

Eur. J. Clin. Chem. Clin. Biochem.
Vol. 29, 1991, pp. 675–683

© 1991 Walter de Gruyter & Co.
Berlin · New York

Practical Implications of Coexistent Different Technologies in Clinical Chemical Laboratories

Solid Phase Chemistry and Conventional Analysis

By J. Kropf, A. M. Marx, J. Hildebrandt and A. M. Gressner

Department of Clinical Chemistry, Philipps University, Marburg, Germany

(Received February 28/June 24, 1991)

Summary: If different analytical methods are alternatively used for the determination of the same analyte, basic differences in test methodology can give rise to an increased number of deviating results. Such coexistence of methods might be necessary, for example, during a transition phase while upgrading to new technologies. We have exemplarily investigated this topic for the comparison of solid phase chemistry ("dry chemistry") versus conventional methods ("wet chemistry"). The Kodak Ektachem 700XR clinical chemistry analyser was compared with the Hitachi 737 analyser from Boehringer Mannheim using 18 clinical chemical analytes and specimens submitted for routine analysis. Before the start of the evaluation, the Ektachem 700XR was adjusted ("calibrated") by the manufacturer for optimal agreement with the Hitachi 737.

Satisfactory agreement was obtained for most investigated analytes as judged by correlation coefficients and three commonly applied regression methods (linear regression, principal components, and *Passing/Bablok* method). For some analytes, however, strongly deviating results were often obtained. Quality control-derived limits (maximum acceptable inaccuracy) and data from biological variation (critical differences) were used for the assessment of the inter-instrument bias for diagnosis and patient monitoring, respectively. For enzymes, 0% (amylase) to 22% (creatine kinase) of all pathologic results differed by more than the maximum acceptable analytical inaccuracy (21%–27%) of these analytes. If more stringent limits derived from biological variation were used, 24% (creatine kinase) – 62% (aspartate aminotransferase) of all differences between paired measurements exceeded the critical difference for enzymes. Deviations greater than the critical differences were also marked for serum concentrations of sodium, calcium, and creatinine.

We conclude that the simultaneous use of analysers based on different analytical methodologies, i.e. solid state chemistry and conventional methods, is generally problematic, especially for purposes of patient monitoring.

Introduction

Analysis using carrier-bound reagents ("solid phase chemistry", "dry chemistry") is an evolving technique, rapidly gaining acceptance in several areas of clinical chemistry (1–3). Advantages over conventional analysis using liquid reagents ("wet chemistry") are (among others) high stability of reagents and calibration curves, rapid analysis, and the easy operation of analytical devices (4).

Early instrumentation was focussed on small devices for bedside analysis or testing in the consulting room, but analysers for the large clinical laboratory are also available. It can therefore be anticipated that dry chemistry technology and traditional wet chemistry methods will – at least during a transition phase – coexist. This coexistence will occur between laboratories, i.e. some laboratories having integrated dry chemistry analyzers, while others use conventional

chemistry alone, and also within single laboratories. This is because economy dictates the continuous use of functionally still reliable equipment, while investments in new instruments eventually will favour dry chemistry-based devices. In the present paper, we wish to concentrate on the latter aspect. The central question can thus be formulated as:

What are the consequences (regarding the medical interpretation of laboratory results) if analytical results are determined (probably within short intervals between determinations) sometimes using solid phase chemistry and sometimes using conventional chemistry?

We have *not* investigated this topic in its full generality, but have taken an exemplary, pragmatic approach by considering two devices, the Hitachi 737 and the Kodak Ektachem 700XR, which we think can serve as representative models for the study of their intrinsic analytical techniques. We have further reduced generality by focussing on analytical principles adapted for German laboratories, which differ, for example, from the respective procedures used in the United States.

One can expect to observe discrepancies caused by different sensitivities to interferences like drugs, or properties of the specimen like viscosity (5). The conversion of enzyme activities from 37 °C (Ektachem 700XR) to 25 °C (Hitachi 737) and different reaction conditions (e. g. buffers) will lead to biased results, if the temperature response is non-linear or if isoenzyme patterns vary (6, 7).

This study is *not* concerned with the establishment of ranks between the two analytical methodologies or the two instruments used as representatives for these methodologies. We have not attempted to evaluate the "true" value of two diverging results by use of, e. g., a definitive or reference method, but have restricted our investigation to reporting the magnitude and frequency of differences found for samples from the daily routine analysis in our laboratory. The reported figures are intended to serve as a guideline for the discrepancies one has to expect, if solid phase chemistry and conventional chemistry are applied simultaneously.

Methods

Devices

A Hitachi 737 analyser (Boehringer Mannheim, Mannheim, Germany) which is used for routine analysis in our laboratory was compared with a Kodak Ektachem 700 XR analyser (Kodak AG, Stuttgart, Germany). For comparison between two conventional analytical methods the results of the Hitachi 737 were compared with those obtained by a Hitachi 717 analyser.

Both Hitachi analysers were calibrated daily before start of routine analysis.

The Ektachem was calibrated once and adjusted for optimal agreement with the Hitachi 737 by the manufacturer before the start of the evaluation. During an initial phase both instruments were operated in parallel. Adjustment was made by two point factorization of the Ektachem 700XR based on linear regression of the collected data. Approximately 80–100 paired samples were utilized for the regression of each analyte. Strongly deviating results ("outliers") were omitted from the computations. As stated by the manufacturer, the calibration was reliable for all analytes during the whole evaluation period.

