Combi Scheme: New Combined Internal/External Quality-Assessment Scheme in The Netherlands

Herman Steigstra,1 Rob T. P. Jansen,2 and Henk Baadenhuijsen1

The Dutch Foundation for Quality Assessment in Clinical Chemistry (SKZL) is the professional organization that conducts external quality-assessment schemes in The Netherlands. However, such schemes in fact assess the performance of the internal quality-control systems of the participating laboratories. In this paper we describe a new concept, relating the data for internal control materials with those for external samples and thereby leading to a combined external/internal scheme (Combi). The statistical principles underlying the Combi scheme are discussed and examples of the graphical presentation of the results are shown. Because the laboratory data are transmitted over the public telephone system to the computers of the SKZL, we also describe the principles of the data communication. At two-month intervals a statistical presentation is sent to all participants. The central database is updated daily with the received results, making possible an on-line consultation regarding the statistics of the accumulated findings of the control materials in use.

Additional Keyphrases: statistics • quality control • proficiency testing

The necessity of external quality assessment of analytical processes in clinical laboratories has long been recognized. Laboratories all over the world participate in external proficiency testing schemes in which unknown control materials are supplied and the results are processed in regional or national centers.

The intensive use of internal control materials to monitor analytical processes is considered normal laboratory practice. Various commercial quality-assessment schemes process the data obtained from these internal samples. This external processing of internal data, which should not be confused with external quality assessment, may give a flatteringly and biased image of the real analytical process, because the internal results are used to intervene in the process, the process is adjusted to the target values of the internal control samples, and specific matrix effects of the controls are masked. Nonetheless, a large amount of information is contained in the internal data.

External control materials serve to assess the matrix and concentration sensitivity of the analytical processes and mimic the analysis of individual unknown patients' samples. They assess the quality of the internal-control procedures.

The SKZL (Dutch Foundation for Quality Assessment in Clinical Chemistry) is the professional organization that conducts several external quality-assessment schemes in The Netherlands. The external scheme for general clinical chemistry (1) was started in 1973 and is still the most important chemistry scheme in The Netherlands. Almost all hospital laboratories participate. Although the current program is essentially the same as the original one, the presentation has been changed to an almost complete graphical individualized layout, printed on a high-resolution multiple light-emitting diode printer.

Since 1973, schemes covering a broad portion of clinical chemistry and hematology have been introduced, including a coupled external/internal scheme (2) for general clinical chemistry and schemes for pH, blood gas analysis, immunochemistry, hematological cell counting, blood smears, transfusion, urinary chemistry, serology, and analyses of renal calculi. Here, we describe a new combined external/internal scheme that relates data for internal control materials to those for external samples.

Materials and Methods

Control Materials

Laboratories are free to choose their own internal control materials. Information on the materials used (manufacturer, batch number, etc.) are to be sent to the SKZL, so that results of laboratories using the same batches of control material can be combined and statistically processed.

The external materials are supplied by the SKZL. The participants analyze one control sample weekly. A different control material is used weekly, thus allowing tests at various concentrations and in different matrices, from time to time including fresh-frozen human sera. In the hematology scheme, fresh human blood samples are used throughout.

Statistical Methods

Scaling. The raw data of the internal as well as the external control materials are first scaled to obtain data sets with comparable means and standard deviations. Results for the internal control materials are scaled to the mean value of all laboratories that use the same batch—or to the individual laboratory mean value, if the laboratory is the sole user of a material—and a reference standard deviation. Results for the external

1 Dutch Foundation for Quality Assessment in Clinical Chemistry, St. Radboud University Hospital, Geert Grooteplein 8, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands.
2 Clinical Laboratory, St. Anna Hospital, Bogaerde 2, 5664 EH Geldrop, The Netherlands.

