Journal of Diabetes Science and Technology
Volume 4, Issue 1, January 2010

SYMPOSIUM

Analytical and Clinical Performance of Blood Glucose Monitors

Suzanne Austin Boren, Ph.D., M.H.A[1,2,3] and William L. Clarke, M.D.[4]

Abstract

Background:
The objective of this study was to understand the level of performance of blood glucose monitors as assessed in the published literature.

Methods:
Medline from January 2000 to October 2009 and reference lists of included articles were searched to identify eligible studies. Key information was abstracted from eligible studies: blood glucose meters tested, blood sample, meter operators, setting, sample of people (number, diabetes type, age, sex, and race), duration of diabetes, years using a glucose meter, insulin use, recommendations followed, performance evaluation measures, and specific factors affecting the accuracy evaluation of blood glucose monitors.

Results:
Thirty-one articles were included in this review. Articles were categorized as review articles of blood glucose accuracy (6 articles), original studies that reported the performance of blood glucose meters in laboratory settings (14 articles) or clinical settings (9 articles), and simulation studies (2 articles). A variety of performance evaluation measures were used in the studies. The authors did not identify any studies that demonstrated a difference in clinical outcomes. Examples of analytical tools used in the description of accuracy (e.g., correlation coefficient, linear regression equations, and International Organization for Standardization standards) and how these traditional measures can complicate the achievement of target blood glucose levels for the patient were presented. The benefits of using error grid analysis to quantify the clinical accuracy of patient-determined blood glucose values were discussed.

Conclusions:
When examining blood glucose monitor performance in the real world, it is important to consider if an improvement in analytical accuracy would lead to improved clinical outcomes for patients. There are several examples of how analytical tools used in the description of self-monitoring of blood glucose accuracy could be irrelevant to treatment decisions.

J Diabetes Sci Technol 2010;4(1):84-97

Author Affiliations: [1]Health Services Research and Development, Harry S. Truman Memorial Veterans' Hospital, Columbia, Missouri; [2]Department of Health Management and Informatics, University of Missouri, Columbia, Missouri; [3]Center for Health Care Quality, School of Medicine, University of Missouri, Columbia, Missouri; and [4]Division of Pediatric Endocrinology, Department of Pediatrics, University of Virginia Health System, Charlottesville, Virginia

Abbreviations: (BG) blood glucose, (CEG) consensus error grid, (DCCT) Diabetes Control and Complications Trial, (EGA) error grid analysis, (FDA) Food and Drug Administration, (HbA1c) hemoglobin A1c, (ISO) International Organization for Standardization, (SBGM) self-blood glucose monitor, (SMBG) self-monitoring of blood glucose

Keywords: diabetes mellitus, glucose, glycemic control, laboratory techniques and procedures, point of care

Corresponding Author: William L. Clarke, M.D., Division of Pediatric Endocrinology, Department of Pediatrics, University of Virginia Health System, Box 800386, Charlottesville, VA 22908; email address wlc@virginia.edu

Introduction

When performed and utilized properly, monitoring of blood glucose (BG) permits people with diabetes to determine their BG level and use the information as part of their treatment program. The overall performance of self-monitoring of blood glucose (SMBG) systems is a combination of the analytical performance of the instrument, quality of the test strips, and proficiency of the user. Considerations for improving the accuracy of glucose monitoring systems emphasize technical improvements and improved patient education to decrease user error.[1] The term "accuracy" as applied to analytical performance is defined by the International Organization for Standardization (ISO) as "the difference between the expectation of measurement results and the true value of the measured quantity." In essence, it is a measure of the difference between obtained results (by the blood glucose monitor) and the true value (determined by an accepted reference method). To our knowledge, an internationally accepted reference method for the determination of whole BG concentration does not exist. There is currently a lack of consensus on the performance standards for blood glucose monitors and whether target values of an expanded ISO 15197 standard would be appropriate.

The authors undertook an analysis of the literature to understand the performance of blood glucose monitors in the real world. This article reports on the review of the published literature to provide information on current blood glucose monitor performance. Also discussed are differences between what level of accuracy is needed for appropriate analytical performance and appropriate clinical performance and why different settings might require different levels of accuracy.

Methods

Data Sources

Medline (January 2000–October 2009) was searched for eligible articles using combinations of the following search terms: (1) diabetes mellitus, or type 1 diabetes mellitus, or type 2 diabetes mellitus; (2) blood chemical analysis, or blood glucose, or blood glucose self-monitoring, or point-of-care systems; (3) calibration, or laboratory techniques and procedures, or quality control, or reference standards, or reference values, or reproducibility of results; and (4) humans. The reference lists of included studies were also searched.

Study Selection and Data Extraction

Titles and abstracts of identified citations were screened, and articles were identified as eligible based on the following criteria. Inclusion criteria were review articles of BG accuracy, original articles that reported the performance of blood glucose meters in laboratory settings, original articles that reported the performance of blood glucose meters in clinical settings, and simulation modeling of BG. The following information was collected from eligible articles for each of the categories: (1) review articles of blood glucose accuracy—specific factors affecting the accuracy evaluation of blood glucose monitors; (2) original articles that reported the performance of blood glucose meters in laboratory settings—blood glucose meters tested, blood sample, meter operators, and performance evaluation measures; (3) original articles that reported the performance of blood glucose meters in clinical settings—blood glucose meters tested, setting, sample (number, diabetes type, age, sex, and race), duration of diabetes, years using a glucose meter, insulin use, recommendations followed, and performance evaluation measures; and (4) simulation modeling of blood glucose—sample, results, and conclusions.

Results

Comprehensive literature searches identified 563 articles. Articles were screened and 31 articles met eligibility criteria. Twenty-three of the 31 articles (74%) were published in the past 4 years.

Review Articles of Blood Glucose Meter Accuracy

Five of the six review articles focused on factors affecting evaluation of the accuracy of blood glucose monitors.[1–5] Factors can be categorized as patient/user knowledge, methods and analyses, sources of interference, and reporting of results. Patient/user knowledge factors include educating patients in the proper techniques of glucose meter use,[2,3] proper handling and storage of test strips,[2,5] proper storage of control solution,[5] meter cleanliness,[5] timing of sample collection after eating or insulin dose,[2–5] size and placement of blood sample,[5] and removal of blood from the strip.[5] Education would provide an opportunity to reduce operator error[2,4] and falsification of results by the patient.[3] A poor user interface[4] can also impact proper glucose meter use. Factors related to the methods and analyses include blood versus plasma/

serum glucose,[1,2,4] source of sample (arterial, capillary, and venous),[1,2,4] calibration of instrument,[1,5] different enzymatic methods,[2] between-lot variability of glucose strips,[1,4,5] analytic imprecision of laboratory reference method,[2] and study design.[2,6]

Several substances and patient factors are sources of interference. Substances include galactose,[3] xylose,[3] sodium fluoride,[1] glutathione,[1] cysteine,[1] uric acid,[1,3] hemolysis,[1,2,5] and drugs (e.g., ascorbic acid,[1–5] aspirin,[2,5] acetaminophen,[1,3–5] dopamine,[4] icodextrin,[4,5] maltose,[3,5] mannitol,[4] and tolbutamide[1]). Patient factors include hypotension,[1,4] hypertriglyceridemia,[1,2] hypothermia,[4] pH,[4] oxygen (e.g., hypoxia and oxygen therapy),[1–4] hematocrit (e.g., anemia and polycythemia),[1,3-5] and bilirubin.[3] Factors of reporting results of accuracy include statistical methods (e.g., correlation coefficients, linear regression, percent deviation, and mean differences),[1,2] bias plots,[2] and error grids.[1,2,4] A sixth review article examined best practices for conducting and reporting glucose monitor evaluation studies.[6] A checklist is presented in the article that outlines an approach for a protocol that is evidence based and provides enough details for reproducibility of methods and results.[6]