Analytical imprecision was determined by analysis of appropriate control materials after every tenth sample. For the Ektachem 700XR two serum pools were used, while Kontrollorgen LU (Behringwerke AG, Frankfurt, Germany) was used for the Hitachi 737. Accuracy was monitored with Kodatrol I or Kodatrol II (Kodak) for the Ektachem 700XR, and with Monitrol I (Baxter Deutschland GmbH, Unterschleißheim, Germany) and Kontrollorgen LP (Behringwerke) control sera for the Hitachi analysers.

Analytes

Visually normal samples from sera submitted for routine analysis were used for all comparisons. Measurements were performed in batches consisting of ≈ 20 specimen on both analysers on consecutive days at the same time. The results from both analysers were compared immediately, and if judged necessary, the measurement was repeated to exclude outliers caused by severe instrumental malfunction or sample mix up. Because all suspect results were confirmed within reasonable limits by the respective analyser, we generally used the values of the first determination for further investigations. This argument was not valid for turbidimetric lipase analysis with the Hitachi 737. The high frequency (8.5%) of obviously false, negative enzyme activities (8, 9) caused us to omit this analyte from further evaluation. A total of 18 different clinical chemical analytes was investigated. Apart from their principally different analytical basis, different chemical methods are used for some analytes by the Hitachi 737 and the Ektachem 700XR:

Enzyme activities are measured at 25 °C by the Hitachi 737 according to modified recommendations of the Deutsche Gesellschaft für Klinische Chemie, while the Ektachem 700XR performs these measurements at 37 °C according to IFCC guidelines. Moreover, alanine aminotransferase¹⁾ and aspartate aminotransferase¹⁾ are determined in presence of pyridoxal 5-phosphate, while this additive is missing in the corresponding methods of the Hitachi 737. Sodium and potassium are determined potentiometrically with ion-sensitive electrodes by both instruments, but a dilution of the sample is used by the Hitachi 737, while the Ektachem 700XR measures the undiluted sample directly. Further differences (Ektachem 700XR / Hitachi 737) exist with respect to enzyme substrates (amylase: dyed starch / *p*-nitrophenylmaltoheptaoside, γ -glutamyltransferase¹⁾: *L*- γ -glutamyl-3-carboxy-*p*-nitroanilide / *L*- γ -glutamyl-*p*-nitroanilide), the determination of glucose (glucose oxidase / hexokinase), creatinine (creatinine amidohydrolase / Jaffe), and calcium (arsenazo III / *o*-cresolphthalein).

Statistical methods

The statistical methods of *Passing & Bablok* (10, 11), linear regression (12), and principal component analysis (13), respec-

¹⁾ Enzymes: alanine aminotransferase (EC 2.6.1.2); alkaline phosphatase (EC 3.1.3.1); aspartate aminotransferase (EC 2.6.1.1); γ -glutamyltransferase (EC 2.3.2.2); lactate dehydrogenase (EC 1.1.1.27)

tively, were applied for method comparison for all investigated analytes.

For the assessment of deviations we computed the frequencies of relative differences d_i between values obtained from the same sample x_i by both analysers:

$$d_i = \frac{x_i^H - x_i^K}{x_i^H},$$

where H denotes Hitachi 737 and K Kodak Ektachem 700XR.

As discussed below, clipped ranges of deviations were used in some instances.

For computations the software packages Statgraphics (STCS Inc., Rockville, MD), SAS (SAS Institute Inc., Cary, USA), and Methodenvergleich (Boehringer) were used.

Results

Method comparisons

The between-run imprecisions of both instruments during the evaluation are given in table 1. The results of the method comparisons are shown in table 2 and exemplarily for serum creatinine in figure 1. Except for sodium and calcium, the coefficients of correlation are close to 1. Approximately the same slopes and intercepts were obtained by the three statistical methods. Differences in the tests for significance emerge as a consequence of the different assumptions of the respective statistics.

Tab. 1. Between-run analytical imprecisions for the Ektachem 700XR and the Hitachi 737. Two serum pools with pathological (high) and normal (low) concentrations (given in parenthesis) were used for the Ektachem 700XR.

| Analyte | Ektachem 700XR CV [%] | | Hitachi 737 CV [%] | |
|-------------------------------|--------------------------|------------|-----------------------|------------|
| | Low | High | Low | High |
| Sodium [mmol/l] | 1.2 (143) | 1.5 (148) | 1.2 (138) | 1.8 (168) |
| Potassium [mmol/l] | 1.2 (4.16) | 1.1 (5.0) | 1.2 (4.24) | 1.7 (6.29) |
| Calcium [mmol/l] | 1.7 (2.47) | 1.3 (2.38) | 4.0 (2.25) | 2.4 (3.2) |
| Glucose [mmol/l] | 0.9 (5.94) | 1.3 (8.16) | 4.2 (4.05) | 1.9 (13.5) |
| Creatinine [μ mol/l] | 4.7 (71) | 0.9 (539) | 4.2 (106) | 2.2 (293) |
| Protein [g/l] | 1.9 (76) | 2.1 (66) | 2.0 (65) | 2.5 (58) |
| Urea [mmol/l] | 1.5 (4.83) | 1.0 (22.3) | 3.3 (5.16) | 3.2 (16.2) |
| Triacylglycerols [mmol/l] | 1.2 (1.39) | 1.6 (4.42) | 1.7 (1.07) | 2.8 (1.15) |
| Bilirubin [μ mol/l] | 0.0 (8.55) | 1.4 (75.2) | 3.2 (18.8) | 1.6 (44.5) |
| Cholesterol [mmol/l] | 1.8 (5.65) | 1.1 (5.46) | 3.1 (4.92) | 4.3 (3.06) |
| Uric acid [μ mol/l] | 1.1 (244) | 1.3 (488) | 2.5 (303) | 3.5 (521) |
| Amylase [U/l] | 1.9 (57) | 1.5 (178) | 3.4 (60) | 2.7 (244) |
| Aspartate aminotransferase | 10.8 (9) | 0.5 (250) | 3.1 (34) | 2.4 (85) |
| Lactate dehydrogenase | 3.5 (121) | 2.1 (729) | 3.2 (222) | 5.3 (419) |
| Alkaline phosphatase | 2.8 (102) | 1.9 (569) | 8.7 (76) | 4.8 (430) |
| γ -Glutamyltransferase | 4.1 (12) | 1.6 (253) | 5.7 (16) | 5.9 (115) |
| Creatine kinase | 2.0 (49) | 1.2 (296) | 3.3 (51) | 2.7 (196) |
| Alanine aminotransferase | 6.8 (7.5) | 0.5 (110) | 2.5 (26) | 3.2 (87) |