Received November 14, 1990; accepted April 10, 1991.
control materials are scaled to a reference value and a reference standard deviation. The following algorithms are used:

Internal samples:

\[ X_{ik}' = (X_{ik} - \bar{X})/SD_{ref} \quad \text{with } SD_{ref} = f(\bar{X}) \]

External samples:

\[ Y_{ij}' = (Y_{ij} - Y_{jref})/SD_{jref} \quad \text{with } SD_{jref} = f(Y_{jref}) \]

in which

- \( X_{ik}' \) is the scaled \( k \)th value of the appropriate internal control sample of laboratory \( i \)
- \( X_{ik} \) is the \( k \)th value of the internal control sample of laboratory \( i \)
- \( \bar{X} \) is the batch group mean value of the internal control sample
- \( SD_{ref} \) is the reference SD for the internal sample
- \( Y_{ij}' \) is the scaled value of the \( j \)th external control sample of laboratory \( i \)
- \( Y_{ij} \) is the value of the \( j \)th external control sample of laboratory \( i \)
- \( Y_{jref} \) is the reference value of the \( j \)th external control sample
- \( SD_{jref} \) is the reference SD for the \( j \)th external control sample

The reference SD is the desired analytical precision at the concentration value of interest, and can be based on the intra-individual biological variance \( \sigma_k \), as advocated by Fraser (3) and Harris (4), who demonstrated that the desired analytical SD should not exceed one-half of the intra-individual SD. On the one hand, this strategy leads to rather sharp analytical criteria: e.g., for sodium a CV\(_{analytical} \) (CV\(_a \)) of 0.3% is required, which is far too demanding, given that the state-of-the-art CV\(_a \) is at best 0.9%. On the other hand, some analytes can now be measured much more precisely than is medically required. For instance, on the basis of intra-individual variation, measurement of the enzyme activity of alanine aminotransferase requires a CV\(_a \) of 13.6%, whereas typical within-laboratory CVs of 4% are achievable. Therefore, in practice, we use a combination of the theoretical goals of Fraser and goals considered as realistic targets for a state-of-the-art challenge.

Yet another aspect must be dealt with. Especially for the external results, the SD\(_{ref} \) has to compensate for the effect that analytical precisions are not constant over the whole concentration/activity range. Therefore, the SD\(_{ref} \) should also take into account the precision/concentration profiles of each of the analytes, which can be derived retrospectively from earlier trial results. This approach is planned for use soon but has not yet been implemented. At present, we still use an analytical function describing a relationship between the concentration of the analyte and the expected analytical imprecision at that concentration. If a linear relationship between concentration and analytical variance is assumed, this function can be formalized as

\[ CV_{ref} = CV_t \sqrt{(C_t/C)} \]

in which \( C \) is the (mean) concentration of concern, \( C_t \) the typical tabulated concentration, and \( CV_t \) the accompanying coefficient of variation. Table 1 lists the values for these target CVs at a stated concentration or activity. For comparison, the theoretical goals [according to Fraser (3)] are also given.

From the scaled values, mean values and SDs are computed for the internal samples as well as for the external samples, within laboratories and between laboratories, within methods and within method groups, and overall, as follows:

\[ \bar{X}_i' = \frac{1}{n_i} \sum_{k=1}^{n_i} X_{ik}'/n_i = \text{scaled mean of laboratory } i \text{ of the } n_i \text{ results for the appropriate internal sample.} \]

\[ SDX_i' = \sqrt{\frac{1}{n_i} \sum_{k=1}^{n_i} (X_{ik}' - \bar{X}_i')^2/(n_i - 1)} = \text{scaled SD of laboratory } i \text{ of the internal sample of concern.} \]