Performance of Blood Glucose Meters in Laboratory Settings

Fourteen articles reported on the performance of blood glucose meters in laboratory settings[7–20] (**Table 1**). In each article, 1[7] to 30[19] blood glucose meters were tested. Most commonly, 3,[11,17,18] 4,[7,9,14,16,20] or 5[10,12,13] blood glucose meters were compared. Meter operators were specified in only three of the articles.[8,9,13] The most common performance evaluation types included accuracy (e.g., bias,[9,10,12,15,18,19] Bland–Altman plots,[7,8,10,14–16,20] or error grid analysis[8–10,15,18]), analytical range,[9,12,16] effect of chemical interference on meter accuracy,[7,8,12,20] effect of hematocrit interference on meter accuracy,[7,11,12,14,17,19,20] effect of PO_2 on meter accuracy,[14,17] locally smoothed median absolute difference curves,[8,14,15] precision,[7,8–12,16,20] and performance criteria (e.g., ISO 15197,[8,11,14–16] The Netherlands Organization for Applied Scientific Research,[19] or other professional and regulatory agencies[9]). Performance evaluation measures of blood glucose meters in laboratory settings are presented in **Table 2**. Two[9,18] of the five[8–10,15,18] studies that provided results from an error grid analysis (EGA) included measurements that fell outside of zones A and B.

Performance of Blood Glucose Meters in Clinical Settings

Nine articles reported 12 studies on the performance of blood glucose meters in clinical settings[21–29] (**Table 3**). In each protocol, 1[23,26,29] to 21[21] blood glucose meters were tested. Settings included hospital,[22,26,29] outpatient clinics,[21,23–26] and home.[27,28] A total of 1990 patients participated in the studies (median 101, range 32 to 715). The patient used a blood glucose meter in all but three studies, where a health care provider used a meter instead.[22,23,29] Each of the studies involved people with type 1 or type 2 diabetes except for studies of women with gestational diabetes,[24] of children with only type 1 diabetes,[22] and of lay users and health care professionals.[26,29] Five studies included children[22,23,25,27,28]; however, only 1 of the studies involved children exclusively.[22] Two of the studies applied their findings to American Diabetes Association guidelines[21,24] and almost all other studies applied findings to the ISO 15197 standard.[22,23,25,26,28,29] Performance measures of blood glucose meters in clinical settings are presented in **Table 4**. Only 5 of the studies (presented in three of the articles)[23,25,26] provided results from an EGA. All measurements in these studies fell into zones A and B.

Simulation Modeling of Blood Glucose

Two articles reported the use of simulation to investigate insulin dosing errors.[30,31] One study calculated that insulin dosage errors would occur 8 to 23% of the time in a meter with a total analytical error of 5%.[30] Insulin dosage errors also were calculated to occur 16 to 45% of the time in a meter with a total analytical error of 10%.[30] This study concluded that glucose meters meeting current quality standards permit a large percentage of administered doses to differ from intended doses.[30] The other study focused on intravenous insulin therapy in critically ill patients and used a sample of 29,920 glucose values to simulate glucose values with error.[31] In the simulation, when 10, 15, or 20% total error was modeled, one category insulin errors occurred with 39, 46, or 46% of simulated glucose values and two category insulin errors occurred with 3, 9, or 16% of simulated glucose values.[31] The study concluded that permitting a 10% total error in glucose measurement would result in safer management of patients on intravenous insulin therapy.[31]

Analytical versus Clinical Performance Accuracy

Self-blood glucose monitors (SBGM) were introduced widely during the early 1980s and became commonly used in the 1990s as a replacement for urine testing as a means for patients with diabetes to determine their current level of glycemia. Patients were taught to use these SBGM readings to guide their decisions regarding immediate treatment. While analytical or statistical accuracy of SBGM systems is necessary for Food and

Table 1.
Blood Glucose Meters in Laboratory Settings [a]

Author	Blood glucose meter tested	Sample	Meter operators	Performance evaluation types
Bewley *et al*.[7]	StatStrip glucose meter (Nova Biomedical, Waltham, MA) compared with three conventional glucose meter technologies	Whole blood samples from 37 patients on peritoneal dialysis in the dialysis center of the nephrology clinic	NR	Accuracy (Bland-Altman plots), effect of chemical interference on meter accuracy, effect of hematocrit interference on meter accuracy, method comparison, precision (within run)
Chan *et al*.[8]	StatStrip (Nova Biomedical, Waltham, MA)	Samples from six clinical locations within a university-affiliated, tertiary care hospital: capillary blood samples (NICU), capillary blood sample (diabetes clinic), arterial blood (CVICU), arterial blood (stat lab), cord blood (delivery suite), arterial/venous blood (dialysis unit)	Four laboratory technologists and 20 front line nurses	Accuracy (Bland-Altman plots), accuracy (error grid analysis), effect of chemical interference on meter accuracy, locally-smoothed median absolute difference curves, performance criteria recommended by ISO 15197, precision (within run)
Chen *et al*.[9]	Four brands of commonly used glucose meters	Whole blood samples from 503 patients	Five medical technologists, three research assistants, and one medical doctor	Accuracy (bias), accuracy (error grid analysis), analytical range, performance criteria recommended by professional and regulatory agencies, precision
Cohen *et al*.[10]	Five currently available blood glucose meters in Australia: Accu-Chek Go (Roche), Accu-Chek Advantage (Roche), CareSens (i-Sens), GlucoMen PC (Menarini), and Optium (Abbott)	Capillary blood samples from 49 patients attending a diabetes clinic	NR	Accuracy (bias), accuracy (Bland–Altman plots), accuracy (error grid analysis), precision
Flore *et al*.[11]	Three representative point-of-care testing systems: Bayer Rapidlab 855 (Seimens, Brussels, Belgium), RapidPoint 400 (Seimens, Brussels, Belgium), and Accu-Chek Inform (Roche Diagnostics, Basel, Switzerland)	Control materials	NR	Effect of hematocrit interference on meter accuracy, performance criteria recommended by ISO 15197, precision, total error
Karon *et al*.[12]	Four glucose meter technologies representing the major hospital-based technologies currently available: Accu-Chek Inform (Roche Diagnostics, Indianapolis, IN), Precision PCx (Abbott Diabetes, Alameda, CA), SureStepFlexx (LifeScan, Milpitas, CA), and StatStrip (Nova Biomedical, Waltham, MA)	Control materials and whole blood samples from 185 patients in the intensive care unit	NR	Accuracy (bias), analytical range, correlation with reference method, effect of chemical interference on meter accuracy, effect of hematocrit interference on meter accuracy, precision (day-to-day), precision (within run)
Kimberly *et al*.[13]	Five of the most common monitors: MediSense Precision Xtra, Ascencia Dex, Prestige Smart System, OneTouch Ultra, and Accu-Chek Advantage	12 capillary blood samples from 22 people without diabetes and 71 people with diabetes	One trained operator	Among strip lot within-monitor CV, total within-monitor CV, within-strip lot within-monitor CV

Table 1 continued →

Table 1 cont.
Blood Glucose Meters in Laboratory Settings [a]