Tab. 2. Method comparison between Ektachem 700XR and Hitachi 737. Results of the regression procedures are given as slope \pm intercept.

| Analyte | N | r (Pearson) | Linear regression | Passing/Bablok | Principal components |
|-------------------------------|-----|----------------|----------------------------------|----------------------------|-------------------------|
| Sodium | 168 | 0.874 | 0.775 ^a + 32.9 | 0.857 + 21.7 | 0.887 + 17.3 |
| Potassium | 174 | 0.975 | 0.989 - 0.01 | 1.000 + 0.00 | 1.014 - 0.12 |
| Calcium | 188 | 0.887 | 0.999 - 0.09 | 1.111 - 0.36 | 1.127 - 0.40 |
| Glucose | 167 | 0.997 | 1.033 - 0.40 | 1.033 - 0.40 | 1.037 - 0.43 |
| Creatinine | 189 | 0.996 | 0.932 + 1.94 | 0.901 + 5.22 | 0.936 + 1.41 |
| Protein | 190 | 0.975 | 0.912 + 7.31 | 0.925 + 6.48 | 0.935 + 5.70 |
| Urea | 185 | 0.998 | 0.907 + 0.14 | 0.889 + 0.29 | 0.909 + 0.12 |
| Triacylglycerols | 179 | 0.984 | 0.949 + 0.05 | 0.948 + 0.04 | 0.964 + 0.02 |
| Bilirubin | 201 | 0.995 | 1.108 - 2.17 | 1.000 \pm 0.00 | 1.114 - 2.33 |
| Cholesterol | 160 | 0.986 | 0.916 + 0.31 | 0.919 + 0.30 | 0.929 + 0.24 |
| Uric acid | 160 | 0.983 | 0.909 + 13.7 | 0.963 + 3.39 | 0.924 + 8.98 |
| Amylase | 118 | 0.993 | 0.975 + 5.33 | 1.063 - 2.22 | 0.982 + 4.62 |
| Aspartate aminotransferase | 168 | 0.995 | 0.961 + 3.58 | 0.996 + 2.15 | 0.966 + 3.37 |
| Lactate dehydrogenase | 196 | 0.981 | 0.874 + 4.51 | 0.886 - 0.12 | 0.891 - 1.04 |
| Alkaline phosphatase | 212 | 0.986 | 0.967 - 10.1 | 0.929 - 3.40 | 0.980 - 13.1 |
| γ -Glutamyltransferase | 172 | 0.993 | 0.998 - 5.17 | 1.011 - 5.42 | 1.005 - 5.68 |
| Creatine kinase | 81 | 0.999 | 0.979 - 8.87 | 0.900 + 0.30 | 0.981 - 9.12 |
| Alanine aminotransferase | 168 | 0.995 | 0.961 + 3.58 | 0.996 + 2.15 | 0.966 + 3.37 |

^a Boldface: significant difference ($p < 0.05$) of slope or intercept from 1.0 or 0.0, respectively

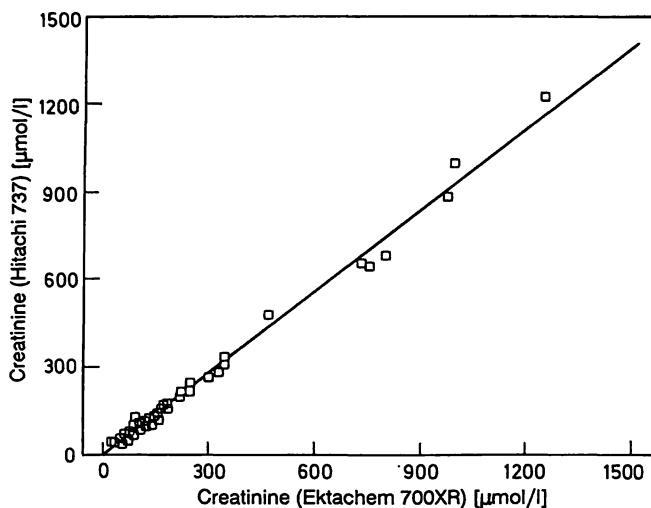


Fig. 1. Linear regression of the analysis of serum creatinine by the Hitachi 737 versus Ektachem 700XR. Regression data are given in table 2.

Significant deviations with respect to slope or intercept were observed for most analytes (especially striking: sodium, aspartate aminotransferase and lactate dehydrogenase). Our results are, however, comparable to those published by others (3, 14), who judge overall conformity to be satisfactory.

