Objective To assess clinical implications of bias and variance of point-of-care glucometric measurements in cardiac surgery patients with wide variations in postoperative hematocrit.

Methods Point-of-care glucose measurements were compared with values from laboratory analysis of the same sample of whole blood obtained from cardiac patients early on postoperative days 1 and 2. Twenty nurses collected 89 arterial blood samples from 58 patients during a 4-month period. Bias was measured by using difference scores between paired measurements. Patients were grouped within 5% increments according to hematocrit, and analysis of variance was used to test for differences. Variation was analyzed by precision-to-tolerance analysis within 3 euglycemic tolerance ranges.

Results Laboratory glucose values were 62 to 224 mg/dL; point-of-care measures were 83 to 253 mg/dL. Bias was 10.85 mg/dL across all hematocrit groups. Pairs of laboratory and point-of-care glucose values differed significantly ($t_{174} = 10.03$; $P < .001$). Bias increased from -2.83 mg/dL for patients with hematocrits exceeding 39% to +16.71 mg/dL for patients with hematocrits between 20% and 24%. The standard deviation of difference scores was 11.59 mg/dL overall. The difference between 5% hematocrit groups was significant ($F_4 = 4.11$; $P = .004$). Precision-to-tolerance capability ratios for specification limits of 70 to 300, 90 to 140, and 80 to 110 mg/dL were 0.30, 1.39, and 2.32, respectively.

Conclusions The direction of bias change between hematocrit groupings was the direction predicted in the manufacturer’s information. Precision-to-tolerance measures indicated that the point-of-care equipment was not suitable for testing glucose within the planned “tighter” glycemic standards. (American Journal of Critical Care. 2009;18:232-239)
Since 1990, many reports and reviews\textsuperscript{1-24} have been published on studies to evaluate the relative equivalence of serum glucose values obtained via point-of-care (POC) glucometers and via standard laboratory tests. Most investigators have reported acceptable differences between results obtained with these glucometers and the laboratory standards. Study teams consistently report the mean difference between values obtained with different measurement devices (bias) and the percentages of patients whose variant POC statistics would cause inappropriate corrections for insulin administration. Correlations are nearly always reported between laboratory and POC values and nearly always exceed 0.90. Although the variance or random error is mentioned in these reports, rarely has it received consideration in relation to a glucometer’s discriminant capability. With precision-to-tolerance, discrimination capability can be tested with variance data determined through test-retest studies. Such analysis defines the limitations of the glucometer in detecting important differences from specification limits surrounding the range of “normal.” After precision-to-tolerance analysis, clinicians can better manage the balance between glucometer imprecision and aggressiveness of an insulin adjustment protocol.

The issue regarding overall precision of our glucometers came to our attention as the result of an investigation that was initially designed to answer a more basic question. Nurses in our cardiovascular intensive care unit knew that blood hematocrit is widely variable in patients immediately after cardiac surgery. In addition, our POC coordinator knew of the warning of the manufacturer of our glucometers that low and high hematocrit values can introduce bias. For 16 years, we have used the same glucometers, which rely on glucose oxidase test strips for the analysis. The nurses questioned whether the amount of bias due to variant hematocrit was clinically important to the welfare of their patients. Did we need to alter our measurement and insulin adjustment processes to compensate for the variability?

Although we were easily able to measure the effects of low hematocrit on glucometer bias, we were surprised to learn simultaneously that the precision of our meters was insufficient to support our plan to “tighten” glucose control both in the critical care and subacute nursing units. In this article, we describe a study that revealed previously unappreciated limitations of our POC devices. We also provide a description of the precision-to-tolerance analysis that caused us to redesign our insulin management processes.

Materials and Methods

The study was conducted at a 384-bed community hospital in Boise, Idaho, with a 10-bed cardiovascular intensive care unit that treats approximately 350 surgical patients per year. Most of these patients experience an acute decrease in hematocrit within the first 2 days after cardiac surgery, and all require daily morning laboratory tests. These laboratory tests include either a measurement of serum level of glucose within a metabolic chemistry panel or a blood cell count alone. If blood chemistry tests including determination of serum glucose level was not ordered on a particular day for our study patients, a “no charge” measurement of serum level of glucose was completed with the unused part of the blood sample. Insulin administration continued to be based solely on POC glucose measures even though laboratory values were available later. The POC glucometer used was a brand that has been used widely for more than a decade throughout the health care system of which our hospital is a member.