Author	Blood glucose meter tested	Sample	Meter operators	Performance evaluation types
Kost *et al.*[14]	Four hospital glucose meter systems: Accu-Chek Inform (Roche Diagnostics, Indianapolis, IN), One Touch SureStep (LifeScan, Milpitas, CA), HemoCue Glucose 201 (HemoCue AB, Angelholm, Sweden), and Precision PCx (Abbott Laboratories, Abbott Park, IL)	613 arterial blood samples from a university hospital blood gas laboratory	NR	Accuracy (Bland-Altman plots), accuracy (error grid analysis), effect of chemical interference on meter accuracy, locally smoothed median absolute difference curves, performance criteria recommended by ISO 15197, precision (within run), effect of O_2 interference on meter accuracy, locally smoothed median absolute difference curves, performance criteria recommended by ISO 15197
Kost *et al.*[15]	StatStrip glucose meter (Nova Biomedical, Waltham, MA) compared with 20 clinical laboratory chemistry analyzer reference instruments	1703 venous samples were analyzed at 35 hospitals that used 20 types of chemistry analyzers	NR	Accuracy (bias), accuracy (Bland–Altman plots), locally smoothed median absolute difference curves, performance criteria recommended by ISO 15197
Lippi *et al.*[16]	Four marketed portable glucose meters: Glucomen PC (A. Menarini Diagnostics, Florence, Italy), One Touch II (LifeScan, Milpitas, CA), Accu-Check Comfort (Roche Diagnostics, Indianapolis, IN), and Ascensia Glocometer Elite XL (Bayer Corporation, Elkhart, IN)	45 whole blood specimens	NR	Accuracy (Bland–Altman plots), analytical range, performance criteria recommended by ISO 15197, precision (within run)
Rao *et al.*[17]	Three blood glucose meters	Venous whole blood samples from healthy volunteers	NR	Effect of hematocrit interference on meter accuracy, effect of O_2 interference on meter accuracy
Savoca *et al.*[18]	3 plasma calibrated blood glucose meters: Precision Xceed (Abbott), Ascensia Contour (Bayer Corporation, Elkhart, IN), Accu-Chek Aviva (Roche Diagnostics, Indianapolis, IN)	115 capillary blood samples	NR	Accuracy (bias), accuracy (error grid analysis)
Slingerland *et al.*[19]	Thirty blood glucose meters available on the Dutch market	50 capillary blood samples and 10 venous whole blood samples	NR	Accuracy (bias), effect of hematocrit interference on meter accuracy, performance criteria recommended by The Netherlands Organization for Applied Scientific Research
Thomas *et al.*[20]	StatStrip glucose meter (Nova Biomedical, Waltham, MA) compared with three glucose meter systems: Advantage (Roche Diagnostics, Indianapolis, IN), Optium Xceed (Abbott Diabetes, Alameda, CA), and Contour TS (Bayer Healthcare Diabetes Care)	109 capillary blood specimens from 39 NICU patients	NR	Accuracy (Bland–Altman plots), effect of chemical interference on meter accuracy, effect of hematocrit interference on meter accuracy, method comparison, precision (within run)

[a] CV, coefficient of variance; CVICU, cardiovascular intensive care unit; NICU, neonatal intensive care unit; NR, not reported in article; O_2, oxygen.

Table 2.
Performance Measures of Blood Glucose Meters in Laboratory Settings[a]

Author	Performance evaluation measures
Bewley *et al.*[7]	The %CV of all four meters in the medium and high level of glucose specimens was less than 5%, but was more variable at low glucose. Linear regression analysis demonstrated a slope of 1.002 for StatStrip, whereas the other strip meter systems had lower slope values: FreeStyle, 0.946; Elite XL, 0.868; and Accu-Chek Aviva, 0.916. The addition of bilirubin did not cause a significant change in glucose readings compared with glucose baseline readings in all four meter systems. The addition of βHB did not cause a significant change in glucose readings compared with glucose baseline readings in all four meter systems. The addition of lactate did not cause a significant change in glucose readings compared with glucose baseline readings in all four meter systems. At medium and high glucose levels, the addition of maltose significantly increased the glucose readings obtained with Accu-Chek Aviva and FreeStyle meters systems, with the difference from baseline varying from 12 to 50%. Varying levels of hematocrit did not adversely affect the accuracy of StatStrip, whereas the accuracy of Elite XL, Accu-Chek Aviva, and FreeStyle glucose meter measurements were significantly adversely affected. For StatStrip, all percent bias readings fell within ISO15197 criteria, whereas for Elite XL and Accu-Chek Aviva, several of the bias plot readings exceeded ISO 15197 criteria.
Chan *et al.*[8]	The within-run imprecision varied from 2.4 to 4.7%. Significant interference was observed only for "free" Hb at 10 g/liter, which resulted in 18 and 9% reductions in the glucose value at concentrations of 8.3 and 21.8 mmol/liter, respectively. All 15 measurements were accurate within the allowable total error of 20%. A total of 386 paired observations were made, giving mean glucose concentrations of 7.03 (±3.9)[means (±SD)] mmol/liter for the glucose meter and 7.11 (±4.0) mmol/liter for the laboratory, respectively. The paired *t* test, however, indicated a significant difference, although small (0.08 mmol/liter), between the two mean glucose concentrations (pb0.05). The Deming linear regression of the laboratory (*X* axis) and the glucose meter (*Y* axis) showed a slope of 1.01 (95% CI: 0.99–1.03) and an intercept of 0.01 (95% CI: −0.13–0.15). Consensus error grid analysis gave an overall 97 and 3% of samples falling within regions A and B, respectively. Three hundred seventy-three out of 386 (96.6%) samples were within the bias limits of ISO 15197.
Chen *et al.*[9]	Meters were precise with a CV of <4% across a wide range of glucose concentrations. Slopes significantly different from 1.0 were observed for two meters with 11–13% and 211–213% at the 95% confidence interval level by the linear regression of meter results versus the HK method from 33 to 481 mg/dl (correlation coefficient >0.98 and standard error of estimation *Sy/x*, 13 mg/dl for both meters). Analysis of clinical significance of bias by Clarke error grid showed that results of the four meters were outside the accurate zone (26.5, 2.4, 1.5, and 5.6%). Only a small number of results showed clinically significant bias in the hypoglycemic range. Meters performed consistently throughout the study and were precise, although precision varied at extremely high or low glucose concentrations. Two of the glucose meters had substantial systematic bias when compared with an HK method, indicating a need for improving calibration and standardization. Analytical performance varied over the physiological range of glucose values. No meter met the <5% bias criteria of the ADA.
Cohen *et al.*[10]	The CVs of most meters were acceptable at <5%. Bias ranged from 4.0 to 15.5% with only one meter satisfying the ADA recommendation of <5% bias. Error grid analysis showed that 94–100% of readings were clinically accurate and that none of the differences from the reference method would lead to clinical errors. Bland–Altman plots showed that the magnitude of the difference between the meter and the reference method increased with increasing glucose values for two meters, but did not change significantly with glucose level for the other three meters.
Flore *et al.*[11]	The total error of the glucometers exceeded by far the desirable analytical specifications (based on a biological variability model). Lot-to-lot variation, interinstrument variation, and interoperator variability contributed approximately equally to total variance. The percentage of outliers exceeded ISO 15197 criteria in a broad glucose concentration range. When evaluating glucose POCT data vs the ISO 15197 guideline, one outlier (7%) was found in the range –4.16 mmol/liter (75 mg/dl) and five outliers (8.5%) in the range 4.16 mmol/liter (75 mg/dl).
Karon *et al.*[12]	Within-run and day-to-day precision assessed at multiple glucose levels resulted in CV values of less than 5% for all meters tested with the exception of day-to-day precision at low glucose on the PCx meter, which was 5.1%. Linear regression analysis demonstrated a slope of 0.90 and an intercept of <10 mg/dl glucose for StatStrip and the Accu-Chek meters. The PCx and SureStepFlexx meters had lower slopes and higher intercepts. There were significantly more values within 10% of the reference method on the StatStrip (170 of 185) compared to the SureStepFlexx (134 of 185), Accu-Chek (127 of 185), or PCx (79 of 185) methods. At low glucose (54 mg/dl), the mean glucose difference changed by more than 10 mg/dl between lowest and highest hematocrit values tested on the PCx and SureStepFlexx meters. At higher (247 and 483 mg/dl) glucose concentrations, the Accu-Chek, PCx, and SureStepFlexx meters demonstrated greater than 10% change in the mean glucose percentage difference between lowest and highest hematocrit values. At low glucose, changes in mean glucose difference were statistically significant for the PCx and SureStepFlexx ($p < 0.001$) between lowest and highest hematocrit tested. Changes in mean glucose percent difference between lowest and highest hematocrit tested were statistically significant ($p < 0.001$) for Accu-Chek, PCx, and SureStepFlexx technologies at higher glucose levels and marginally significant ($p = 0.0203$) at a glucose concentration of 483 mg/dl for the StatStrip. Regression analysis resulted in slopes and intercepts that were significantly different from zero ($p < 0.0001$) for the PCx and SureStepFlexx meters. For the StatStrip and Accu-Chek meters, the slope of percent bias versus hematocrit was not significantly different from zero ($p > 0.05$). Acetaminophen did not produce a clinically significant interference on any of the four meter technologies studied.