Frequencies of discrepant results

Inter-instrument bias related to analytical inaccuracy

Desirable goals for analytical errors have been related to biological variation (15). According to *Tonk's* rule (16), for example, the allowable limits of error should not exceed $\frac{1}{4}$ of the reference interval. These concepts have been applied to analytical inaccuracy by *Stamm* (17) and have recently been adopted as criteria for the establishment of tolerance limits in intra-laboratory quality control in German laboratories (18). In short, the maximum acceptable percental *inaccuracy* is three times the maximum coefficient of variation (CV) from day to day which is defined by $\frac{1}{4}$ of the respective reference range. For sodium, e.g., with a reference interval of 135–144 mmol/l, this gives:

$$\frac{144 - 135}{4} = 2.25 \text{ mmol/l.}$$

As a percentage of the reference range (with some rounding off) the maximum CV is computed as:

$$100 \times \frac{2.25}{139.5} = 1.61 \approx 2\%.$$

Further, the maximal percent *inaccuracy* of a single result is three times this maximal CV, which in the case of sodium will give 6%.

This inaccuracy can be interpreted as the maximum deviation from the true value the clinician has to take into account if she/he compares patient results with the reference range of the analyte. The application of these limits as cut-off values for the assessment of deviating results is aimed at demonstrating the number of excessively high differences which can lead to possibly wrong diagnostic decisions.

Because our primary interest is in diagnostically relevant inter-instrumental bias, only determinations with at least one of both results beyond the reference interval (sodium, potassium, calcium, glucose, protein) or above the upper reference limit (enzymes, creatinine, urea, triacylglycerols, bilirubin, cholesterol, uric acid) of the respective analyte were considered for the computation of the frequencies of differences. Unfortunately, for some analytes (potassium, calcium, protein, creatine kinase), there was only a small number of pathological results, thereby lowering the accuracy of the respective frequencies.

Table 3 shows the relative frequencies of differences between paired measurements which exceed the quality control-derived limits. For most metabolites and electrolytes these frequencies are small. However, for potassium, protein, and most enzymes considerably high frequencies (up to 22% of all investigated samples) were found.

As shown exemplarily in figures 2a–d for some analytes, these differences were not distributed perfectly symmetrically with respect to zero. This might be explained by a systematic offset of one instrument. We have not corrected for such bias, using e. g. linear regression, because

- (i) both analysers had been adjusted optimally before start of the evaluation,
- (ii) the calibration of the Ektachem was stable during the whole evaluation phase, and
- (iii) the day-to-day imprecision of both analysers was also satisfactorily during this phase.

The elimination of systematic and proportional bias by use of linear regression would of course have decreased the frequencies of deviating results, but using the evaluation-data themselves for such "re-calibration" would produce both over-optimistic and unrealistic results. Under routine conditions it would make no sense to perform *all* measurements on both analysers. On the other hand, if only some samples were randomly selected for the adjustment, the results of just these samples might by chance be strongly divergent, leading to even worse overall performance.

Tab. 3. Comparison of Ektachem 700XR vs Hitachi 737. Relative frequencies of differences beyond tolerance limits are given for the inaccuracy criterion and the critical difference criterion.

| Analyte | Relevant for diagnosis | | | Frequency | Relevant for monitoring | |
|-------------------------------|-----------------------------------|------------------------------|-------|-----------|----------------------------------|-----------|
| | Inaccuracy criterion ^a | Reference range ^b | | | Critical difference ^c | Frequency |
| Sodium | 0.06 | 135 – 144 | (50) | 0 | 0.02 | 0.49 |
| Potassium | 0.08 | 3.6 – 5.1 | (15) | 0.13 | 0.13 | 0 |
| Calcium | 0.10 | 2.2 – 2.6 | (44) | 0.07 | 0.05 | 0.33 |
| Glucose | 0.15 | 3.9 – 6.1 | (69) | 0 | 0.38 | 0 |
| Creatinine | 0.18 | < 106 | (50) | 0.06 | 0.10 | 0.43 |
| Protein | 0.09 | 66 – 83 | (74) | 0.07 | 0.07 | 0.14 |
| Urea | 0.24 | < 7.65 | (76) | 0 | 0.14 | 0.04 |
| Triacylglycerols | 0.21 | < 2.28 | (55) | 0 | 0.40 | 0 |
| Bilirubin | 0.21 | < 3.4 | (60) | 0.05 | 0.38 | 0.02 |
| Cholesterol | 0.18 | < 6.47 | (51) | 0 | 0.08 | 0.15 |
| Uric acid | 0.18 | < 416 | (26) | 0 | 0.15 | 0.02 |
| Amylase | 0.27 | < 130 | (24) | 0 | 0.12 | 0.27 |
| Aspartate aminotransferase | 0.21 | < 17 | (83) | 0.15 | 0.13 | 0.62 |
| Lactate dehydrogenase | 0.21 | < 240 | (106) | 0.06 | 0.14 | 0.28 |
| Alkaline phosphatase | 0.21 | < 190 | (73) | 0.12 | 0.08 | 0.59 |
| γ -Glutamyltransferase | 0.21 | < 28 | (103) | 0.04 | 0.16 | 0.54 |
| Creatine kinase | 0.24 | < 80 | (9) | 0.22 | 0.23 | 0.24 |
| Alanine aminotransferase | 0.21 | < 23 | (71) | 0.14 | 0.15 | 0.48 |

^a Maximal allowable relative deviation for analytical inaccuracy according to the quality control guidelines for German laboratories

^b Number of cases outside the reference range are given in parenthesis

^c Change in serial results from the same patient necessary for statistical significance ($P \leq 0.05$)

