laboratories fulfilled the goal for diagnostic testing. Although antithrombin has been a frequently used test in hemostasis laboratories, we conclude that only a minority of the laboratories can fulfill the widely accepted analytical goals based on the biological variation.

In conclusion, the model presented here characterizes the long-term analytical performance for hemostasis field methods. The ACD was introduced to highlight the impact of the long-term analytical performance on the daily laboratory routine. We believe that this linear regression model with a consensus value may be applied more generally for analytes lacking a reference system.

Further data will be required to find the sources of the observed long-term variability of antithrombin measurements.

We gratefully acknowledge all participants in the EQA program of the ECAT Foundation for their contributions.

References