Table 2 continued →

Table 2 cont.
Performance Measures of Blood Glucose Meters in Laboratory Settings[a]

Author	Performance evaluation measures
Karon *et al.*[12] cont.	Lactate did not produce a clinically significant interference on any of the four meter technologies studied. βHB did not produce a clinically significant interference on any of the four meter technologies studied. At low glucose (70 mg/dl), ascorbic acid produced a clinically significant (>10 mg/dl) interference with the Accu-Chek, PCx, and SureStepFlexx glucose meters. At higher (141 and 237 mg/dl) glucose concentrations, ascorbic acid produced a clinically significant (>10%) interference on the Accu-Chek and PCx glucose meters. Maltose produced a clinically significant interference only on the Accu-Chek meter. Epinephrine levels up to 1 μg/dl had little effect on any of the four glucose meters.
Kimberly *et al.*[13]	Total CVs and within-strip lot CVs were not statistically different among monitors, ranging from 3.1 to 11.3% and from 2.1 to 8.5%. There were statistically significant differences among monitors for among-strip lot CVs (range 0 to 7.5%). The degree of significance increased as the concentration range increased [3.9–5.5 mmol/liter: $p < 0.05$; 5.6–7.7 mmol/liter: $p = 0.003$; 7.8–11.1 mmol/liter: $p < 0.001$]. The average percent difference between monitor pairs was statistically significant ($p < 0.05$) in more than half of the paired comparisons (range of significant differences 5.7 to 32.0%).
Kost *et al.*[14]	Performance in hypoglycemic and hyperglycemic ranges erratically exceeded the recommended LS MAD error tolerance limit (5 mg/dl). Some systems showed acceptable (within LS MAD tolerance) or nearly acceptable performance in and around a tight glycemic control interval of 80–110 mg/dl. Performance patterns varied in this interval, creating potential for discrepant therapeutic decisions. Bias typically changed from positive to negative and then decreased as hematocrit increased.
Kost *et al.*[15]	ISO 15197 guidelines were met; 98.6% (410 of 416) of observations were within tolerance for glucose <75 mg/dl and 100% were within tolerance for ≥75 mg/dl. Paired differences (handheld minus reference) averaged −2.2 (SD 9.8) mg/dl; the median was −1 (range, −96 to 45) mg/dl. LS MAD curve analysis revealed satisfactory performance below 186 mg/dl. Above 186 mg/dl, the recommended error tolerance limit (5 mg/dl) was not met. All points fell in Clarke error grid zone A. Linear regression showed $y = 1.018x - 0.716$ mg/dl, and $r^2 = 0.995$.
Lippi *et al.*[16]	The within-run imprecision ranged from 2.2 to 3.2%. Passing and Bablok regression analysis yielded slope values from 0.93 to 1.07 and correlation coefficients between 0.994 and 0.998. When compared with the secondary reference analyzer, mean variations were between −4.9 and 14.1%, fulfilling the 5.5% current desirable analytical quality specifications for total error in three out of four cases.
Rao *et al.*[17]	Different O_2 tensions did not significantly affect glucose measured in meter 1 and meter 2. Meter 3 had significantly increased glucose concentrations at normal O_2 tension compared to saturated concentrations. Meter 2 showed a decreasing trend in mean percentage error when hematocrit levels were increased from 30 to 55%. Meters 1 and 3 did not show noticeable differences at various hematocrit levels. Hematocrit measured by meter 3 had good correlation ($r = 0.998$) with a slope of 0.989 and intercept 0.827 and was in agreement with the reference method.
Savoca *et al.*[18]	Glucose values from all three POCT devices were found to be significantly higher than results from the laboratory method ($p < 0.0001$), but there was no significant difference among the glucometers ($p < 0.2$–0.36). An absolute relative deviation of more than 10% (8%) from the reference method was seen in 37% (46%), 35% (44%), and 42% (50%) of values from the Abbott, Bayer, and Roche devices. Parkes error grid analysis revealed therapeutically relevant deviations in 13, 8.7, and 10.4% of values from the Abbott, Bayer, and Roche devices that fell in zone B or C.
Slingerland *et al.*[19]	Sixty percent passed for accuracy (maximum 15% deviation from the HK method). Eighty-three percent passed for reproducibility (maximum CV 10%). Eighty-three percent passed for hematocrit dependency (range 0.35–0.50 L/L)(maximum CV 10% at maximum for glucose values <6.5 mmol/liter). Seventeen percent passed for the hematocrit dependency range stated by manufacturers. Twenty percent passed underfilling protection (maximum 10% from result at minimal volume or error mark).
Thomas *et al.*[20]	The percent coefficient of variance (%CV) for all four meters across the three glucose ranges was similar (≤5%) with the exception of the Contour glucose meter, which had a %CV of >10% in the low glucose range. Linear regression analysis demonstrated a slope of 0.960 for Contour, 0.920 for StatStrip, 0.705 for Optium Xceed, and 0.791 for Advantage. The presence of βHB had a minimal effect on the accuracy of all four glucose meters. The presence of ascorbate had a minimal effect on the accuracy of all four glucose meters. The presence of maltrose did not adversely affect the accuracy of the StatStrip, Contour, and Optium meters, but did affect the accuracy of the Advantage glucose meter measurements. Varying levels of hematocrit did not adversely affect the accuracy of StatStrip and Countour glucose measurements, whereas the accuracy of Optium and Advantage glucose meter measurement were adversely affected to a clinically significant degree. Using a Bland–Altamn plot analysis of percent bias, the StatStrip showed closer accordance with ISO 15197 criteria and closer accuracy to the laboratory HK method compared with Contour.

[a] ADA, American Diabetes Association; βHB, β-hydroxybutyrate; CV, coefficient of variance; HK, hexokinase; LS, locally smoothed; MAD, medial absolute difference.