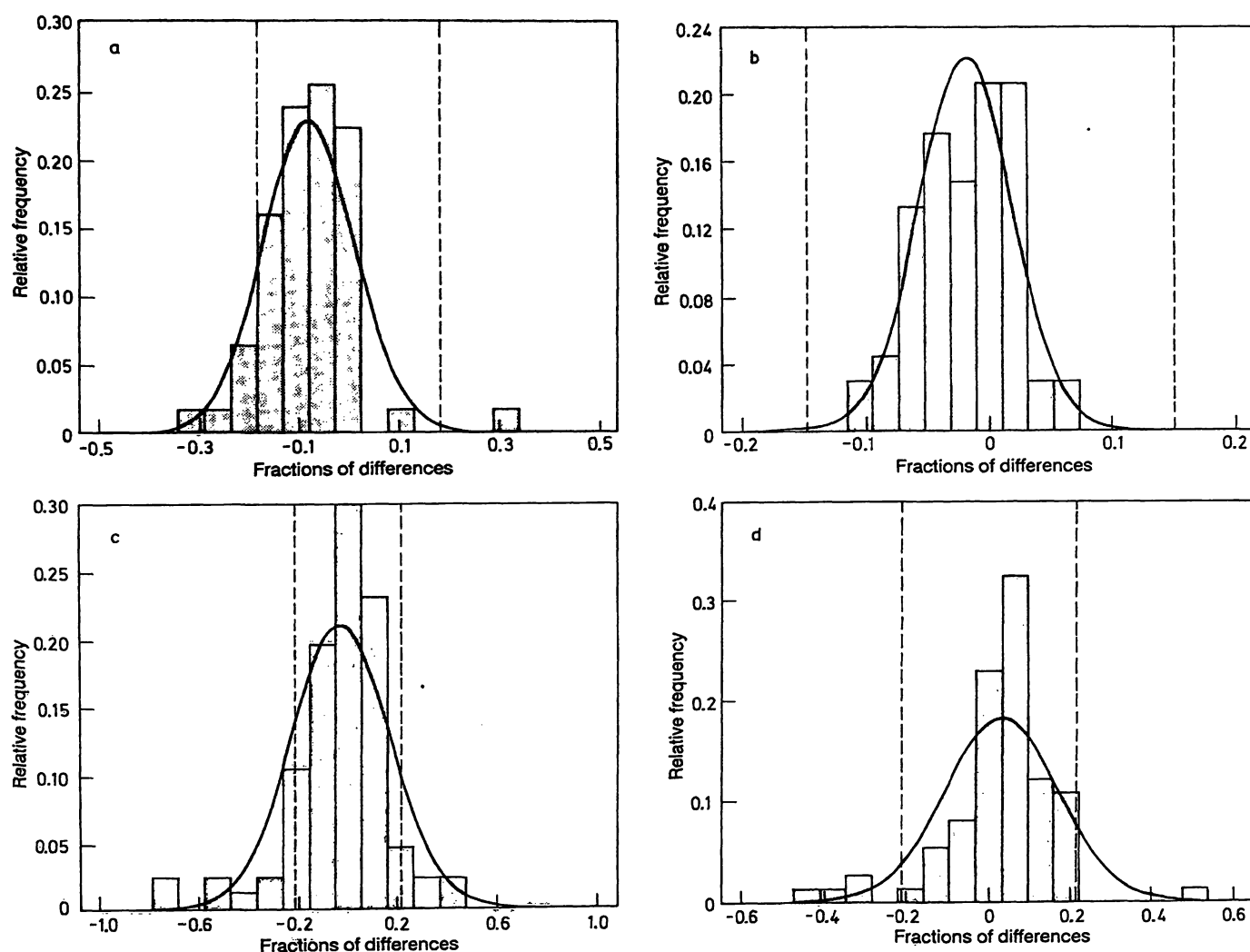


Fig. 2. Frequency distributions of differences between results from Hitachi 737 and Ektachem 700XR. Only pathological values were considered. The dotted lines indicate the maximum acceptable inaccuracy limits. a: Creatinine, b: Glucose, c: Aspartate aminotransferase, d: Alanine aminotransferase.

Inter-instrument bias related to critical differences

The concept of the critical difference can be used to follow the course of laboratory parameters in one individual in consecutive measurements (19). The critical difference d_K is dependent on the intra-individual variance of serial measurements in one individual and is given with $p < 0.05$ by $d_K = 2.8 \sqrt{s_i^2 + s_a^2}$, where s_i and s_a are the intra-individual and analytical variances, respectively. For duplicate measurements of the same sample the intra-individual component of d_K is zero and any difference between determinations is due to the analytical variance only. As a conservative criterion we assume that the analytical variance should not produce differences exceeding the critical difference given by biological variation alone, i.e. the difference between measurements on both analysers should be less than $2.8 \times s_i$ or $2.8 \times CV_i$, if the coefficient of variation is used.

In table 3, the relative frequencies of differences \geq the critical differences are shown for all investigated analytes. Data on intra-individual variances during one day were taken from Costongs et al. (19). Since patient monitoring is not necessarily restricted to the observation of pathological values, results within reference ranges were also considered in the computations.

With respect to the critical differences, the frequencies of deviating results were generally higher than those obtained using the inaccuracy criterion. For enzyme activities they range from 0.24–0.64. Additionally, the frequencies for sodium, calcium, and creatinine (fig. 3) were much higher than those obtained using the inaccuracy criterion (fig. 2a).

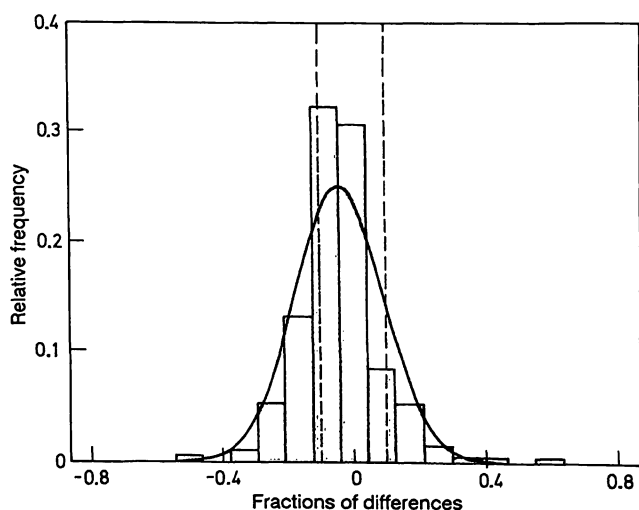


Fig. 3. Frequency distribution of differences between creatinine results from Hitachi 737 and Ektachem 700XR. All values were considered. The dotted lines indicate the limits defined by critical difference.