A convenience sample of patients experiencing cardiac surgery was used. Blood was sampled by

About the Authors

Catherine Prinzing is a clinical nurse specialist for diabetes, Sarah Rosenlund is manager of the cardiovascular intensive care unit, Vicki Sukeena is the point-of-care testing coordinator, Cynthia Malinowski is a nurse educator in the cardiovascular intensive care unit, and Lowell C. Wise is a Six Sigma black belt at Trinity Health-Saint Alphonsus Regional Medical Center in Boise, Idaho.

Corresponding author: Lowell C. Wise, RN, DNSc, Six Sigma Department, Trinity Health-Saint Alphonsus Regional Medical Center, 1055 N Curtis Rd, Boise, Idaho 83706 (e-mail: lowewise@sarmc.org).
using customary syringe withdrawals from arterial catheters. Following standard procedure, nurses activated a mechanism within our in-line venous arterial blood management protection system that removes 10 mL from the arterial catheter before a sample is obtained for testing. Because the original intent was to answer a simple question about accuracy in patients with a low hematocrit, patients were selected to ensure that sufficient numbers of samples were collected from patients with hematocrits between 20% and 40%. Sample size could not be estimated before the study started because the effect size and variation were unknown. Comparison of sample sizes from similar published studies provided much of our guidance.

On postoperative days 1 and 2, two blood samples taken from the same syringe were used for testing. One was sent to the central laboratory for a serum glucose test and the other whole blood sample was applied to the POC reagent strip. To replicate normal operations, each primary care nurse collected the blood sample and did the POC testing. A study coordinator recorded the results for each patient. Nurses on the night shift generally retrieved the sample and did the POC testing between 4 AM and 6 AM. Data consisted of paired laboratory and POC glucose results, patients’ hematocrits, time of sample collection, and identity of the nurse who obtained the sample. The study was reviewed by the quality improvement investigational review board at Saint Alphonsus Regional Medical Center.

The average size of difference scores (bias) was greatest at the lowest hematocrit levels and closest to matching lab values above 35%.

### Data Analysis

Difference scores were calculated for each pair of POC and laboratory values. Laboratory results and POC scores were compared by using bivariate correlation. Patients’ hematocrits were used to assign comparative POC and laboratory glucose scores to 5 hematocrit-based categories: 20% to 24%, 25% to 29%, 30% to 34%, 35% to 39%, and greater than 39%. Difference scores were compared between hematocrit groups by using analysis of variance with Minitab statistical software. In each analysis, α was set at .05.

Bland and Altman scatter diagrams were created to provide a visual representation of the relationship between difference scores and categories based on hematocrit groupings and at different glucose levels. Scatter diagrams also were created to illustrate the correlation between glucose values obtained by laboratory and POC methods.

Precision-to-tolerance calculations were added to the analysis to judge the capability (gauge capability) of our POC devices in discriminating normal and abnormal serum or blood glucose values under narrower control limits. Although this consideration has not received much attention in earlier POC glucometric studies, we did this analysis because of Bland and Altman’s admonition that random error be considered in evaluations of measurement devices. Two competing models are used to evaluate precision-to-tolerance of measurements. Many authors endorse the stringent standard of gauge capability whereby a measurement device is “gauge capable” only if random error remains less than 30% of total measurement variation at the 6-sigma (standard deviation [σ] x 6) level. This standard is customarily used in production industries to evaluate process control measures and is used only rarely in health care operations. The purpose of this analysis is to decide whether to use or replace a measurement device.

The other model of measurement analysis uses a more comprehensive statistical method that establishes the amount of error uncertainty in a measurement device without a predetermined score to decide gauge capability. Instead, the resulting analysis provides boundaries of statistical certainty that help decision makers select a particular device in specific operational circumstances. The questions to be answered were (1) whether degree of uncertainty due to random error subverts our ability to distinguish normal from abnormal glucose values, and (2) whether we can adjust clinical protocols sufficiently to accommodate these limitations of precision. We used both models to broaden our understanding of what gauge capable meant in our particular context.

With the Six Sigma standard method, precision-to-tolerance was calculated from a pooled standard deviation to produce a ratio as follows:

\[
\frac{6 \times \sigma}{USL-LSL}
\]

In this ratio, σ is the standard deviation of the difference scores, USL is the upper specification limit, and LSL is the lower specification limit. The formula for precision-to-tolerance ratio was applied to 3 separate sets of USLs and LSLs that represented 3 different ranges of interest. One reflected our former euglycemic range (70-300 mg/dL), and 2 other tighter ranges represented our goals for adoption in critical care and medical units, namely, 80 to 110 mg/dL and 90 to 140 mg/dL.
Using the alternative statistical error analysis applied to the 3 glycemic ranges of interest required only that we calculate the median or probable error from the standard deviation \( (= 0.67 \times \sigma) \), select a level of confidence (95%) that a value is above or below a specification boundary, then calculate a 95% range of uncertainty on either side of each boundary limit. This process allowed us to visually depict the range of measured values for which the meaning is clear and to separate ranges of uncertain (error) values, the implications of which require expert clinical judgment.