Table 3.
Blood Glucose Meters in Clinical Settings [a]

Author	Blood glucose meter tested	Setting	Sample	Age	Sex	Duration of diabetes	Years using glucose meter	Using insulin
Alto *et al.*[21]	Twenty-one different glucose monitors compared with laboratory value from handheld glucose monitor (One Touch II hospital blood glucose monitoring system)	Two family practice residency sites	111 patients with T1 or T2 diabetes	56.0 (±14.1) years	54.6% female	NR	2.8 (±3.0) (range 0 to 15 years)	45 (40.9%)
DirecNet[22]	Two blood glucose meters: FreeStyle Flash (Abbott Diabetes Care, Alameda, CA) and OneTouch Ultra (LifeScan, Milpitas, CA)	Inpatient (24-hour stay)	50 children with T1 diabetes	14.8 (±1.7) years (range 10 to 17 years)	44% female	7.0 (±3.7) years	NR	50 (100%)
Garg *et al.*[23]	Ascensia Breeze (Bayer HealthCare LLC, Elkhart, IN)	Two outpatient diabetes centers	100 patients with T1 or T2 diabetes	39 years (range 15 to 79 years)	50% female	12 years (range <1 to 56 years)	NR	NR
Henry *et al.*[24]	Six different unspecified home blood glucose meters compared with the HemoCue B glucose analyzer	Outpatient clinic visits	500 patients with diabetes during pregnancy	32.4 years	100% female	NR	NR	NR
Kilo *et al.*[25] (lay user protocol)	Two blood glucose meters: Ascensia Contour and Ascensia Microlet Vaculance (Bayer HealthCare LLC, Elkhart, IN)	Four outpatient diabetes clinics	101 patients with diabetes T1 diabetes, T2 diabetes, or gestational diabetes	44 years (range 17 to 79 years)	48% female	9 years (range 0.1 to 50 years)	8.1 years (range 0.1 to 45 years)	NR
Kilo *et al.*[25] (high altitude protocol)	Two blood glucose meters: Ascensia Contour and Ascensia Microlet Vaculance (Bayer HealthCare LLC, Elkhart, IN)	Four outpatient diabetes clinics	54 patients with T1 or T2 diabetes	56 years (range 31 to 80 years)	52% female	NR	NR	NR
Kilo *et al.*[25] (alternative site protocol)	Two blood glucose meters: Ascensia Contour and Ascensia Microlet Vaculance (Bayer HealthCare LLC, Elkhart, IN)	Four outpatient diabetes clinics	40 patients with T1 or T2 diabetes	Range 12 to 77 years	68% female	NR	NR	NR
Kilo *et al.*[26] (lay user protocol)	Ascensia Countour (Bayer HealthCare LLC, Elkhart, IN)	Two outpatient clinical sites	101 lay users and health care professionals	45.1 years (range 17 to 79)	48% female	9 years (range 0.1 to 50 years)	8.1 years (range 0.1 to 45 years)	NR

Table 3 continued →

Table 3 cont.
Blood Glucose Meters in Clinical Settings [a]

Author	Blood glucose meter tested	Setting	Sample	Age	Sex	Duration of diabetes	Years using glucose meter	Using insulin
Kilo *et al.*[26] (hospital bedside protocol)	Ascensia Countour (Bayer HealthCare LLC, Elkhart, IN)	Inpatient hospital bedside	60 patients with T1 diabetes, T2 diabetes, or hyperglycemic condition	56 years (range 18 to 93)	40% female	NR	NR	NR
Kristensen *et al.*[27]	Six blood glucose meters: OneTouch Basic/II/Profile and GlucoTouch (LifeScan, Milpitas, CA); Accu-Chek Sensor (Roche Diagnostics, Basel, Switzerland); Glucometer Dex/Dex2 and Ascensia Elite (Bayer Diagnositcs, Leverkusen, Germany); Precision QID, MediSense Pen, MediSense Card, and Precision Xtra (Abbott Laboratories/MediSense, Abbott Park, IL)	Home	126 patients with T1 or T2 diabetes	55 years (range 6 to 84)	50% female	9 years (range 2 to 61 years)	9 years (range 2 to 25 years)	83 (66%)
Kristensen *et al.*[28]	Nine blood glucose meters: OneTouch Ultra and OneTouch Gluco Touch (LifeScan, Milpitas, CA), Precision Xceed/Precision Xtra Plus (Abbott, Abbott Park, IL), Accu-Chek Aviva, Accu-Chek Sensor/Accu-Chek Sensor Comfort Glucose, and Accu-Chek Compact Plus (Roche Diagnostics, Basel, Switzerland), Glucocard X-Meter/Glucocard X-Sensor (Arkray, Kyoto, Japan), HemoCue Monitor (HemoCue AB, Angelholm, Sweden), Ascensia Contour/Ascensia Microfill (Bayer Healthcare, Leverkusen, Germany)	Home	715 patients with T1 or T2 diabetes	53.8 years (range 10 to 75)	45% female	NR	NR	NR
Maynaar *et al.*[29]	AccuChek (Roche Diagnostics, Mannheim, Germany)	Intensive care unit in a general hospital	32 critically ill patients	71.6 (±11.9) years	72% female	NR	NR	NR

[a]ADA, American Diabetes Association; NR, not reported in article; T1, type 1; T2, type 2.

Table 4.
Performance Measures of Blood Glucose Meters in Clinical Settings[a]