Inter-instrument differences between two conventional analysers

To investigate whether the reported discrepancies are due to the analytical methodology, we additionally compared two instruments employing identical methods. Restricted to five of the most deviating analytes, this evaluation was performed as described above for the comparison between Ektachem 700XR and Hitachi 737. The results are given in table 4 and exemplarily for serum creatinine in figure 4a–b. With respect to the inaccuracy criterion, there was high conformity between the two instruments. Much higher figures were observed using the criterion of critical differences. But, except for the analysis of creatinine, these fractions were only about 50% of those found for the former comparison.

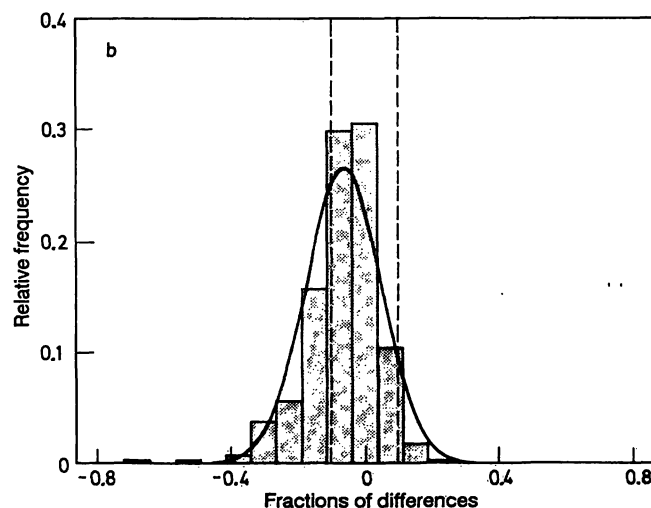
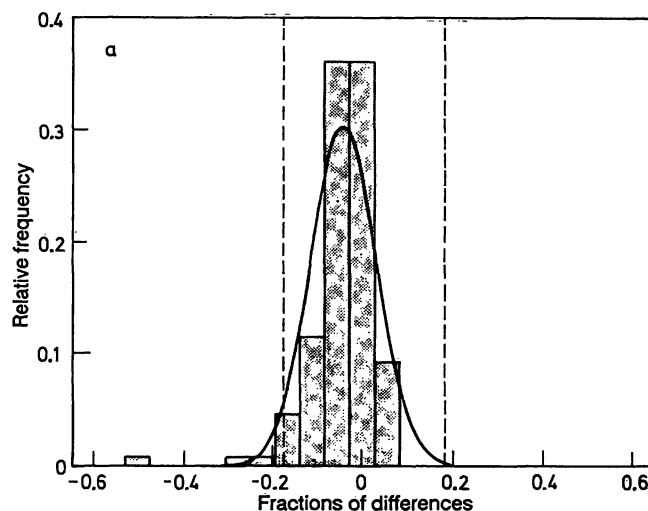


Fig. 4. Frequency distributions of differences between creatinine results from Hitachi 737 and Hitachi 717. The dotted lines indicate the maximum acceptable inaccuracy limits (a) and limits defined by critical difference (b), respectively.

Discussion

This paper was initiated by experience of the introduction of an analyser based on dry chemistry sheets into our laboratory, which otherwise employed solely conventional wet chemistry-based analysis. The plan was to use a Kodak Ektachem 700XR in the emergency laboratory in parallel with two Hitachi 737 machines for routine analysis.

After adjustment of the Ektachem 700XR to the currently used analysers (see above), satisfactorily correlation and regression data were found. The results shown in table 2 are in agreement with the data found by others (3, 14). Also with respect to internal quality control, both techniques give similar results, as shown in table 1. Fulfillment of these criteria is commonly accepted for sufficient comparability between methods (20, 21). Nevertheless, during parallel use a notable number of discrepant results occurred which were reproducible and thus could not be explained by chance alone. This therefore raised the question of the frequency of medically relevant differences (15) caused by the coexistence of diverse analytical methodologies.

In general, physicians request laboratory tests for two reasons: to assist in diagnosis or to monitor patients (22, 23). Because inter-instrument bias will have different effects in these two areas, we established two sets of criteria for the assessment of excessive differences between both instruments.

For diagnosis, analytical results should agree with respect to decisions based on established reference intervals. As described above, maximal tolerable inaccuracy in German laboratories is defined by $\frac{1}{2}$ of the respective reference ranges (18). This criterion could not be achieved for 10 out of 18 investigated analytes (tab. 3). If routinely applied for analysis the outcome of this "inaccuracy" would lead to diverse classifications of pathological and normal cases. For example, only 74 out of 84 sera with elevated aspartate

aminotransferase activities were classified as pathological by both methods (results not shown). Although an attempt was made to adjust the Ektachem 700XR to the Hitachi 737 for optimal agreement of the results, systematic bias was observed (fig. 2a–d). Obviously, for some analytes the adjustment of the Ektachem 700XR was performed incorrectly or at least suboptimally by the manufacturer. For potassium, creatinine, protein, and lactate dehydrogenase a recalibration would have led to clearly decreased frequencies of discrepant results, which should be apparent from the respective figures in table 3. We have, nevertheless, reported the frequencies obtained during the original adjustment by the manufacturer, because this procedure is generally performed during installation of the Ektachem 700XR and an extensive evaluation followed by eventual recalibration might not necessarily be undertaken by all laboratories.

