

# Comparative Accuracy of 17 Point-of-Care Glucose Meters

Journal of Diabetes Science and Technology  
2017, Vol. 11(3) 558–566  
© 2016 Diabetes Technology Society  
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/1932296816672237  
journals.sagepub.com/home/dst  


Laya Ekhlaspour, MD<sup>1</sup>, Debbie Mondesir, BS<sup>1</sup>,  
Norman Lautsch, BS, MS<sup>1</sup>, Courtney Balliro, RN<sup>1</sup>,  
Mallory Hillard, BS<sup>1</sup>, Kendra Magyar, NP<sup>1</sup>,  
Laura Goergen Radocchia, MSN PMHNP-BC<sup>1</sup>,  
Aryan Esmaeili, MD<sup>1</sup>, Manasi Sinha, MD, MPH<sup>1</sup>,  
and Steven J. Russell, MD, PhD<sup>1</sup>

## Abstract

**Background:** The accuracy of point-of-care blood glucose (BG) meters is important for the detection of dysglycemia, calculation of insulin doses, and the calibration of continuous glucose monitors. The objective of this study was to compare the accuracy of commercially available glucose meters in a challenging laboratory study using samples with a wide range of reference BG and hemoglobin values.

**Methods:** Fresh, discarded blood samples from a hospital STAT laboratory were either used without modification, spiked with a glucose solution, or incubated at 37°C to produce 347 samples with an even distribution across reference BG levels from 20 to 440 mg/dl and hemoglobin values from 9 to 16 g/dl. We measured the BG of each sample with 17 different commercially available glucose meters and the reference method (YSI 2300) at the same time. We determined the mean absolute relative difference (MARD) for each glucose meter, overall and stratified by reference BG and by hemoglobin level.

**Results:** The accuracy of different meters widely, exhibiting a range of MARDs from 5.6% to 20.8%. Accuracy was lower in the hypoglycemic range, but was not consistently lower in samples with anemic blood hemoglobin levels.

**Conclusions:** The accuracy of commercially available glucose meters varies widely. Although the sample mix in this study was much more challenging than those that would be collected under most use conditions, some meters were robust to these challenges and exhibited high accuracy in this setting. These data on relative accuracy and robustness to challenging samples may be useful in informing the choice of a glucose meter.

## Keywords

diabetes, blood glucose, plasma glucose, glucose meter, accuracy, MARD, hemoglobin

Tight glycemic control with intensive insulin therapy reduces the risk of diabetic complications.<sup>1,2</sup> Self monitoring of blood glucose plays an important role in the safety and efficacy of current therapy for diabetes mellitus because most dosing decisions are based on blood glucose value obtained by measuring a fingerstick capillary blood specimen on a home glucose meter.<sup>3–5</sup> The safety and effectiveness of continuous glucose monitors, and automated glucose management systems that utilize them as component parts, are also critically dependent on accurate BG measurements for calibrations. In this setting, the impact of 1 inaccurate measurement can be magnified.

The increase in the global prevalence of diabetes,<sup>6</sup> the importance of the accuracy of self monitoring glucose systems, and price pressures, have led the medical device

manufacturers to develop more brands and types of glucose meters.<sup>7</sup> The accuracy of results, the ease of technique and maintenance, and the price of both the meter and the strips are factors that are considered when choosing a glucose meter.

The purpose of this study was to evaluate the comparative accuracy of 17 commercially available glucose meters across a wide range of reference plasma glucose (PG) values

<sup>1</sup>Massachusetts General Hospital Diabetes Research Center, Boston, MA, USA

## Corresponding Author:

Steven J. Russell, MD, PhD, Massachusetts General Hospital Diabetes Research Center, 50 Staniford St, Ste 301, Boston, MA 02114, USA.  
Email: sjrussell@mgh.harvard.edu

and hemoglobins (Hb). We specifically aimed to choose a meter for calibration of the continuous glucose monitor component of a bionic pancreas system for automated glucose management.

## Methods

The study was performed September 2014 through December 2014 at Massachusetts General Hospital's Diabetes Research Center, Boston, MA. Fresh discarded whole blood samples collected in blood gas syringes and anticoagulated with heparin were obtained from the STAT lab. The protocol was reviewed by the Partners Human Research Committee and determined to be non-human subjects research.

