phoethanolamine) should be considered for the assessment of chronic alcoholism.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: No authors declared any potential conflicts of interest.

References


Frequency of Extreme Differences and Clinical Performance of Glucose Concentration Measurements Judged from 21 000 Duplicate Measurements

To the Editor:

Duplicate measurements of patient samples are not common in routine clinical chemistry. The performance of measurements is continuously monitored by the internal quality control (IQC), which is not designed to find occasional dropouts. Therefore, the frequency of such errors is not usually known. In the present study, we attempted to estimate the error frequency at commonly suggested imprecision targets by evaluating the difference between 2 adjacent repeated measurements. The frequency is expected to be very low, and therefore the study of this problem required a large number of samples.

We measured glucose concentration over a 6-month period in samples obtained from an unselected population of hospitalized patients and outpatients. We used 3 Dimension Vista 1500 analyzers connected to StreamLAB (Siemens Healthcare Diagnostics), in accordance with the manufacturer’s instructions. To include measured values in our study, we required that

Fig. 1. Box-and-whisker plots for different subgroups of transferrin phenotypes for the ratio of the %CDT measured by the N Latex CDT assay to the %CDT measured by CZE.

* P < 0.001 and ** P < 0.0001, Mann–Whitney U-test for independent samples compared with the wild-type subgroup. Data are presented as the median, interquartile range, and range.
glucose assay performance comply with the RiLiBAEK (Richtlinien der Bundesärztekammer) IQC limits (1) over the measuring interval of 1 mg/dL (0.06 mmol/L) to 500 mg/dL (27.75 mmol/L).

We measured 21 653 samples twice, with each sample randomly allocated to one of the instruments by the laboratory automation system. The repeat measurement, which was performed on the same instrument as the first measurement, was not reported for clinical purposes. Data analyses performed with Microsoft® Excel® (2007) included standard descriptive statistics, linear regression analysis, difference plots, and the SD calculated from the duplicate values. Minor differences in median concentrations were observed between instruments (sample medians of 101, 101, and 103 mg/dL for the 3 instruments).

The regression analysis showed an apparent linear relationship between the first and repeat results, with a correlation coefficient of 0.996–0.997, a slope of 0.996–0.998, and a y intercept 0.000 mg/dL. Assay performance was expressed as the number of differences within a predetermined zone around the line of equivalence, which we termed the “zone of acceptance” (A zone) (2). Observations that fell outside the A zone were regarded as extreme differences.

We counted the number of samples with differences that would exceed the recommended limits derived from several national and international organizations. We evaluated the results obtained from all 3 instruments as a single data set, because we observed no important differences between instruments. Of the observed differences, 395 fell below the first percentile or above the 99th percentile, and 44 were outside the 0.1st and 99.9th percentiles. These results yielded rates of 18 and 2 per 1000 measurements, respectively, as expected.

With the RiLiBAEK IQC target imprecision for glucose (CV, ±11%), 1 in 1000 of our differences would have been excluded as outliers, and 0.3 in 1000 results would have been excluded by the CLIA 88 recommended imprecision (3) [±6 mg/dL (±0.83 mmol/L) or ±10% <76 mg/dL (<4.2 mmol/L), whichever is reached first; and ±20% >76 mg/dL (>4.2 mmol/L)]. According to the “desirable imprecision” criterion recommended by Ricos et al. (4, 5), 25 in 1000 observations would have been excluded: ±4.6% (coverage factor, k = 2, corresponding to a 95% confidence interval) (Fig. 1). The imprecision required to include 95% of the observations in the A zone (i.e., 50 in 1000 measurements outside of the A zone) was ±3.8% (Fig. 1).

Another approach to define targets of allowable extreme differences would be to consider the minimal difference (MD): MD = k × \sqrt{2} × SD^2 = k × SD × \sqrt{2}. This is the smallest significant difference between 2 arbitrary measurements at a level of confidence defined by the coverage factor k, assuming the same measurement uncertainty (i.e., the imprecision of the measurement procedure). Yet another target would be the reference change value (RCV) (6), which adds the biological variation (BV) to the MD: RCV = k × \sqrt{MD^2 + 2 × BV^2}, thereby quantifying the statistically significant difference between 2 consecutive results for the same individual. The estimated MD (k = 2) was 6.8 mg/dL (0.38 mmol/L) or 7.5% at 90 mg/dL (5.0 mmol/L). Use of this MD to delimit an A zone would translate into a frequency of extreme differences in our experiment of approximately 2 per 1000. Assuming a biological variation of 4.9% (4, 5) (i.e., the same magnitude as the analytical variation), we estimated the RCV at 90 mg/dL (5.0 mmol/L) as 14.4% or 13 mg/dL (0.76 mmol/L) (k = 2). Use of this RCV to delimit the A zone would translate into an extreme difference frequency of about 0.4 in 1000 measurements, approximately the same magnitude as applying the RiLiBAEK or CLIA requirements to the duplicates.
Previously, duplicates were regularly measured to improve the imprecision and to spot dropouts. Our study shows a measurable frequency of dropouts when glucose, as an example analyte, is measured with the Siemens Dimension Vista 1500, which performs well within RiLiBAEK limits. Even when measurements comply with IQC rules, the experience of this study indicates about 1 extreme difference in 1000 measurements would be expected. Half of these extreme differences exceed the RCV and might have an impact on patient care.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. 

Disclosures and/or potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: Reagents were supplied by Siemens Healthcare Diagnostics.
Expert Testimony: None declared.
Patents: None declared.

References


Astrid Petersmann2*
Christina Wasner2
Matthias Nauck2
Anders Kallner3

2 Institute of Clinical Chemistry and Laboratory Medicine
University Medicine Greifswald
Greifswald, Germany

3 Department of Clinical Chemistry
Karolinska University Hospital
Stockholm, Sweden

* Address correspondence to this author at:
Institute of Clinical Chemistry and Laboratory Medicine
University Medicine of Greifswald
Ferdinand-Sauerbruch-Straße
17475 Greifswald, Germany
Fax +49-3834-865302
E-mail astrid.petersmann@uni-greifswald.de

Previously published online at
DOI: 10.1373/clinchem.2012.197996