Author	Performance evaluation measures
Alto *et al.*[21]	52.8% of patients had SMBG values that varied less than 10% from control monitor values. 31.5% of SMBG values varied 10 to 20% from the control values. 16% of patients had SMBG values that varied in excess of 20% from control values. 49 (46%) of the random blood glucose values were less than 180 mg/dl. 89 (82.4%) of the participants' monitors reported glucose values that were less than those of the control monitor.
DirecNet[22]	The FreeStyle meter tended to read slightly higher than the reference method [median difference = +3 mg/dl ($p < 0.001$)]. The Ultra meter tended to read slightly higher than the reference method [median difference = +2 mg/dl ($p < 0.015$)]. The Ultra and FreeStyle meters had similar accuracy results (median values compared with reference glucose values: absolute difference 6 vs 6 mg/dl, relative difference +1% vs +2%, RAD 5% and ISO criteria met 99% vs 98%). Both the Ultra and the FreeStyle were within ±10% of the reference for 81% of the pairs. During hypoglycemia (162 cases with reference glucose ≤70 mg/dl) the median absolute difference was 5 mg/dl for both HGM devices and ISO criteria were met by 96 and 99% of Ultra and FreeStyle values, respectively. For the 70 pairs where the reference was ≤60 mg/dl, 67 (96%) of Ultra and 70 (100%) of FreeStyle values were ≤70 mg/dl. False positive rates (HGM ≤60 mg/dl, but reference >70 mg/dl) were 5% (3/61) for the Ultra and 5% (3/62) for the FreeStyle. There were 18 Ultra glucose measurements and 20 FreeStyle glucose measurements from capillary blood, mostly made during hypoglycemia. All of the Ultra and 90% (18/20) of the FreeStyle measurements met ISO criteria when compared with reference glucose values obtained from venous blood.
Garg *et al.*[23]	For the subject meter glucose results, 93.4% (n = 196) met the ISO 15197 accuracy criteria, and 94.4% (n = 196) of health care professional meter results met criteria. In the error grid analysis, 92% of data points fell in Zone A and 8% fell in Zone B for both subject and health care professional results. No results were in Zone C, D, or E for either testing group. Each of the 100 subjects tested the low, normal, and high glucose control solutions in duplicate. The subjects obtained acceptable precision results. The within-run CV was 3.4% for the low control, 2.8% for the normal control, and 2.9% for the high control. The total CV was 6.4% for the low control, 4.5% for the normal control, and 4.7% for the high control.
Henry *et al.*[24]	At the 10.5% deviation level, 34% of SMBG meter readings were out of range; 54% of these would have implied erroneous treatment. At the 15.5% deviation level, 18% were out of range; 63% of these would have implied erroneous management.
Kilo *et al.*[25] (lay user protocol)	Lay users obtained results within the recommended limits (±15 mg/dl for BG values <75 mg/dl; ±20% for BG values ≥75 mg/dl) 96.9% (n = 392) of the time. HCP results met ISO criteria. All measurements fell in zones A and B on the Parkes error grid, indicating that biases obtained were all clinically insignificant. For lay users, 97.4% of SMBG results were in zone A and 2.6% in zone B; for HCPs, 97.2% of SMBG results were in zone A and 2.8% in zone B. The 95% confidence intervals of the slopes show the absence of any hematocrit effect at low glucose levels, but suggest a small effect at high glucose levels.
Kilo *et al.*[25] (high-altitude protocol)	Ninety-seven percent of measurements met ISO 15197 accuracy criteria. The slope and y intercept of the regression line of combined data and 95% confidence intervals [$y = 0.92x$ (0.90 to 0.94) + 1.83 (–0.83 to 4.20), r = 0.984] suggest a systematic bias of about –8% compared with the laboratory method. However, all Ascensia Contour glucose measurements fell within zones A and B on the error grid, indicating clinically accurate results.
Kilo *et al.*[25] (alternative site protocol)	Six of the 40 subjects required assistance for at least one of the alternative site measurements. Four of those instances were due to manifest physical impediments. The remainder, 34 of 36 (94.4%) subjects, were able to perform all of the measurements. Most of the subjects were able to obtain an acceptable blood sample on the first attempt, but a significant fraction (8–36%, depending on the sampling site) required two or more punctures. Samples from the palm gave the closest agreement to fingerstick data, with 97.5% of measurements meeting ISO criteria for accuracy. Palm site success data were followed by the thigh (90.0%), the abdomen (87.2%), and the forearm (82.5%).
Kilo *et al.*[26] (lay user protocol)	Using criteria in ISO 15197, lay users obtained results within the recommended accuracy limits (±0.83 mmol/liter for BG values <4.2 mmol/liter; ±20% for BG values ≥4.2 mmol/liter) 97.2% (n = 392) of the time. HCP results met ISO accuracy criteria 96.7% (n = 392) of the time. For lay users, 97% of SMBG results were in zone A and 3% in zone B; for HCPs, 98% of SMBG results were in zone A and 2% in zone B.
Kilo *et al.*[26] (hospital bedside protocol)	Ninety-four percent of the readings of the plasma glucose calibrated meters met ISO accuracy criteria, falling just short of the recommended 95% level. Plasma glucose data exhibited a proportional bias (slope = 1.10) and a small off-setting constant bias (y intercept = –0.49 mmol/liter), possibly due to negative bias of the laboratory method. None of the measurements fell into zones C, D, or E. Three percent were in zone B and 97% were in zone A, implying that all measurements were clinically acceptable.

Table 4 continued →

Table 4 cont.
Performance Measures of Blood Glucose Meters in Clinical Settings

Author	Performance evaluation measures
Kristensen *et al.*[27]	Decreased the percentage of poor results among diabetes patients significantly, from 11.2 to 1.9% in the first and last surveys, respectively. Between-participant CVs improved from 5.5 to 3.7% and were comparable to results from office laboratories.
Kristensen *et al.*[28]	The imprecision obtained by patients (CVs of 3.2–8.1%] was generally higher compared to that by MLT (CVs of 2.3–5.9%). Three of the nine instruments did not achieve the quality goal based on the recommendation in the ISO 15197 guideline in the hands of diabetes patients. The bias from the comparison method ranged from –10.4 to +3.2%.
Maynaar *et al.*[29]	Mean AccuChek whole blood glucose was 126 ± 36 mg/dl (7.0 ± 2.0 mmol/liter); mean central laboratory serum glucose was 137 ± 38 mg/dl (7.6 ± 2.1 mmol/liter). Mean difference was 11 mg/dl (0.61 mmol/liter)(8%) (95% confidence interval = 9–13 mg/dl, $p < 0$ .001), with the intraclass correlation coefficient being 0.934. Of Accu-Chek measurements, 225 (94.1%) were within ISO 15197 margins. Mean hematocrit was 0.30 ± 0.05. Hematocrit did not influence AccuChek results in the 0.20 to 0.44 range.

[a]CV, coefficient of variance; EQA, external quality assessment; HCP, health care provider; MLT, medical lab technician; RAD, relative absolute difference.

Drug Administration (FDA) approval and may be useful for demonstrating relative differences among systems, presentations of accuracy using analytical terms alone may not describe accuracy in its entirety.

Several examples of how analytical tools used in the description of SBGM accuracy could be irrelevant to treatment decisions are presented. First, correlation coefficients, which describe relationships between two variables—reference BG and SBGM BG— can be close to unity and highly significant when large numbers of data points are analyzed, yet individual data point may differ by large amounts.[32] Pohl and colleagues[33] have shown that even when the correlation coefficient of a large set of reference and SBGM determinations is highly significant across the entire BG range, correlation coefficients for that same data may differ significantly in the three critical BG ranges—hypoglycemia, euglycemia, and hyperglycemia. Linear regression equations present the slope of the "best-fitting" line relating two sets of data, but even when that slope approaches unity, the relationship between two data points is not assured. ISO standards acknowledge that percentage differences between reference and SBGM BG values do not apply across the entire BG reference range; however, these criteria are not sufficient to prevent the possibility of serious errors in treatment decision making.[34] For instance, if the reference value is 74 mg/dl, a level that would signal impending hypoglycemia, a SBGM BG value would be considered accurate if it were between 59 and 89 mg/dl. These two values signal entirely different clinical responses. Likewise if the reference BG were 76 mg/dl, then a 20% deviation of SBGM values from 51 to 91 mg/dl would be considered accurate.

These examples clearly suggest that traditional analytical methods, even though widely accepted as standards for reporting accuracy of SBGM systems, may confer a level of accuracy that could complicate a patient's ability to achieve BG targets, including the detection and prevention of hypoglycemia.

Patients and clinicians use SBGM systems to guide very specific clinical decisions. Thus a different method for describing the accuracy of SBGM is needed to reduce potential clinical errors that could be associated with traditional statistical analyses. EGA was the first method developed to quantify the clinical accuracy of patient-determined BG values.[32] EGA categorizes the relationship between a patient-generated BG level and a reference BG levels in terms of the clinical status that would result from a treatment decision based on the patient-generated result. Subsequently, Parkes and colleagues[35] developed the consensus error grid (CEG), a similar method for describing the accuracy of SBGM based on clinical decision making. Both of these analyses are designed to emphasize the importance of obtaining clinically accurate information across the entire BG range (hypoglycemia, euglycemia, and hyperglycemia).[36]

The EGA is a graph of reference vs SBGM BG data pairs, which is divided into five zones of clinical accuracy (**Figure 1**). The basic assumptions of EGA are that the target BG range is between 70 and 180 mg/dl and that any patient-generated BG value outside that range will be treated according to protocols or algorithms selected by the health care provider and the patient. Zone A (upper and lower) data pairs represent patient-generated values, which are within 20% of the reference values and/or