<table>
<thead>
<tr>
<th>Component</th>
<th>Unit</th>
<th>( C_t )</th>
<th>CV(_a ) %</th>
<th>Goals, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium</td>
<td>mmol/L</td>
<td>140</td>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>Potassium</td>
<td>mmol/L</td>
<td>5</td>
<td>1</td>
<td>2.4</td>
</tr>
<tr>
<td>Chloride</td>
<td>mmol/L</td>
<td>100</td>
<td>1</td>
<td>0.7</td>
</tr>
<tr>
<td>Calcium</td>
<td>mmol/L</td>
<td>2.5</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>Phosphorus</td>
<td>mmol/L</td>
<td>2</td>
<td>2</td>
<td>4.0</td>
</tr>
<tr>
<td>Magnesium</td>
<td>mmol/L</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Iron</td>
<td>( \mu )mol/L</td>
<td>30</td>
<td>4</td>
<td>15.9</td>
</tr>
<tr>
<td>Iron-bind. capacity</td>
<td>( \mu )mol/L</td>
<td>50</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td>mmol/L</td>
<td>5</td>
<td>3</td>
<td>6.3</td>
</tr>
<tr>
<td>Creatinine</td>
<td>( \mu )mol/L</td>
<td>200</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Uric acid</td>
<td>mmol/L</td>
<td>0.5</td>
<td>4</td>
<td>4.2</td>
</tr>
<tr>
<td>Glucose</td>
<td>mmol/L</td>
<td>10</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Total protein</td>
<td>( g )L</td>
<td>50</td>
<td>2</td>
<td>1.4</td>
</tr>
<tr>
<td>Albumin</td>
<td>( g )L</td>
<td>40</td>
<td>3</td>
<td>1.4</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>( \mu )mol/L</td>
<td>25</td>
<td>4</td>
<td>11.3</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>mmol/L</td>
<td>5</td>
<td>2</td>
<td>2.7</td>
</tr>
<tr>
<td>HDL cholesterol</td>
<td>mmol/L</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Triglycerides</td>
<td>mmol/L</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Osmolality</td>
<td>mmol/kg</td>
<td>300</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Alk. phosphatase</td>
<td>U/L</td>
<td>100</td>
<td>2</td>
<td>3.4</td>
</tr>
<tr>
<td>Aspartate aminotrans.</td>
<td>U/L</td>
<td>100</td>
<td>2</td>
<td>7.2</td>
</tr>
<tr>
<td>Alanine aminotrans.</td>
<td>U/L</td>
<td>100</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>Lactate dehydrog.</td>
<td>U/L</td>
<td>200</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>( \gamma )-Glutamyltransferase</td>
<td>U/L</td>
<td>100</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Creatine kinase</td>
<td>U/L</td>
<td>100</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

HDL, high-density lipoprotein.
* As published by Fraser (3).
\[ \bar{Y}_i = \frac{\sum_{j=1}^{m_i} Y_{ij} / m_i}{m_i} = \text{scaled mean of laboratory } i \text{ for the } m_i \text{ external results.} \]

\[ SDY_i' = \sqrt{\sum_{j=1}^{m_i} (Y_{ij} - \bar{Y}_i')^2 / (m_i - 1)} = \text{scaled SD of laboratory } i \text{ of the external samples.} \]

\[ X' = \frac{\sum_{i=1}^{n_{lab}} n_i X_i}{\sum_{i=1}^{n_{lab}} n_i} = \text{scaled group mean of internal samples of the } n \text{ laboratories using the same method, method group, or batch of internal material.} \]

\[ SDX' = \sqrt{\sum_{i=1}^{n_{lab}} (n_i - 1)SDX_i' / \sum_{i=1}^{n_{lab}} n_i} \frac{SDX_i'}{n_{lab}} = \text{scaled group within-laboratory SD of the internal samples of the } n \text{ laboratories using the same method, method group, or batch of internal material.} \]

\[ Y_i' = \frac{\sum_{i=1}^{n_{lab}} m_i \bar{Y}_i'}{\sum_{i=1}^{n_{lab}} m_i} = \text{scaled mean of external results for the } n \text{ laboratories using the same method or method group.} \]

\[ SDY' = \sqrt{\sum_{i=1}^{n_{lab}} (m_i - 1)SDY_i' / \sum_{i=1}^{n_{lab}} m_i} \frac{SDY_i'}{n_{lab}} \]

\[ = \text{scaled group within-laboratory SD of the external results for the } n \text{ laboratories using the same method or method group.} \]