### Hemoglobin Estimation

The hemoglobin content of the samples was tested with the HemoCue® B-Hemoglobin system (HemoCue AB, Ängelholm, Sweden). If the hemoglobin level in the sample was between 7 and 16.5 g/dl, the sample was transferred to a coded tube for glucose testing. A total of 137 samples were collected for glucose testing.

### Laboratory PG Estimations

Reference glucose measurements on whole blood were performed with the YSI 2300 STAT PLUS Glucose & L-Lactate Analyzer (Yellow Springs Instruments, OH). We chose this instrument because the US Food and Drug Administration (FDA) accepts it as the reference method, and most glucose meters are factory calibrated using this device as the standard. Daily quality assurance tests were performed on the YSI as suggested by the manufacture guidelines using glucose standards traceable to the NIST (National Institute of Standards and Technology). We used a conversion based on hemoglobin to convert YSI measurements on whole blood to plasma equivalents: Hemoglobin corrected glucose = YSI measurement  $\times$  (0.84 / (0.93-0.22  $\times$  Hemoglobin %)).<sup>8</sup>

### Preparation of Samples

Samples could be used with or without modification. To obtain additional hyperglycemic samples, blood was spiked with a solution of 3.5 M glucose in 150 mM NaCl. To obtain additional hypoglycemic samples, blood was incubated in a 37°C warm water bath. After incubating or spiking procedures, samples were retested using the YSI. Within each 1 g/dl hemoglobin bin in the 7-15 g/dl range, samples with an even distribution across reference PG levels from 20 – 440 mg/ were generated. A total of 346 reference measurements were available from 351 samples after samples with reference PG measurements less than 20 mg/dl or greater than 440 mg/dl were removed.

## Glucose Meter Measurements

Once a value in the desired reference PG range was obtained, the samples were immediately tested on 17 different commercially available blood glucose meters systems, consisting of a meter and test strips (Table 1). Using a pipette, a drop of blood from the sample was placed on a sheet of Parafilm (Bemis Company, Inc, Neenah, WI). All 17 meters were used to test this sample within 30 seconds and the values were recorded. At least 3 lots of each strip type were used during the study.<sup>9</sup>

## Data Analysis

We calculated the mean absolute relative difference (MARD = mean (|reference value – meter value|) / (reference value)  $\times$  100) versus the PG measurement derived from the YSI as the primary outcome measure. For the purposes of converting YSI whole blood glucose measurements to PG measurements we used the hemoglobin value equal to the midpoint of each hemoglobin bin. Some measurements with the point-of-care glucose meters did not produce a numerical reading; rather, they displayed a “low” or “high” message. To include these results in the numerical analyses, the results were censored by setting them to 1 mg/dl lower than the lowest value in the meter's reported glucose concentration range (eg, 19 mg/dl for a meter with a lower limit of 20 mg/dl) or 1 mg/dl higher than the upper limit (eg, 445 for a meter with an upper limit of 444 mg/dl), respectively.

Although our testing method was not the one specified by the standard, we determined whether each meter met numerical criteria for accuracy described in International Organization for Standardization (ISO) 15197: 2003 (>95% within 15 mg/dl of reference <75 mg/dl or within 20% of reference  $\geq$ 75 mg/dl)<sup>10</sup> and ISO 15197: 2013 (>95% within 15 mg/dl of reference <100 mg/dl or within 15% of reference  $\geq$ 100 mg/dl).<sup>11</sup>

To compare the MARDs of different meters ANOVA (analysis of variance) was performed. The Bonferroni correction was used to correct for multiple comparisons; adjusted *P* values < .05 were considered significant. Relative difference (RD = [meter PG – reference PG] / reference PG) plots were constructed with the YSI results converted to PG on the x-axis and relative difference of the meter versus the YSI on the y-axis. To assess the significance of differences in the accuracy of each meter in low versus normal hemoglobin ranges Student's t-test was used. The relationship between FDA approval date and MARD was analyzed using linear regression analysis. Statistical analysis was performed in Stata version 13.1 (StataCorp LP, College Station, TX, USA).