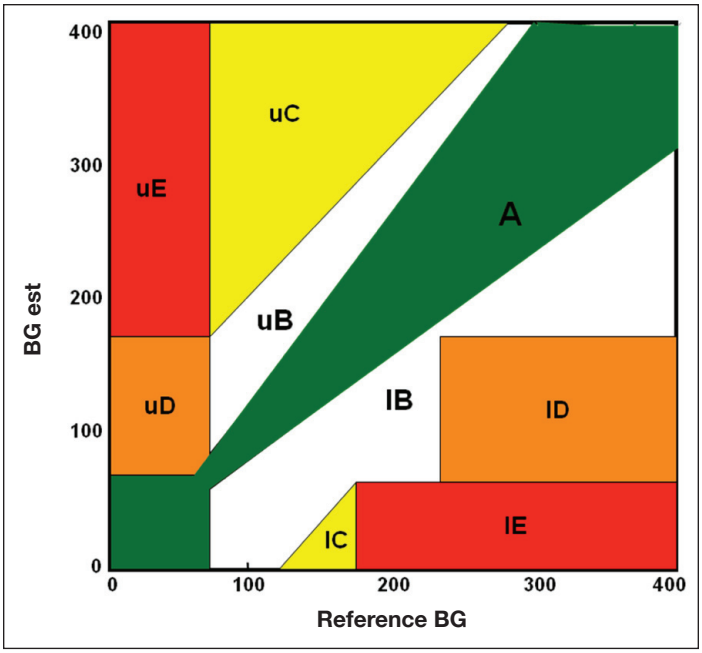

Figure 1. Five zones of clinical accuracy of EGA.

≤70 mg/dl when the reference is ≤70 mg/dl. Zone A points are categorized as clinically accurate because they could lead to accurate treatment decisions. Zone C (upper and lower) data pairs represent possible "overcorrection errors," as patient-generated values in these zones could trigger treatment responses that might result in the subsequent BG value being outside the target range. Zone D (upper and lower) values are "failure to treat" errors because the patient-generated values are within the target range when the reference value is either low (≤70 mg/dl) or high (≥240 mg/dl). Zone E values are "erroneous errors" where the patient-generated values are either high (>180 mg/dl) when the reference is low (≤70 mg/dl) or low (≤70 mg/dl) when the reference is high (>240 mg/dl). Patient self-treatment based on these errors could result in serious hypoglycemia or hyperglycemia. Zone B data pairs are those where the patient-generated value deviates from the reference by more than 20% but may not result in clinically significant treatment errors. They are designated clinically acceptable.

Since the late 1980s, EGA has been used by most manufacturers of SBGM devices to demonstrate the clinical accuracy of their systems and reported along with more traditional statistical analyses to the FDA as part of premarketing applications.[37] In the original presentation of the EGA, results from a variety of SBGM were presented.[32] In no case was the clinically accurate/acceptable (zones A+B) percentage less than 94%. Indeed, even when results of visually interpreted BG strips were plotted on the EGA, few errors in clinical decision making would have been expected. Thus, it appears, from over 20 years of data collection, that the clinical accuracy of SBGM systems analyzed using either EGA or CEG is sufficient to permit patients to make appropriate treatment. SBGM systems were used by subjects participating in the Diabetes Control and Complications Trial (DCCT) to achieve intensive control of their BG levels and to reduce their risk of diabetic complications.[38]

It is important to point out that clinical accuracy depends on the BG target range. Although the target range used most commonly for FDA submissions of clinical accuracy of SBGM systems using EGA is 70 to 180 mg/dl, that range is not fixed. It may differ depending on the clinical situation and the treatment goals of an individual patient. In the original description of EGA, an example is presented where the target range of a data set is changed from 70–180 to 60–120 mg/dl, a target being used as a goal for managing glycemia during pregnancy. As a result of that change, the clinically accurate/acceptable (zones A+B) percentage increased while the clinically inaccurate (zones C, D, and E) percentage decreased.[32] Thus selecting a different target range for BG alters the size and/or position of the EGA zones and the clinical accuracy of a given SBGM system, but does not change its statistical accuracy. A careful examination of the EGA (**Figure 1**) and its zones of clinical accuracy permits one to evaluate the potential consequences of selecting different target ranges. For instance, use of a narrow target range (80–110 mg/dl) for hospitalized intensive care patients would shift lower zone C (overcorrection) as well as lower zone E (erroneous) to the left such that the reported fivefold increase in hypoglycemia might have been anticipated.[39]

Because the clinical accuracy of SBGM assumes that the patient will take treatment action to return his/her BG into the target range, it is clear that patient education and technical performance play a role in the clinical accuracy of SBGM systems. These patient factors and their contribution to accuracy have been reviewed in this article.

Discussion

The American Diabetes Association has suggested that SBGM systems be developed to achieve an analytical plus user error of less than 10% at BG levels between 30 and 400 mg/dl.[40] The analytical error goal for such SBGM systems is 5% or less. We have reviewed recent publications describing the analytical and clinical

performance accuracy of current systems and the numerous factors, physiologic, environmental, and educational, that affect accuracy. Researchers and clinicians must ask the following obvious questions. "What are the potential benefits to patients that would accompany an improvement in analytical accuracy? Would an improvement in analytical accuracy be logically accompanied by any improvement in clinical outcomes, such as reduced morbidity and mortality in hospitalized patients, or lower hemoglobin A1c (HbA1c) levels, reductions in glycemic verticality, and/or reductions in acute and chronic complications in outpatients?"

Simulation studies suggest that a reduced SBGM analytical error could be associated with fewer incorrect insulin doses given to ICU patients being treated with tight glucose control protocols.[39] However, in these simulations, SBGM error had to be reduced to <2% to ensure 95% correct insulin dosing. A lower analytical error might also permit a more rapid and accurate diagnosis of diabetes based on a single fasting BG level.

As stated earlier, SBGM systems were not developed to be substitutes for the precise analytical instruments used to determine BG in hospital laboratories. They were developed to provide immediate BG information to patients with diabetes so that they might make their own treatment decisions. SBGM systems have been used since the late 1970s with varying degrees of success by educated and motivated patients in clinical trials such as the DCCT and in routine self-management. The authors are unaware of any clinical studies that demonstrate a difference in clinical outcomes—HbA1c, glycemic variability, pregnancy, hypoglycemia, diabetic ketoacidosis, or chronic complications—when subjects used SBGM systems with different analytical accuracy. It is possible that this may be due to the limited amount of information provided by a single BG determination. Without information regarding recent BG values and the current rate and direction of BG change, as provided by continuous glucose monitoring technology, reductions in the analytical error of single BG determinations may be insufficient to affect changes in clinical outcomes.

Funding:

This work was supported in part by NIH Grants R01DK5162, RO1DK28288, and M01 RR00847.

