Method Precision and Frequent Causes of Errors
Observed in Point-of-Care Glucose Testing

A Proficiency Testing Program Perspective

Berna Aslan, MD, Julia Stemp, ART, Paul Yip, PhD, and Jane Gun-Munro, MLT, MBA

From the Ontario Medical Association, Institute for Quality Management in Healthcare, Toronto, Canada; and Department of Laboratory Medicine and Pathobiology, University of Toronto and University Health Network, Toronto, Canada.

Key Words: POC glucose; External quality assessment; Proficiency testing; Analytic quality goals; Performance standards

ABSTRACT

Objectives: Method imprecision, error rates, and explanatory causes that were identified in the Institute for Quality Management in Healthcare point-of-care (POC) glucose proficiency testing (PT) program were assessed in comparison with results obtained from laboratory glucose PT surveys.

Methods: POC and laboratory glucose PT data were assessed from September 2009 to June 2011. Peer group means and coefficients of variation (CVs) were estimated using the robust algorithm recommended in the International Organization for Standardization/International Electrotechnical Commission 13528(E). Discordant finding investigations were also reviewed to determine the causes of significant and recurring errors.

Results: POC glucose CVs were higher than laboratory method CVs (median CV, 4.5% and 1.6%, respectively). While all laboratory glucose results were within the performance limits, 305 (0.59%) of 51,379 POC glucose results exceeded limits. Investigations were required for 277 (0.53%) POC results. Pre- and postanalytical errors accounted for 76% of the discordant findings. Using wrong PT items, sample mix-up on the bench, and reporting results for the wrong sample were the most frequent reasons, while 21% of discordant findings identified manufacturer issues, and 3% were of unknown origin.

Conclusions: Both method CVs and error rates were higher in POC than in laboratory glucose methods, even though larger performance limits were used for the assessment of POC glucose.

Point-of-care (POC) testing is a rapidly developing division of the diagnostic laboratory sector. Major advantages brought by this technology are an increase in accessibility to laboratory tests, decreased test turnaround times (TATs), and use of smaller sample volumes. These advantages make POC testing a method of choice in primary care; homecare; locations where ready access to laboratory services is limited; emergency departments (EDs) and intensive care units (ICUs), where TAT has vital importance; and neonatal ICU and pediatric wards, where low sample volumes and timely results are crucial.

POC glucose testing, one of the earliest examples of this technology, was originally developed for self-monitoring of blood glucose for diabetic patients who were taking subcutaneous insulin. After studies indicated that tight glycemic control decreased the mortality rates in critically ill patients in the ICU, intravenous insulin treatment for tight glucose control was accepted as a standard of care. Tighter targets for blood glucose concentrations in ICU patients have caused an increased risk of hypoglycemia and require frequent glucose monitoring in these patients. This has increased the use of POC glucose testing.

In addition to the ICU setting, POC glucose testing is also used in other sections of tertiary health care institutions, such as wards, EDs, obstetrics and gynecology services, diabetic clinics, and general inpatients. POC glucose is used for monitoring patients treated in tertiary care centers, often interchangeably with laboratory glucose values. Accurate measurement of blood glucose is essential for correct treatment decisions for glycemic control. Despite the frequent use of POC glucose testing in hospital settings, its quality assurance is still challenging for many institutions, with
very limited data available on error rates. In this study, we investigated POC glucose testing method precision, as well as causes of recurrent and gross errors that have occurred in the Institute for Quality Management in Healthcare (IQMH) POC glucose Proficiency Testing (PT) program.

Materials and Methods

IQMH POC glucose and laboratory glucose PT data obtained between September 2009 and June 2011 were used for the assessments. IQMH is a not-for-profit corporation wholly owned by Ontario Medical Association. IQMH proficiency testing programs are accredited against International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 17043:2010: Conformity Assessment—General Requirements for Proficiency Testing.10 Clients in Canada and other countries may voluntarily apply for accreditation and/or participate in PT surveys through the IQMH.

IQMH accreditation requirements are based on ISO 1518911 and ISO 22870:2006,12 which contain general medical laboratory and POC testing requirements, respectively. All participants in the POC and the laboratory glucose surveys included in this study were accredited Ontario laboratories, and as a regulatory requirement, all POC testing devices in the facilities were managed by the institutions’ respective laboratory regardless of the location of the devices within each institution.