## Results

### Accuracy Analysis

We tested the accuracy of 17 different commercially available blood glucose meters systems, consisting of a meter and

**Table 1.** Glucose Meters Evaluated in This Study.

Glucose meter	Company <sup>a</sup>	Retail cost per strip (\$) <sup>b</sup>	Strip enzyme	Range (mg/dl)	Market share <sup>c</sup>
FreeStyle Freedom Lite	Abbott	0.5	GDH-FAD	20-500	13.8% <sup>d</sup>
FreeStyle Lite	Abbott	0.5	GDH-FAD	20-500	13.8% <sup>d</sup>
AgaMatrix JAZZ	AgaMatrix	1.1	GOx	20-600	<0.4%
AgaMatrix PRESTO	AgaMatrix	0.2	GOx	20-600	0.4%
BREEZE @2	Bayer	0.2	GDH-FAD	20-600	1.1%
Contour Next	Bayer	0.2	GDH-FAD	10.8-599	6.8%
HemoCue Glucose 201	HemoCue	1.5	GDH-PQQ	0-444	— <sup>e</sup>
SideKick	Nipro	0.4	GOx	20-600	PL
TRUResult	Nipro	0.2	GDH-PQQ	20-600	PL
Nova Max	Nova	0.3	GOx	20-600	<0.4%
StatStrip Xpress	Nova	1.0	GDH-FAD	10-600	— <sup>e</sup>
OneTouch Ultra2	LifeScan	1.0	GOx	20-600	39%
OneTouch VerioIQ	LifeScan	0.5	GDH-FAD	20-600	3.4%
ReliOn Micro	ReliOn	0.4	GOx	20-500	PL
ReliOn Prime	ReliOn	0.3	GOx	20-600	PL
Accu-Chek Aviva Plus	Roche	1.0	GDH-PQQ	20-600	9.0%
Accu-Chek Nano	Roche	0.9	Mut Q-GDH	20-600	4.2%

Meters are listed in alphabetic order by manufacturer. Nielsen OTC Store Brand represents 70% of the OTC Retail Market 12 months ending 04/2015. Nielsen OTC Retail Market = 12.8% of Total Retail including products dispensed on Rx. Data are not available for hospital meters.

<sup>a</sup>Abbott, Abbott Diabetes Care; Bayer, Bayer HealthCare; HemoCue, HemoCue-Radiometer Group; Nipro, Nipro Diagnostics; Nova, Nova Biomedical; Roche, Roche Diagnostics.

<sup>b</sup>The prices are the minimum of the retail store and online prices (as in May 2016), except for the HemoCue Glucose 201 and the StatStrip Xpress, which were purchased through a hospital distributor.

<sup>c</sup>Source: IMS Health NPA Market Dynamics™, National Prescription Audit™ 12 months ending 04/2015.

<sup>d</sup>Same strip.

<sup>e</sup>Intended for use in hospitals.

test strips (Table 1), using the 346 samples that were made from 137 specimens, many after spiking or incubating. The meters exhibited a wide range of MARDs ranging from 5.6% to 20.8% (Table 2, Figure 1). The range of MARDs for the high reference PG ( $\geq 180$  mg/dl) samples (5.4-20.4%) was similar to the MARDs overall. There were 8 meters with overall MARDs less than 10%. There was no significant difference between the MARDs of the 7 meters with the lowest overall MARDs ( $P > .151$ , Figure 1).

There were 5 meters with MARDs above 15%. Of these 5 meters, the MARDs of 4 were not significantly different from each other (Figure 1). For all meters, the MARD in hypoglycemic range was higher than the overall MARD, and the MARD in the normoglycemic and hyperglycemic ranges, and the spread of MARDs was wider (8.9-39.2%). However, the rank order of meter accuracy, as expressed by the MARD, did not differ markedly between PG ranges.

### Relative Difference Plots

To visualize any biases that might be present in meter PG measurements, RD plots were generated to display the differences between reference PG values at meter PG values (Figure 2). The mean RDs for all meters were between -0.11 and 0.13. The mean RD for meters with MARDs <10% was

between -0.06 and +0.06. The slope of the linear interpolation of the RD plots for all meters with MARDs less than 10% were  $\leq 0.0003$ , indicating no significant differences in the bias across different reference PG ranges (Figure 2). The plots showing the RD versus Hb are provided in Supplemental Figure 1.