For monitoring of patient laboratory results, differences between consecutive measurements can be assessed by comparison with the intra-individual variation of a given analyte (24, 25). We calculated the frequencies of significant deviations according to the concept of critical differences. Most analytes are effectively controlled by homeostatic mechanisms, leading to smaller fluctuations within one individual than in population-based reference ranges. Therefore, clearly elevated frequencies of results exceeding the critical difference criterion could be expected (tab. 3). These frequencies are further increased by the fact that normal results were considered in the calculations. It should be stressed, however, that under certain conditions patient monitoring within the reference range is necessary, e. g. glucose concentration in diabetics, creatinine concentration after kidney transplantation, or control of infusion therapy. For 6 analytes more than 40% of all results differed by amounts greater than the respective critical difference!

Found discrepancies can be explained at least in part by differences in analytical methodology. Measure-

Tab. 4. Comparison of Hitachi 717 vs Hitachi 737. Relative frequencies of differences beyond tolerance limits are given for the inaccuracy criterion and the critical difference criterion.

| Analyte | Inaccuracy criterion ^a | Reference range ^b | Frequency | Critical difference ^c | Frequency |
|----------------------------|-----------------------------------|------------------------------|-----------|----------------------------------|-----------|
| Creatinine | 0.18 | 106 (115) | 0.03 | 0.10 | 0.37 |
| Aspartate aminotransferase | 0.21 | 17 (83) | 0 | 0.13 | 0.33 |
| Alkaline phosphatase | 0.21 | 190 (120) | 0 | 0.08 | 0.26 |
| Creatine kinase | 0.24 | 80 (45) | 0 | 0.23 | 0.15 |
| Alanine aminotransferase | 0.21 | 23 (78) | 0 | 0.15 | 0.17 |

^a Maximal allowable relative deviation for analytical inaccuracy according to the quality control guidelines for German laboratories

^b Number of cases outside the reference range are given in parenthesis

^c Change in serial results from the same patient necessary for statistical significance ($P \leq 0.05$)

ment of enzyme activities at different temperatures is cumbersome, since biased total activities will emerge from differences in the temperature-dependence of isoenzymes. In addition, variation of substrates or cofactors, e. g. supplementation of the assay sample with pyridoxal 5-phosphate (aspartate aminotransferase, alanine aminotransferase), can alter the outcome of the analysis in certain patients (7, 26). Owing to different sensitivity to interference, this is also true if quite different methods are applied, e. g. *Jaffe* and creatinine amidohydrolase in the analysis of creatinine (27–30). Differences in the matrix effects on dry chemistry sheets and in conventional analysis are well known (3, 5, 7, 14, 31, 32). Since most of these factors show no systematic influence, a general correction by "factorization" is impossible. This explains diverse results in spite of optimal adjustment of the instruments by linear regression. Only if identical methods are used can an optimal adaptation of instruments be expected, as further demonstrated by the higher coincidence of the Hitachi 737 / Hitachi 717 results (tab. 4). At this point it must be stated clearly that by observing frequencies of *differences* only, the evaluated instruments are treated symmetrically, i. e. we do not blame one method for producing false results. Depending on the respective analyte, one or the other

methodology might be advantageous. Judgement of true results is only possible with respect to definitive methods. Unfortunately, such methods are not available for all analytes and method comparisons are performed with respect to nationally recommended procedures. Therefore assessment of accuracy is only possible with respect to these procedures, and implementation of new technologies can only occur after modification of the respective recommendations.

In conclusion we have shown that several problems arise if different methodologies are used simultaneously in the same laboratory, even if acceptable correlation data exist. Since method-dependent bias causes considerably divergent results, a simultaneous use of both instruments should be carefully organized. For example, patient-dependent sample distribution (same instrument for all specimens of one patient) might be appropriate. Another possibility would be the distribution of the applied methods between instruments, e. g. sole use of the most convenient method. Finally, an unselected use of both instruments combined with careful control and perhaps re-measurement with the alternative instrument seems possible. But all these measures implicate an increase in workload, costs, turnaround time, and chance of mistakes.