The IQMH provides POC glucose and laboratory glucose surveys (ie, glucose tests performed on an automated analyzer in a hospital laboratory). The POC glucose survey consists of two shipments per year with three PT items in each shipment. The laboratory glucose test PT program consists of three shipments with three levels of PT material each time.

Commercially prepared bovine plasma products (Eurotrol WB Glucose, Ede, the Netherlands) and pooled human serum samples were used as PT items in the POC and laboratory glucose surveys, respectively. It is a challenge for PT providers to provide capillary whole blood or a commutable PT item that is compatible with all different brands of POC glucose devices. PT items for the POC glucose survey contain a physiologic buffered salt solution with a small amount of purified plasma of bovine origin. Because of the high probability of a matrix effect in this artificial PT item, only imprecision within instrument peer groups is evaluated, not method bias.

Peer group means and coefficients of variation (CVs) were estimated and compared. These statistical parameters were calculated using the robust algorithm recommended in ISO/IEC 13528(E).13 This nonparametric statistical algorithm consists of an iterative statistical function that minimizes the effects of the outliers on peer group mean, SD, and CV. Calculations start with a median (x*) as a statistical parameter for central tendency and s* = 1.483 × median |x*–x| as a parameter for the dispersion around the median. The next step is that all values in the original data set that are lower than x* – 1.5 s* and higher than x* + 1.5 s* are trimmed to x* – 1.5 s* and x* + 1.5 s*, respectively. In the following step, new x* and s* are estimated as the mean and SD of a new data set using conventional parametric statistics. The trimming step is repeated several times until there is no change from one iteration to the next in the third significant figure of mean and SD.

Laboratories’ performances were assessed using predefined allowable performance limits (APLs) (also known as performance specifications, analytic quality specifications, or analytical goals) and reflect total allowable error. APLs that are used in the survey assessment are determined by the IQMH scientific committees. Committees use one or more of the following methods in the determination of an APL: clinical outcomes in specific clinical settings, biological variation, recommendations from professional expert groups and the regulatory bodies, and current state of the art as demonstrated by PT program providers. Table 1 shows the allowable performance limits used at the time of the PT challenges included in this study.

The standard uncertainty of assigned value (UAV) was also estimated using the method recommended in the ISO 13528, and when peer groups’ UAV were higher than 10% of the allowable performance range, peer groups were excluded from the study. The number of participants in each peer group was changing from 5 to 1,500 in the POC glucose and 5 to 47 in the laboratory glucose PT surveys. The Mann-Whitney U test was used for statistical comparison of the laboratory and POC glucose method CVs.

Participants are required to submit formal discordant investigation forms when their results exceed allowable limits at a given level, indicating large variations from the assigned mean. This investigation must include the cause of the discordant result and its corresponding corrective action. The program’s advisory or expert committee reviews the

<table>
<thead>
<tr>
<th>Table 1</th>
<th>IQMH Allowable Performance Limits for Laboratory and POC Glucose Testing at the Time of the Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method and Glucose Concentration</td>
<td>Allowable Performance Limits</td>
</tr>
<tr>
<td>Laboratory glucose</td>
<td></td>
</tr>
<tr>
<td>≤4 mmol/L</td>
<td>±9%</td>
</tr>
<tr>
<td>&gt;4 mmol/L</td>
<td>±7.5%</td>
</tr>
<tr>
<td>POC glucose</td>
<td></td>
</tr>
<tr>
<td>≤5 mmol/L</td>
<td>±1 mmol/L</td>
</tr>
<tr>
<td>&gt;5 mmol/L</td>
<td>±20%</td>
</tr>
</tbody>
</table>

IQMH, Institute for Quality Management in Healthcare; POC, point of care.
The difference was statistically not significant (concentration group and 4.3% in the high-concentration group; *P* > .05). Groups was statistically significant (1.6% [range, 0.6%-3.2%], respectively) at the concentration glucose CVs (median CV, 4.5% [range, 0.8%-14.5%] and laboratory glucose and POC glucose CV values were 15.4 mmol/L. The second group was used for comparison to 4.4 mmol/L, and the second group had a range of 4.6 to 15.4 mmol/L. The first group consisted of a concentration range of 2.3 to 15.4 mmol/L in the glucose surveys included lower levels than the laboratory glucose surveys. Therefore, only the data for a similar range of concentrations were compared between the two groups. \( P < .001 \).