The scaled values are used in the graphical presentation of results. Only a few values are presented in a table. The internal laboratory mean value \( X'_i \) and SD (SDX'\(_i\)), and batch group mean \( X' \) and average within-laboratory SD (SDX'\(_i\)) for each of two internal control concentrations, are scaled back to their respective \( X \) values. The external laboratory mean \( Y'_i \) and SD (SDY'\(_i\)), and the external method mean \( Y' \) and average within-laboratory SD (SDY'\(_i\)), are scaled back to the mean value of the target values \( Y'_\text{ref} \).

**Outlier detection.** Second, the scaled data are processed to detect and remove outlying values. The outlier detection procedures are performed at the level of individual laboratories, individual analytical methods, and groups of analytical methods, respectively. In the outlier detection procedures, a scaled value is compared with a scaled reference value, which is the scaled mean value (consensus value), or with a scaled target value, obtained from a definitive or reference method for a given component. The following algorithm is used:

\[ Z_a \text{ is considered as the most extreme outlying value } (Z_a > Z_b) \text{ if} \]

\[ |Z_a - \bar{Z}| > k_aSD \]

with \[ |Z_a - \bar{Z}| > |Z_b - \bar{Z}| ||b = 1,n \land b \neq a \]

and SD \[ = \sqrt{\sum_{b=1}^{n} (Z_b - \bar{Z})^2 / (n - 2)} \]

\( k_a \) is a constant, the value of which depends on the number of values \( n \).

In this algorithm, the value at the greatest distance from the mean value \( Z \) is omitted to compute a new mean and SD. If the omitted value lies outside the interval \( Z \pm k_aSD \), the value is marked as an outlier. The procedure is reiterated for the \( n - 1 \) remaining values. The constants \( k_a \) extracted from the table of tolerance intervals for Gaussian distributions (5), are chosen such that the chance of inadvertently marking a result as an outlier is 1%. Some typical values are \( k_8 = 77.96, k_{10} = 3.54, k_{50} = 2.71, \) and \( k_{10} = 2.57 \). The outlier detection procedure is repeated until no further outliers are detected.

The procedure is applied to the internal results as well as to the external results within individual laboratories after scaling. An outlying value is removed from the calculation of the final laboratory mean value and laboratory SD. The procedure is also applied to internal (only users of the same batch of control material) and external laboratory mean values, grouped by analytical method. All of the results of a laboratory are removed from the computation of the final mean value and SD for a method if the laboratory mean is marked as an outlier at this level. Finally, the outlier procedure is applied to laboratory mean values within a group of analytical methods (e.g., all of the direct ion-selective methods for sodium). The following list clarifies the notation of the explicit use of \( Z_a \) and \( Z \) in the different contexts of use of the global formula for outlier detection.

\[ Z_a \]

\[ Z \]

\[ SD \]

**Result population**

<table>
<thead>
<tr>
<th>( X'_i )</th>
<th>( X' )</th>
<th>( SDX' )</th>
<th>( Y'_i )</th>
<th>( Y' )</th>
<th>( SD )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal sample within lab.</td>
<td>( X'_i ) ( \times )</td>
<td>( X' ) ( \times )</td>
<td>( SDX' ) ( \times )</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
</tr>
<tr>
<td>Internal sample between labs within batch group</td>
<td>( X'_i ) ( \times )</td>
<td>( X' ) ( \times )</td>
<td>( SDX' ) ( \times )</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
</tr>
<tr>
<td>External samples within lab.</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
<td>( SD ) ( \times )</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
</tr>
<tr>
<td>External samples between labs within method group</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
<td>( SD ) ( \times )</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
</tr>
<tr>
<td>External samples between labs overall</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
<td>( SD ) ( \times )</td>
<td>( Y'_i ) ( \times )</td>
<td>( Y' ) ( \times )</td>
</tr>
</tbody>
</table>