References

- Curme, H. G., Columbus, R. L., Dappen, G. M., Eder, T. W., Fellows, W. D., Figueras, J., Glover, C. P., Goffe, C. A., Hill, D. E., Lawton, W. H., Muka, E. J., Pinney, J. E., Rand, R. N., Sanford, K. J. & Wu, T. W. (1978) Multilayer film elements for clinical analysis: General concepts. *Clin. Chem.* 24, 1335–1342.
- Libeer, J. C. (1985) Solid phase chemistry in clinical laboratory tests: a literature review. *J. Clin. Chem. Clin. Biochem.* 23, 645–655.
- Sonntag, O. (1988) *Trockenchemie. Analytik mit trägergebundenen Reagenzien*. Georg Thieme, Stuttgart–New York.
- Sandberg, S., Christensen, N. G., Thue, G. & Lund, P. K. (1989) Performance of dry-chemistry instruments in primary health care. *Scand. J. Clin. Lab. Invest.* 49, 483–488.
- Easte, S. J., Benson, P. J., Preese, L. M. & Apple, F. S. (1989) Factitious high sodium activities on the Ektachem 400 owing to interferences by high gamma-globulin concentrations. *Clin. Chem.* 35, 333–334.
- Hafkenscheid, J. C. M. & van der Ven-Jongerkrijg, J. (1988) Two dry-reagent systems evaluated for determinations of enzyme activities. *Clin. Chem.* 34, 155–157.
- Wallinder, H., Angbäck, K., Kahan, J., Onica, D. & Waldenlind, L. (1989) Measurement of alkaline phosphatase activity by Kodak Ektachem 700 XR analyzer compared with other routine analytical methods. *Clin. Chem.* 35, 328–329.
- Weiß, T. & Lorentz, K. (1984) Vergleichende Lipasebestimmungen mit turbidimetrischer Technik und Enzymimmunoassay. *Lab. Med.* 8, 63–67.
- Junge, W., Leybold, K. & Kraack, B. (1983) Influence of colipase on the turbidimetric determination of pancreatic lipase catalytic activity. *J. Clin. Chem. Clin. Biochem.* 21, 445–451.
- Bablok, W. & Passing, H. (1983) A new biometrical procedure for testing the equality of measurements from two different analytical methods. *J. Clin. Chem. Clin. Biochem.* 21, 709–720.
- Haeckel, R. & Passing, H. (1985) Statistische Probleme beim Vergleich klinisch-chemischer Analysenverfahren. *J. Clin. Chem. Clin. Biochem.* 23, 307–318.
- Sachs, L. (1984) *Angewandte Statistik: Anwendung statistischer Methoden*. 17th Ed., Springer, Berlin.
- Feldmann, U., Schneider, B. & Klinkers, H. (1981) A multivariate approach for the biometric comparison of analytical methods in clinical chemistry. *J. Clin. Chem. Clin. Biochem.* 19, 121–137.
- Kadlinger, C. L. & O'Kell, R. T. (1983) Clinical evaluation of Eastman Kodak's Ektachem 400 analyzer. *Clin. Chem.* 29, 498–501.
- Fraser, C. G. (1983) Desirable performance standards for clinical chemistry tests. *Adv. Clin. Chem.* 23, 299–339.
- Tonks, D. B. (1963) A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. *Clin. Chem.* 9, 217–223.
- Stamm, D. (1982) A new concept for quality control of clinical laboratory investigation in the light of clinical requirements and based upon reference method values. *J. Clin. Chem. Clin. Biochem.* 20, 817–824.
- Stamm, D. (1988) Qualitätssicherung der quantitativen Bestimmungen im Laboratorium. *Dt. Ärztebl.* 85, 697–712.

19. Costongs, G. M. P. J., Janson, P. C. W., Bas, B. M., Hermans, J., van Wersch, J. W. J. & Brombacher, P. J. (1985) Short-term and long-term intra-individual variations and critical differences of clinical chemical laboratory parameters. *J. Clin. Chem. Clin. Biochem.* 23, 7–16.
20. Westgard, J. O. & Hunt, M. R. (1973) Use and interpretation of common statistical tests in method-comparison studies. *Clin. Chem.* 19, 49–57.
21. Koch, D. D., Hassemer, D. J., Wiebe, D. A. & Laessig, R. H. (1988) Testing cholesterol accuracy. *J. Am. Med. Ass.* 260, 2552–2557.
22. Harris, E. K. (1975) Some theory of reference values. I. Stratified (categorized) normal ranges and a method for following an individual's clinical laboratory values. *Clin. Chem.* 21, 1457–1464.
23. Harris, E. K. (1976) Some theory of reference values. II. Comparison of some statistical models of intraindividual variation in blood constituents. *Clin. Chem.* 22, 1343–1350.
24. Harris, E. K., Cooil, B. K., Shakarji, G. & Williams, G. Z. (1980) On the use of statistical models of within-person variation in long-term studies of healthy individuals. *Clin. Chem.* 26, 383–391.
25. Harris, E. K. & Yasaka, T. (1983) On the calculation of a "reference change" for comparing two consecutive measurements. *Clin. Chem.* 29, 25–30.
26. Gressner, A. M. & Sittel, D. (1985) Plasma pyridoxal 5'-phosphate concentrations in relation to apo-aminotransferase levels in normal, uremic, and post-myocardial infarct sera. *J. Clin. Chem. Clin. Biochem.* 23, 631–636.
27. Bissell, M. G., Ward, E., Sanghavi, P., Toffaletti, J., Bird, C., Berg, C., Jerkunica, I. & Hutton, P. F. (1987) Multilaboratory evaluation of the new, single-slide enzymatic creatinine method on the Kodak Ektachem analyzer. *Clin. Chem.* 33, 951.
28. Kroll, M. H., Roach, N. A., Poe, B. & Elin, R. J. (1987) Mechanism of interference with the Jaffé reaction for creatinine. *Clin. Chem.* 33, 1129–1132.
29. Roberts, R. T., Alexander, N. M. & Keiner, M. J. (1988) Definitive liquid-chromatographic demonstration that n-ethylglycine is the metabolite of lidocaine that interferes in the Kodak sarcosine oxidase-coupled method for creatinine. *Clin. Chem.* 34, 2569–2572.
30. Apple, F. S., Benson, P., Abraham, P. A., Rosano, T. G. & Halstenson, C. E. (1989) Assessment of renal function by inulin clearance: comparison with creatinine clearance as determined by enzymatic methods. *Clin. Chem.* 35, 312–314.
31. Doumas, B. T., Hause, L. L., Simuncak, D. M. & Breitenfeld, D. (1989) Differences between values for plasma and serum in tests performed in the Ektachem 700 XR analyzer, and evaluation of "Plasma Separator Tubes (PST)". *Clin. Chem.* 35, 151–153.
32. Berg, J. D., Romano, G., Bayley, N. F. & Buckley, B. M. (1988) Heparin interferes with aspartate aminotransferase activity determination in the Ektachem 700. *Clin. Chem.* 34, 174.

Dr. J. Kropf
Department of Clinical Chemistry
Philipps University
Baldingerstraße
W-3550 Marburg
Bundesrepublik Deutschland