In the POC glucose surveys, 95% of the peer group’s CVs were below 10%, while 95% of the peer group’s CVs were less than 2.6% in the laboratory glucose surveys (Table 2). In the low concentration range (2.3-4.4 mmol/L), only 17% (\( n = 4 \)) of the POC method peer group CVs were equal to or less than 3%, and 57% (\( n = 13 \)) were 5% or less. In the higher concentration range (4.6-15.4 mmol/L), these rates were 10% (\( n = 15 \)) and 64% (\( n = 92 \)), respectively. When all POC peer groups were evaluated, 11% (\( n = 19 \)) and 62% (\( n = 105 \)) of the peer group CVs were 3% or less and 5% or less, respectively. In laboratory glucose surveys, 99% of the CVs were 3% or less, and all were 5% or less (Figure 2).

Three hundred five (0.59%) results exceeded performance limits (±1 mmol/L, ≤5 mmol/L; 20%, >5 mmol/L) in the POC glucose surveys. Discordant findings forms were submitted for 277 (0.53%) results (Figure 3).

One cause of discordant findings arose from failure to use the correct PT items (58%). In two different institutions, in two different surveys, PT items from previous surveys were used mistakenly instead of the current set. This affected 161 results. The second most common cause was a manufacturer- or supplier-related issue (19%), which affected 52 results in eight institutions. These institutions had switched to new devices from the same vendor but, as advised by the vendor, continued to use strips produced for the old devices to consume their existing stocks. They reported that during their verification studies, they found a positive bias using old strips in the new devices but showed acceptable performance compared with the laboratory glucose test. However, they were evaluated in a new instrument peer group in PT surveys, and bias between strips caused an increased number of outliers.

The third most frequently observed cause of discordance was switching the results of the PT items in the same survey (13%). This was caused by sample mix-up on the bench in preparation for distribution, or the exact reason could not be identified whether samples were switched prior to testing or results were recorded for the wrong sample after testing (Figure 3).

### Table 2
Glucose Analyzers and Point-of-Care Devices Used by Participating Laboratories

<table>
<thead>
<tr>
<th>Glucose meters used</th>
<th>Laboratory instruments used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbott Precision PCx-PCx, Precision PCx-PCx Plus, Precision QID, Precision Xceed Pro-PCx Plus, Precision Xceed Pro-PxP, and Precision Xtra (Abbott Point of Care, Princeton, NJ)</td>
<td>Abbott (Abbott Park, IL) Architect 8000</td>
</tr>
<tr>
<td>Bayer Elite and Elite XL (Bayer Healthcare, Whippany, NJ) LifeScan One TouchII, Surestep FLEXX, Surestep Hospital, and ULTRA (Lifescan, New Brunswick, NJ)</td>
<td>Beckman Coulter (Brea, CA) UniCel and Synchron Instruments</td>
</tr>
<tr>
<td>Nova Stat Strip (Nova Biomedical, Waltham, MA) Roche Accu-Chek Advantage, Accu-Chek Inform, Accu-Chek Inform II, Accu-Soft Advantage, and Accutrend GC (Roche Diagnostics Canada, Quebec, Canada)</td>
<td>Ortho Clinical Diagnostics (Rochester, NY) Vitros systems</td>
</tr>
<tr>
<td>Siemens (Oakville, Canada) ADVIA, Dimension, and Vista Systems</td>
<td>Roche (Quebec, Canada) Cobas Integra and Modular Systems</td>
</tr>
</tbody>
</table>

### Table 3
Comparison of Laboratory and POC Glucose

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>POC Glucose</th>
<th>Laboratory Glucose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration range, mmol/L</td>
<td>4.6-15.4</td>
<td>4.6-19.0</td>
</tr>
<tr>
<td>CV, median (range), %</td>
<td>4.5 (0.8-14.5)</td>
<td>1.6 (0.6-3.2)</td>
</tr>
<tr>
<td>95th percentile, %CV</td>
<td>10.7</td>
<td>2.6</td>
</tr>
<tr>
<td>No. of results assessed</td>
<td>43,206</td>
<td>2,661</td>
</tr>
<tr>
<td>No. of PT challenges</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>No. of peer groups assessed</td>
<td>143</td>
<td>179</td>
</tr>
</tbody>
</table>

CV, coefficient of variation; POC, point of care; PT, proficiency testing.