Disclosure:

The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

References:

1. Bergenstal RM. Evaluating the accuracy of modern glucose meters. Insulin. 2008;3:5-14.
2. Arabadjief D, Nichols JH. Assessing glucose meter accuracy. Curr Med Res Opin. 2006;22(11):2167-74.
3. Bode BW. The accuracy and interferences in self-monitoring of blood glucose. US Endocr Dis Touch Brief. 2007;4-6.
4. Dungan K, Braithwaite SS, Chapman J, Buse J. Glucose measurement: confounding issues in setting targets for inpatient management. Diabetes Care. 2007;30(2):403-9.
5. Hirsch IB, Bode BW, Childs BP, Close KL, Fisher WA, Gavin JR, Ginsberg BH, Raine CH, Verderese CA. Self-monitoring of blood glucose (SMBG) in insulin- and non-insulin-using adults with diabetes: consensus recommendations for improving SMBG accuracy, utilization, and research. Diabetes Technol Ther. 2008;10(6):419-39.
6. Mahoney JJ, Ellison JM. Assessing glucose monitor performance--a standardized approach. Diabetes Technol Ther. 2007;9(6):545-52.
7. Bewley B, O'Rahilly S, Tassell R, DuBois J, Donald E. Evaluation of the analytical specificity and clincal application of a new generation hospital-based glucose meter in a dialysis setting. Point Care. 2009;8(2):61-7.
8. Chan PC, Rozmanc M, Seiden-Long I, Kwan J. Evaluation of a point-of-care glucose meter for general use in complex tertiary care facilities. Clin Biochem. 2009;42(10-11):1104-12.
9. Chen ET, Nichols JH, Duh SH, Hortin G. Performance evaluation of blood glucose monitoring devices. Diabetes Technol Ther. 2003;5(5):749-68.
10. Cohen M, Boyle E, Delaney C, Shaw J. A comparison of blood glucose meters in Australia. Diabetes Res Clin Pract. 2006;71(2):113-8.
11. Flore KM, Fiers T, Delanghe JR. Critical evaluation of connectivity-based point of care testing systems of glucose in a hospital environment. Clin Chem Lab Med. 2008;46(12):1763-8.
12. Karon BS, Griesmann L, Scott R, Bryant SC, Dubois JA, Shirey TL, Presti S, Santrach PJ. Evaluation of the impact of hematocrit and other interference on the accuracy of hospital-based glucose meters. Diabetes Technol Ther. 2008;10(2):111-20.
13. Kimberly MM, Vesper HW, Caudill SP, Ethridge SF, Archibold E, Porter KH, Myers GL. Variability among five over-the-counter blood glucose monitors. Clin Chim Acta. 2006;364(1-2):292-7.
14. Kost GJ, Tran NK, Abad VJ, Louie RF. Evaluation of point-of-care glucose testing accuracy using locally-smoothed median absolute difference curves. Clin Chim Acta. 2008;389(1-2):31-9.
15. Kost GJ, Tran NK, Louie RF, Gentile NL, Abad VJ. Assessing the performance of handheld glucose testing for critical care. Diabetes Technol Ther. 2008;10(6):445-51.
16. Lippi G, Salvagno GL, Guidi GC, Negri M, Rizzotti P. Evaluation of four portable self-monitoring blood glucose meters. Ann Clin Biochem. 2006;43(Pt 5):408-13.
17. Rao LV, Jakubiak F, Sidwell JS, Winkelman JW, Snyder ML. Accuracy evaluation of a new glucometer with automated hematocrit measurement and correction. Clin Chim Acta. 2005;356(1-2):178-83.
18. Savoca R, Jaworek B, Huber AR. New "plasma referenced" POCT glucose monitoring systems--are they suitable for glucose monitoring and diagnosis of diabetes? Clin Chim Acta. 2006;372(1-2):199-201.
19. Slingerland RJ, Muller W, Meeues JT, Van Blerk I, Gouka-Tseng C, Dollahmoersid R, Witteveen C, Vroonhof K. The quality of blood glucose meters in the Netherlands 5 years after introduction of the CE/IVD directive. Ned Tijdschr Klin Chem Labgeneesk. 2007;32:202-4.

20. Thomas A, Sall S, Roberts C, Drayton M, DuBois J, Clampitt R. An evaluation of the analytical performance of a new-generation hospital-based glucose meter and an assessment of its clinical reliability in a neonatal care unit. Point Care. 2009;8:68-73.

21. Alto WA, Meyer D, Schneid J, Bryson P, Kindig J. Assuring the accuracy of home glucose monitoring. J Am Board Fam Pract. 2002;15(1):1-6.

22. Weinzimer SA, Beck RW, Chase HP, Fox LA, Buckingham BA, Tamborlane WV, Kollman C, Coffey J, Xing D, Ruedy KJ; Diabetes Research in Children Network Study Group. Accuracy of newer generation home blood glucose meters in a diabetes research in children network (DirecNet) inpatient exercise study. Diabetes Technol Ther. 2005;7(5):675-83.

23. Garg SK, Carter JA, Mullen L, Folker AC, Parkes JL, Tideman AM. The clinical performance and ease of use of a blood glucose meter that uses a 10-test disk. Diabetes Technol Ther. 2004;6(4):495-502.

24. Henry MJ, Major CA, Reinsch S. Accuracy of self-monitoring of blood glucose: impact on diabetes management decisions during pregnancy. Diabetes Educ. 2001;27(4):521-9.

25. Kilo C, Pinson M, Joynes JO, Joseph H, Monhaut N, Parkes JL, Baum J. Evaluation of a new blood glucose monitoring system with auto-calibration. Diabetes Technol Ther. 2005;7(2):283-94.

26. Kilo C Sr, Dickey WT Jr, Joynes JO, Pinson MB, Baum JM, Parkes JL, Parker DR. Evaluation of a new blood glucose monitoring system with auto-calibration for home and hospital bedside use. Diabetes Res Clin Pract. 2006;74(1):66-74.

27. Kristensen GB, Nerhus K, Thue G, Sandberg S. Results and feasibility of an external quality assessment scheme for self-monitoring of blood glucose. Clin Chem. 2006;52(7):1311-7.

28. Kristensen GB, Monsen G, Skeie S, Sandberg S. Standardized evaluation of nine instruments for self-monitoring of blood glucose. Diabetes Technol Ther. 2008;10(6):467-77.

29. Meynaar IA, van Spreuwel M, Tangkau PL, Dawson L, Visser SS, Rijks L, Vlieland TV. Accuracy of AccuChek glucose measurement in intensive care patients. Crit Care Med. 2009;37(10):2691-6.

30. Boyd JC, Bruns DE. Quality specifications for glucose meters: assessment by simulation modeling of errors in insulin dose. Clin Chem. 2001;47(2):209-14.

31. Karon BS, Deobald G, Klee GG. Effect of different glucose meter error tolerances on dosing errors during intravenous insulin administration. Point Care. 2008;7(3):124.

32. Clarke W, Cox D, Gonder-Frederick L, Carter W, Pohl S. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. Diabetes Care. 1987;10(5):622-8.

33. Pohl S, Gonder-Frederick L, Cox D, Evans W. Self-measurement of blood glucose concentration: clinical significance of patient-generated measurements. Diabetes Care. 1985;8(6):617-9.

34. International Organization for Standardization. *In vitro* diagnostic test systems: requirements for blood glucose monitoring systems for self-testing in managing diabetes mellitus (ISO 15197). Geneva: International Organization for Standardization; 2003.

35. Parkes JL, Pardo S, Slatin SL, Ginsberg BH. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. Diabetes Care. 2000;23(8):1143-8.

36. Klonoff D. The need for separate performance goals for glucose sensors in the hypoglycemic, normoglycemia, and hyperglycemic ranges. Diabetes Care. 2004;27:834-6.

37. US Food and Drug Administration. Review criteria assessment of portable blood glucose monitoring in vitro diagnostic devices using glucose oxidase, dehydrogenase, or hexokinase methodology (version 02/14/96). Available from: *www.fds.cdrh/ode/glu.html.*

38. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes. N Engl J Med. 1993;329(14):977-86.

39. Scott MG, Bruns DE, Boyd JC, Sacks DB. Tight glucose control in the intensive care unit: are glucose meters up to the task? Clin Chem. 2009;55(1):18-20.

40. American Diabetes Association. Self-monitoring of blood glucose. Diabetes Care. 1994;17(1):81-6.