The internal quality-control data are subjected to trend analysis. After a simple linear regression of the raw data, the significance of the deviation of the slope from zero is tested. If a statistically significant trend exists with a slope value \( > \beta \text{SD}_\text{ref} \), all of the laboratory
results are removed from the computations of method and method-group statistical parameters.

Hardware and Software

The program is written in Turbo Pascal 5.5 and run on an IBM PS/2 80 microcomputer system. The graphical presentation of results is printed on an AGFA P400 multiple light-emitting diode printer at the end of each two-month interval.

Laboratory data are transmitted electronically over the public telephone system to the SKZL computers. A dedicated data communication program, QBase, written in Turbo Pascal, is run on MS-DOS IBM-compatible personal computers. Data transmission takes place via modems at 1200 or 2400 baud, but the data are reformatted to minimize the number of bytes to be transmitted. A dedicated IBM AT microcomputer, connected in a local area network with the PS/2 80 system, supports the data communication at the SKZL office. Data are collected and processed overnight. The revised statistics are made available to the remote QBase workstations of the individual participants by the next morning.

Results and Discussion

The basic idea underlying the new combined internal/external scheme of the SKZL is to relate the information contained in the internal control materials of the laboratories to their results obtained in the external proficiency testing part of the concept. The design of the scheme is such that it can be used for general chemistry as well as hemocytometry, coagulation, endocrinology, urinary chemistry, etc. In this scheme, laboratories use internal control materials of their own choice. The external samples are sent at regular intervals, e.g., four samples monthly. In the general chemistry scheme, the laboratories should analyze one sample weekly. In the hemocytometry scheme, fresh donor samples are distributed, which should be analyzed without delay. The relatively high frequency of supplying external control material makes it possible to cover a wide range of analyte concentrations in a relatively short time. In addition, with this chosen design, fresh or fresh-frozen human samples can be distributed at regular intervals.

The apparent problem of comparing results obtained with different control materials (two internal and eight external samples) having different concentrations of analytes has been overcome by using a scaling procedure, in which all values are scaled to a common mean and standard deviation.

Control Results

Internal control results within laboratories. Results obtained for the internal control samples are processed according to the normal quality-control procedures of the laboratory. In addition, results of at most two concentrations of internal control samples are sent to the SKZL computer center. Laboratories are urged (and in the near future will be required) to make use of the SKZL data communication software to transmit their data. To participate in the Combi scheme, laboratories should send at least monthly mean values and standard deviations for the two concentrations of internal control samples. However, the more data that are sent, the more information can be extracted from them. Thus laboratories are urged to send daily values for two concentrations each, or daily mean values and standard deviations for the two concentrations. In the general chemistry scheme, controls for 25 analytes are supplied and results are reported bimonthly.

The extent to which the internal data sent to the SKZL center are evaluated statistically depends on the amount of data submitted. If daily values are submitted, the statistical analysis includes an outlier detection procedure (see Materials and Methods) and trend analysis to detect short-term (two months) and long-term (one year) drifts in the analytical process. If only monthly mean values and standard deviations are submitted, this analysis is, of course, not possible.

Internal control results between laboratories. External comparison of internal control results is possible if a group of laboratories all use the same batches of internal control materials. This may be the case for (e.g.) a regional group of cooperating laboratories that collectively purchase their internal control material or a geographically more-widespread group of laboratories that coincidentally use the same batch of a control material.

If a laboratory submits information on the batch numbers of its materials or on the regional group of laboratories in which it participates, the results from that laboratory are related to the relevant group of laboratories. A batch group mean value and batch group within-laboratory SD are computed after removal of outliers as described in Materials and Methods. This mean and SD are used as reference values in the graphical presentation of the results.