* The range of glucose concentrations in the POC glucose surveys included lower levels than the laboratory glucose surveys. Therefore, only the data for a similar range of concentrations were compared between the two groups. \( P < .001 \).
Other causes reported were selection of an incorrect device code during result entry (4%) and therefore being evaluated in an inappropriate peer group, random or unexplained errors (3%), instrument problems (2%), and computer keying errors (1%) (Figure 3). All laboratory glucose results were within the APLs, despite use of tighter limits (Table 1). If these same performance limits were applied to POC glucose results, 6,888 (13%) results would be outliers.

Discussion

In this study, we investigated the precision performance of current POC glucose systems and their compliance with quality requirements. We reviewed reasons for recurrent and large deviations from the peer group mean to identify improvement opportunities. Participant data contained in this study include POC glucose systems used in Ontario hospitals and licensed medical laboratories, and users of the instruments were health care professionals with capillary blood glucose testing within their scope of competency.

POC Glucose Method Imprecision

In our group of participants, the median CV (4.5%) for POC glucose methods was significantly higher than that of the laboratory glucose methods (1.6%). In another PT program in which its participants were patients and physician office laboratories, method CVs obtained in the peer groups were in the range of 3.9% to 7.7% and 4.6% to 7.2%, respectively.14 Method peer group CVs obtained in PT programs, which are derived from the results of different users and represent interparticipant imprecision, are expected to be higher than CVs obtained in a single institution. However, published intralaboratory between- and within-day POC glucose CVs observed in single institutions ranged from 3.7% to 5.8% at a concentration of 1.8 to 22.2 mmol/L and 1.8% to 4.7% at a concentration of 2.9 to 16.6 mmol/L.3,15,16 All were higher than laboratory glucose CVs that were obtained in PT surveys.

Since there is inherently high imprecision, which may cause falsely high and low glucose results, and erroneous therapeutic decisions, the National Academy of Clinical Biochemistry (NACB) guideline does not recommend the use of POC glucose as a diagnostic or screening test for diabetes.9

Analytical Performance Criteria

Performance limits that have been recommended for PT of POC glucose consist of values ranging between 10% and 20%.6,13,17-19 A published hierarchical list of methods for the determination of analytical performance limits includes the methods based on clinical needs, biological variation, expert opinion, and analytical state of the art, with the first on the list being the method of choice where available.20 Performance limits based on clinical needs determined for a specific clinical condition establish that the test is used for diagnosis or monitoring of existing disease progression.

Studies have determined the performance limits for POC glucose testing based on clinical needs for two specific clinical uses: self-monitoring of blood glucose and tight glycemic control. For self-monitoring of blood glucose, patient-derived performance limits satisfying the needs of 75% of the patients were determined as a CV and bias of 5% or less at nonhypoglycemic concentrations and as an analytical CV of 3.1% or less at hypoglycemic concentrations.21 Many current POC glucose methods did not meet the expectations of the patients in hypoglycemic
There is no published guideline for POC glucose performance limits for tight glycemic control; however, the effects of POC glucose analytical performance on the adjustment of insulin dosing in tight glycemic control were evaluated in two simulation studies. Based on these data, to administering the correct dose of insulin 95% of the time, the method bias and imprecision should be less than 1% or 2%, depending on glucose concentration and insulin dosing rules. In the second study, the relation between insulin dosing errors during tight glycemic control and POC glucose total error (TE) levels of 10%, 15%, and 20% was evaluated. Investigators separated glucose values they used for simulation modeling into 12 insulin dosing categories based on the institution’s tight insulin protocol that defined insulin doses for target glucose concentrations. They observed that one-category insulin dosing errors were common in all error conditions. On the other hand, larger errors were occasionally seen in the systems operating within the 15% TE tolerance limits.