**Results**

**Bias and Hematocrit**

A total of 89 pairs of blood samples from 58 different patients were evaluated during the 4-month study period. Laboratory serum glucose measurements varied between 62 and 224 mg/dL, and POC measurements varied between 83 and 253 mg/dL. The mean difference between laboratory and POC measurements was 12.32 (SD, 11.586) mg/dL across all groups of hematocrit. The difference exceeded 20% in 10 (11% of cases). During this study, 3 episodes of hypoglycemia (<70 mg/dL) were detected via laboratory samples; none of the 3 were discovered by POC testing. The overall correlation between pooled laboratory and POC results was high \( (r = 0.93; P < .001; \text{Figure 1}) \).

A Bland and Altman scatterplot of difference scores was created to show whether the hematocrit had any relationship to the amount of difference between bedside POC glucose values and clinical laboratory values (Figure 2). The scatterplot shows the anticipated negative relationship between hematocrit and difference scores.

A similar scatterplot was created to illustrate the relationship between the amount of difference in comparative scores over the entire range of glucose values. The scatterplot showed no pattern indicating a relationship (Figure 3).

Analysis of variance yielded significant differences between patients’ glucose measurements and paired laboratory values when results were grouped into 5% hematocrit categories \( (F_A = 4.11; P = .004; \text{Figure 4}) \). The mean size of difference scores (bias) was greatest at the lowest hematocrit levels (20%-24%) and closest to matching laboratory values at hematocrits exceeding 35%.

**Precision-to-Tolerance Capability**

The standard deviation of difference scores was 11.59 mg/dL overall and did not vary significantly between hematocrit groups. From this statistic, precision-to-tolerance ratios were calculated for...
specification limits of 70 to 300, 90 to 140, and 80 to 110 mg/dL. The ratios were 0.30, 1.39, and 2.32, respectively. Only at the 70 to 300 range was the ratio approximately 30%, the minimum necessary to establish gauge capability by using the Six Sigma evaluation standard.

After these results were presented to management and clinicians at affiliated hospitals, we searched for meters that would offer better control over bias and variability. We used this opportunity to validate our findings while we tested competing glucose analyzers for use in critical care and general nursing departments. Two laboratory grade analyzers priced at $6000 for use in critical care and general nursing departments. Two laboratory grade analyzers priced at $6000 and $13000 were compared alongside 2 new glucometers from vendors and our own bedside meter, all of which cost less than $1000. Validation tests were conducted by using same-sample test-retest methods. The 2 laboratory-grade analyzers and the 2 newer glucometer products showed negligible bias (all <3 mg/dL). A total of 223 comparisons with our current meter produced a consistent bias of nearly 14 mg/dL. This bias was further validated by comparing all laboratory glucose measurements with POC blood glucose results occurring within 1 hour of each other (n = 1786). Once again, we found a significant positive bias of 16 mg/mL.

### Discussion

The variation in glucose values obtained with our standard glucometer was similar to that found during the initial hematocrit study. Variation among the 2 competing, less-expensive glucometers was essentially the same as that for our existing meter. Variation was between 5 and 6 mg/mL in both the laboratory grade analyzers, which prompted one of our critical care units to limit the use of traditional POC glucometers to patients receiving subcutaneous insulin. In this unit, laboratory-grade analyzers are now used for all patients treated with insulin infusions.

In one unit, laboratory-grade analyzers are now used for all patients on insulin infusion therapy.
not provide the statistical tools that would help us manage error in quantitative terms.

Wheeler and Lyday provide a compelling model that is both statistically sound and instructive. With their model, it is easy to produce illustrations of dynamic relationships between specification limits, levels of measurement error, and the degree of certainty that is needed to interpret measured values safely.

When error represents an approximately normal distribution, the median error can be expressed as the point at which 50% of error occurs between -0.67 \( \sigma \) and +0.67 \( \sigma \). The median error is also referred to as the probable error, and its value varies proportionately with the standard deviation of the error. When the probable error is large in proportion to the unit of measurement, the resolution of the measurement process becomes increasingly imprecise. Values for probable error that encroach on the upper and lower specification limits of "normal" increase the uncertainty of determining whether a specific measured value is within or outside the defined limits of euglycemia.