External control results within laboratories. The external control samples should be treated as unknown patients' samples. A different control sample is analyzed once weekly in the general chemistry scheme. The external control samples are selected so as to cover the concentration range of clinical interest for as many analytes as possible in each trial period. The external results are a measure of the validity of using the internal results to assess the state of the analytical process. The scaling procedure offers the possibility to estimate a laboratory mean value and within-laboratory SD from the results for the external samples. The external laboratory mean value and SD are related to the estimations obtained from the internal samples in the graphical printout.

External control results between laboratories. If possible, target values are assigned to the external samples. The target values are those obtained from reference methods for some analytes (e.g., cholesterol), or from the method mean for laboratories that use a method known to have no significant bias (e.g., atomic absorption spectrophotometry for calcium). If no such target values

CLINICAL CHEMISTRY, Vol. 37, No. 7, 1991 1199
are available, the reference point is determined by the method group mean value (e.g., the IFCC-recommended method for aspartate aminotransferase).

After an outlier removal procedure (see Materials and Methods), the mean values and average within-laboratory standard deviations are computed from the scaled external results within methods, within method groups, and overall. The standard deviations and the bias of the mean values from the target values represent the state of the art. In the graphical presentation of results, these deviations are compared with the relevant reference standard deviation, as described before.

Graphical Presentation of Results

Examples of the graphs, histograms, and tables that the laboratories receive at the end of each two months are presented in Figures 1 and 2. Figure 1 shows an example of a complete page, presenting all the relevant information for a particular analyte in six labs, designated by the large numbers. Part 1 deals with the trial and user identification. Part 2 is further elaborated in Figure 2, which shows an example of the results of the internal control samples and of the external samples in a trial. In this example, the laboratory submitted 45 results for the internal sample (represented by the open circles). Two outliers have been detected, plotted outside the white rectangle (which shows the individual laboratory’s mean ± 3 SD). The position of this rectangle, relative to the gray background rectangle (external/internal reference interval: target value(s) ± 3 SDref) shows that the laboratory mean is lower than the reference mean value. The laboratory internal SD is favorably smaller than both the batch group within-laboratory SD and the reference SD. In the two preceding trials, this laboratory obtained comparable results for this internal sample, as shown by the white bars at the left of the graph. Although the SD is relatively constant, the mean value remains about 1 SD lower than the reference mean. Also plotted in the graph are the results of the laboratory for the eight external samples in the present period (black squares with accompanying numerical values) together with the mean and 3 SD range for these samples, represented by the black bar to the right of the white rectangle. The overall consensus (or reference) mean values for the external samples, together with their average within-laboratory ± 3 SD ranges, are depicted by the dotted bars. This relates the external results to the internal values directly. The external mean values ± 3 SD of the previous two trials are shown by the black bars at the left of the graph. In this example, the external results for both the two previous trials (902, 903) and the current trial (904) are somewhat discordant with the internal results (open circles), and therefore do not confirm the image of the analytical process obtained from the internal sample. The external-method-specific mean ± 3 (within-laboratory) SD of the present trial is represented by the horizontally striped bar. Finally, the overall batch group mean value ± 3 SD is represented by the diagonally striped bar just to the left of the white rectangle of the internal results of this laboratory.

If a laboratory belongs to a group of laboratories (e.g., regional) using the same batches of internal control material, the results for the other users are presented in the right part of the graph (Figure 1, part 2, labs. 15, 25, etc.). Again, the white bars represent the internal mean ± 3 SD for each laboratory for this internal sample; the black bars represent their external results. The broad gray rectangle in the background represents both the internal batch group mean value ± 3 times the reference within-laboratory SD and the external reference interval, defined by the external target values ± 3 times the reference within-laboratory SD. This reference within-laboratory SD is defined by the average target value and the clinically desired analytical variance. The percentage of results that is expected to fall within this range, based on the external results for the laboratory, determines the score for the analyte of concern for this laboratory.