The NACB guideline (2011) recommends that in 95% of the samples, differences between reference laboratory glucose and POC glucose results should stay within 15% or less of the TE limits at a glucose concentration of 5.6 mmol/L (100 mg/dL) or more and ±0.83 mmol/L (15 mg/dL) at a glucose concentration of less than 5.6 mmol/L (100 mg/dL). However, until recently, manufacturers have been using performance limits for TE that are recommended in previous Clinical and Laboratory Standards Institute (CLSI) C30-A2 or ISO 15197:2003 documents which are larger than NACB limits, for the development of the POC glucose test systems. This might be the cause of the high analytical CV in POC glucose tests (Table 3).

The ISO has recently updated its standard as ISO 15197:2013: In Vitro Diagnostic Test Systems—Requirements for Blood Glucose Monitoring Systems for Self-Testing in Managing Diabetes Mellitus. The CLSI also published the third edition of POCT 12-A3: Point-of-Care Blood Glucose Testing in Acute and Chronic Care Facilities; Approved Guideline in January 2013. In both documents, performance limits were tightened, as indicated in Table 4.

Even these new performance criteria are higher than those established for laboratory glucose tests. The CLIA requirement for glucose is 10% for concentrations higher than 60 mg/L (3.33 mmol/L) and 6 mg/dL below this level.
Discordant Findings

Using the wrong PT samples, manufacturer reagent-related issues, sample mix-up on the bench, or reporting a result for the wrong sample were most frequent causes of discordant findings. Pre- and postanalytical errors could not be easily identified in our data set because when results of the PT samples were switched by the laboratory, it could not clearly be determined if the cause was a preanalytical error arising from sample mix-up or a postanalytical error arising from reporting the result for the wrong sample. However, 76% of all discordant findings were either pre- or postanalytical errors. This might be the result of varying levels of training and experience of the users of the POC systems in comparison to laboratory personnel who perform laboratory glucose tests, as well as difficulty in training a high number of staff using POC glucose meters (Figure 2).

Failure to identify correct PT items occurred in two facilities and affected 161 (58%) results. It is probable that the high discordant results are due to sample mix-up and then distributing them to the POC testing sites. This sort of occurrence of error affecting a lot of results is very unlikely during patient testing with POC systems.

In contrast, mixing up samples on the bench or transcription errors while entering the results electronically were another frequent cause of important errors. This type of error can occur during patient testing. Incorporation of POC devices with a barcode reader for positive patient identification and increased use of connectivity features with various laboratory information systems should facilitate to reduce these pre- and postanalytical errors.

Analytical errors accounted for 21% of the discordant findings, 2% were equipment related, and 19% were manufacturer reagent related. This relatively low analytical error rate might be a result of using larger performance limits, which allows recognition of only gross errors, therefore possibly limiting the ability to identify relatively small method-related problems.

Limitations of the Study

Today, an increasing number of medical treatment decisions are made based on POC glucose testing performed outside of the central laboratory settings. It is very important to participate in a PT program for independent control of quality of the test results. However, the POC glucose PT programs have difficulty providing capillary whole blood or commutable PT items that are compatible with all different brands of POC glucose devices. IQMH PT items contain a physiologic buffered salt solution containing a small amount of purified plasma of bovine origin and glucose and red-colored dye. Because of the artificial nature of these samples and high probability of occurrence of a matrix effect, we did not evaluate method trueness but only imprecision among peer groups consisting of particular instruments.

Conclusion

The number of PT errors observed in POC glucose surveys was drastically higher than errors associated with laboratory glucose measurements. Not surprisingly, imprecision was higher in POC glucose measurement compared with laboratory glucose since the design and intended use of glucose meters cannot match laboratory analyzers in the current state of the art. Also, the observed imprecision is likely a manifestation of less stringent performance criteria allowed for POC testing.

Address reprint requests to Dr Aslan: Institute for Quality Management in Healthcare, 1500-393 University Ave, Toronto, ON, Canada MSG 1E6; baslan@iqmh.org.

This article was presented as a poster at the American Association for Clinical Chemistry Annual Meeting: July 31, 2013; Houston, TX.

References