Wheeler and Lyday are critical of formulas such as the one for precision-to-tolerance ratio described earlier. If such formulas are consistently applied without consideration for their practicality, Wheeler and Lyday imply that these rules limit our ability to accommodate imperfections in a measurement process. They argue that under certain circumstances, imperfect measurement systems can be made to accommodate probable error levels. After determining the level of uncertainty of correct valuation that a process can tolerate, one can adjust the specification limits by tightening the "within tolerance” limits and widening the limits for discriminating values that are outside the range of tolerance.

Figure 5 illustrates this concept with our new target range for euglycemia, which was increased as a result of our analysis from 80 to 130 mg/dL to 90 to 140 mg/dL. Additional repeated measures were added to our original data set as we compared our current meter with devices provided by a competing manufacturer of glucose analyzers. The resulting 167 comparisons produced a bias of 13.6 mg/dL and a standard deviation (error) of difference scores of 13.3 mg/dL.

For a 95% probability or certainty that a value is “within limits,” the upper and lower specification limits must be adjusted downward and upward, respectively, by twice the probable error (1.34 \( \sigma \)). This adjustment would mean that around the lower and upper limits of normal (90 and 140 mg, respectively), a 35.6-mg range of uncertainty (95%) must be taken into account. Any POC glucose value falling within this range of uncertainty must be regarded as being neither within or outside of the normal range. Thus, if insulin adjustment were contemplated for hypoglycemia defined by 90 mg/dL, one would need to decide whether the actual measurement must be less than 72.2 mg/dL before executing an insulin correction. The new upper threshold would likewise be 140 + 17.8 mg/dL. Treatable hyperglycemia based on an upper limit of 140 mg/dL might involve delaying adjustment until a value of 157.8 mg/dL has been exceeded. One can be 95% certain that any blood glucose value between 108.6 and 122.2 mg/dL is truly within the limits of normal. The practical and clinical issue is how to regard blood glucose values falling within the 35.6-mg “uncertainty range” surrounding the 2 specification limits.

When we applied this degree of random error and the error adjustment to “tight glucose control” (80-110 mg/dL), the error we found for our bedside...
may have affected the comparison score results. Although the POC tests were conducted solely by registered nurses from a single nursing unit, the number of operators (n = 23) was too large for the small sample size to conduct a credible analysis of interoperator differences. Finally, these results were limited to whole blood compared with serum determinations. The findings would probably have been different had fingerstick samples been used in the glucometer method. Because fingerstick POC glucose testing is much more common in hospitals, generalization from these findings to the more common use of POC testing is certainly limited. The housewide validation analysis showed larger differences in bias and difference score variation when fingerstick and catheter samples were combined with hundreds of POC test operators.

**Conclusion**

Our medical and nursing staff have used the same POC glucometers in the acute and critical care units for years, and all assumed that the measurement devices were accurate and precise. Nonetheless, anecdotal accounts have occurred of clinically relevant discrepancies between laboratory results that became evident only hours after POC testing was completed and laboratory results were posted. In some cases, patients’ insulin infusions or subcutaneous doses were changed in response to POC values, and those adjustments would not have been indicated if laboratory results had been available.

Whether one adopts the Six Sigma analysis or Wheeler and Lyday’s limit adjustment model, it is useful to apply measurement variance analysis to all testing devices considered for application to a specific test and a specified range for discrimination. When we want to change the range or increment of measure within which we want to move clinical care, the same measurement tools may suddenly fail us. They may become incapable of delivering reliable information for managing the quality of care we provide.

**FINANCIAL DISCLOSURES**

None reported.

**REFERENCES**

2. Chrousos G, Kaltas G. How accurate are currently used methods of determining glycemia in critically ill patients.

---

**Figure 6** Illustration of applying to a “tight glucose control” process Wheeler’s instrument adjustment model for our current glucose oxidase point-of-care device and its measured variance. No point-of-care glucose values between 80 and 110 mg/dL can be regarded as within range with 95% certainty.
Precision-to-Tolerance Capability: An Important Consideration in Tight Glucose Control
Catherine Prinzing, Sarah Rosenlund, Vicki Sukeena, Cynthia Malinowski and Lowell C. Wise

Am J Crit Care 2009;18 232-239 10.4037/ajcc2009262
©2009 American Association of Critical-Care Nurses
Published online http://ajcc.aacnjournals.org/

Personal use only. For copyright permission information:
http://ajcc.aacnjournals.org/cgi/external_ref?link_type=PERMISSIONDIRECT

Subscription Information
http://ajcc.aacnjournals.org/subscriptions/

Information for authors
http://ajcc.aacnjournals.org/misc/ifora.xhtml

Submit a manuscript
http://www.editorialmanager.com/ajcc

Email alerts
http://ajcc.aacnjournals.org/subscriptions/etoc.xhtml