A similar graph (Figure 1, part 3) shows the results for the second internal control material, together with the results for the external samples. This makes obvious any discrepancies between the results for the two internal controls and the external samples.

Part 4 shows the frequency histograms of the results of all laboratories for each of the external samples (A-H). The first set of eight histograms highlights (light shading) the results from laboratories using the same method. The second set of eight histograms highlights the distribution of the external results from the group of laboratories using the same internal samples. The laboratory results, represented by the black boxes "hanging" below the baseline, show whether the laboratory results are biased towards the target values (hairlines with accompanying numerical values), the batch group laboratories, or the method.

Part 6 shows the consistency of the laboratory results along the concentration range of the external samples (A-H, as above). In case of constant bias, all of the results will be above or below the central target line. In case of nonlinearity, results in the higher or lower concentration region will deviate from the target.

Finally, part 5 shows the laboratory results in tabular form. As described in Materials and Methods, the internal values are scaled to their respective consensus values, and the external results are scaled to the average target value. The small scoring icon indicates the position of the laboratory score relative to all other participants, the score being based on the percentage of external results that are within the target value ± 3 SDref (the clinically desirable analytical SD). In the example in Figure 1, about 40% of the laboratories had a higher score for the analyte of concern than that of the laboratory depicted here. If the score of a laboratory is below a predefined value determined for each analyte (e.g., 95% for sodium in serum), the laboratory is judged to fail the required standard in the trial. Thus an objective measure is introduced, which may be used by a certification or accreditation agency. In the example
Fig. 1. Example of one complete report page with graphical and tabular information

Part 1: trial and user identification. Part 2: graphical presentation of internal (Level 1) and external results. Left side of dotted vertical line: see Fig. 2. Right side of dotted vertical line: white rectangles = internal mean values ± 3 SD of labs. using same internal control material; black rectangles = external mean values ± 3 SD of labs. using same internal control material. Part 3: same as part 2 for Level 2 of internal control sample. Part 4: frequency histograms of the results of all laboratories for the external samples. Lightly shaded areas represent laboratories using the same internal control material (upper histograms), or laboratories using the same method (lower histograms). Part 5: table with internal and external results. The scoring icon shows the position of the score of the particular participant relative to the other laboratories. In the example, 40% of the laboratories score better. The gray area denotes the percentage of laboratories (20%) scoring below a predefined (certification) limit. Part 6: bias of individual external results with respect to target values.
shown, 20% of the participants fail to meet the criteria, as denoted by the gray area printed in the scoring icon.

Practical Experiences with the Program

To give insight into the practical usefulness of the program, we include a few recent examples of participant presentations and discuss the following three possible situations:

- Discrepancies between external and internal results.
- Trends experienced both in the internal and external results.
- Problems due to nonlinear behavior of the analytical methodology.

Figure 3 shows the results for sodium during September/October 1990. The internal results in no way lead one to expect any potential problem; whereas the external results show a definite positive drift during the first six weeks of the period, ranging from -1.8 to +3.2 mmol/L. Analyses for potassium showed identical behavior. During week 6 the flame photometer was serviced, which probably accounts for the lack of further increases in the external results. On a regional level, the results for the laboratory depicted are in reasonable agreement with the sodium results from the other laboratories, although the individual precision for this analyte is rather widespread.

Figure 4 deals with the results for alkaline phosphatase. For several months a positive trend for both the external and internal results has been discernible (see black and white bars for the preceding trial periods). In contrast, the current period (905) shows a trend toward decreasing values for both the internal and external values. This also leads to the poor precision relative to the reference range. This laboratory has thus far not been able to determine the cause of the long-term fluctuations.

The situation in Figure 5 illustrates some inherent problems with the calcium determinations. For calcium analyses in The Netherlands, the target value is the trimmed overall consensus mean. Because of the relative overrepresentation of chromogenic methods, this consensus mean also reflects falsely high calcium values at the high concentration range, a drawback of chromogenic methods. The laboratory depicted uses a semi-automated titration technique with ethylene glycol bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid, and therefore presents too low results for the calcium values in the higher concentration range. This effect is also reflected by the relative position of the two distinctive internal controls. The results for Level I control (upper panel) lie much more within the reference region (bias = -0.6 SDref) than those for Level II (bias = 2.0 SDref).

This short compilation of practical possibilities could be amplified with many other examples. They are pre-
Combi trial 906
Trial date: November - December 1990
Print date: January 8, 1991
Samples: Precinorm 162975 - ZOB

Fig. 5. Demonstration of nonlinearity in the calcium determination in the high-concentration range
Effects are discernible both in the external and internal results
sented here merely to demonstrate the potential power of this integrated program as a unified package for quality assessment.

Data Communication and Timeframe of Data Logistics

The large number of data submitted with each trial necessitates the use of electronic data transmission. A dedicated program, QBase, has been developed to enable two-way communication between the SKZL office computers and the PCs of the individual participants over the public telephone system. An Xmodem-like protocol maximizes the integrity of the transmitted data. The batchwise communication protocol and the partial encryption of user identification prohibit intruders from breaking into the system. Unauthorized copying of the software will enable that user only to transmit quality-control results with the identification number of the participant that made the software available. The regular distribution of new versions of QBase with changing version numbers and personal user identification will discourage the unauthorized use of QBase.

The communication program QBase allows the individual participant to enter the data obtained for the internal and external control materials. The program not only enables the transmission of quality-control results from the individual participant to the SKZL office, but also offers an updating of the participants' semi-permanent listings of method, reagent, and apparatus codes. The communication also supports the two-way exchange of messages. Whenever a participant's microcomputer contacts the SKZL computer, the method, reagent, and instrument tables are updated. Any messages are dumped in message files that may be viewed from the QBase program.

The first version of QBase that was officially distributed (version 1.03) required the more or less cumbersome manual input of quality-control results. Later versions (1.04 and later) allow the on-line retrieval of quality-control files already present in laboratory information systems. The latest version (2.01) allows the on-line retrieval of the statistics for a large number of internal control material data, as well as the method-specific statistical data of the external proficiency testing samples after the closure date of the trial, from the central SKZL computer system.

Results can be sent to the SKZL office any day of the week. However, the statistical processing for the printed reports is activated once every two months. For the intermediate retrieval of statistics by the QBase users during the second month of each trial period, the statistics are updated each night. Future versions will also support local initial statistical treatment of each participant's internal quality-control results.

In conclusion, the SKZL Combi scheme combines the information content of internal and external control results. Because no demands are made on the origin of the internal control materials, no interlaboratory matrix effects are to be expected, which would appear if a common internal control material had been obligatory. In addition, the design of the scheme allows the use (from time to time) of fresh human material as external samples. The almost complete graphical presentation of results, prepared with a high-resolution printer, facilitates the interpretation. Laboratory performance data for internal control analyses are related to users of the same batches of internal control material. The laboratory results for the external samples are related to the same group of laboratories, to laboratories using the same method, and to the national consensus mean and within-laboratory SD. In addition, a reference range is presented, defined by the average external target value and the clinically desirable analytical SD. Finally, a scoring system, based on the reference range, rates the laboratory performance objectively. By listing the scoring rank of the laboratory and indicating the percentage of laboratories scoring below a predefined minimally allowable limit, an objective tool for certification is introduced. The dedicated computer program, QBase, can be linked to laboratory information systems to prevent duplicate data input. To supply the quality-assurance officer with information on materials other than those used by the laboratory itself, the data bank of the SKZL can be consulted on-line.

References