### Accuracy Based on ISO Criteria

Of the 17 meters, 7 met the numerical criteria for accuracy set out in ISO 2003, although it is important to note that the samples we used were not collected as recommended by the standard (Table 3). The MARDs of these 7 meters were all less than 9%. Only Contour Next and StatStrip Xpress met the ISO 2013 numerical criteria for accuracy in the context of this study and both of these meters had MARDs less than 7%, with no statistically difference in the accuracy between them.

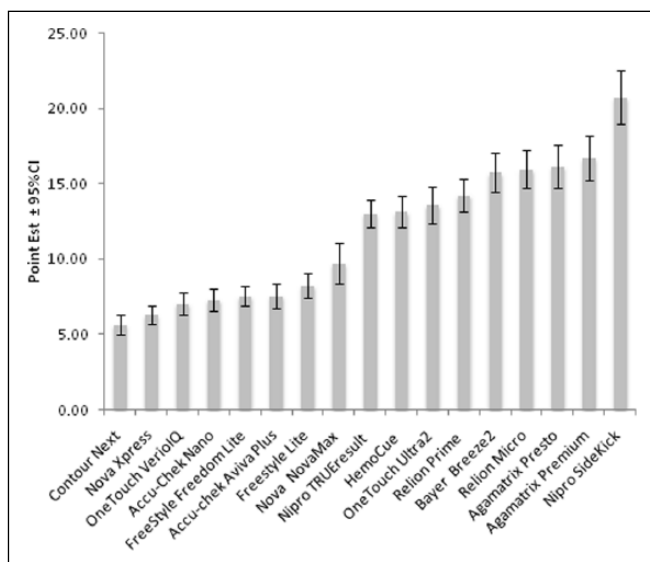
### Accuracy Comparison in Normal Versus Low Hemoglobin Ranges

There was no consistent difference in MARDs between hemoglobin ranges (normal hemoglobin versus anemic values) (Table 4). The rank order of MARDs was similar but not

**Table 2.** Accuracy of Glucose Meters in Different Reference Glucose Ranges.

Glucose meter	MARD (%)		MARD (%)		MARD (%)		MARD (%)	
	Overall	SD	BS < 70 mg/dL	SD	BS 70-179 mg/dL	SD	BS ≥180 mg/dL	SD
Contour Next	5.6	6.4	8.9	13.2	4.8	3.5	5.4	5.2
StatStrip Xpress	6.3	6.1	11.1	9.6	4.9	3.9	6.0	5.5
OneTouch VerioIQ	7.1	6.9	9.9	13.0	6.3	4.7	6.8	5.9
Accu-Chek Nano	7.3	7.1	11.6	12.8	6.7	5.8	6.8	5.8
FreeStyle Freedom Lite	7.5	6.4	15.4	7.1	7.3	6.8	6.2	4.8
Accu-Chek Aviva Plus	7.6	7.9	12.1	15.2	7.3	5.8	6.9	6.4
FreeStyle Lite	8.2	8.1	16.6	6.5	7.9	5.7	6.8	8.3
Nova Max	9.7	12.6	27.0	24.6	8.9	10.5	6.9	6.0
TRUEresult	13.0	8.4	13.2	10.8	12.8	7.3	13.1	8.4
HemoCue Glucose 201	13.2	9.9	19.9	19.4	11.6	7.9	12.6	7.3
OneTouch Ultra2	13.6	11.5	30.3	14.9	15.7	10.8	9.7	7.4
ReliOn Prime	14.3	10.2	16.2	8.8	11.4	7.9	15.1	11.1
BREEZE ®2	15.8	12.4	25.7	19.1	13.3	9.7	14.9	11.0
ReliOn Micro	16.0	12.2	29.9	11.8	14.2	11.2	14.1	10.9
AgaMatrix PRESTO	16.2	13.7	39.2	14.4	21.2	11.3	9.8	7.8
AgaMatrix JAZZ	16.7	13.9	38.0	16.3	22.3	11.6	10.5	8.3
SideKick	20.8	16.6	31.7	16.1	16.7	12.9	20.4	17.2

Meters are listed in order of overall increasing MARD.



**Figure 1.** MARD of evaluated glucose meters, shown as the point estimate of the MARD and the 95% confidence interval. Meters are listed in order of increasing overall MARD.

identical in both ranges. In both normal and anemic ranges, the MARDs of the 2 most accurate meters (Contour Next and StatStrip Xpress) were similar and not statistically different ( $P > .05$ ). Some meters (9 of 17) were nominally more accurate in the anemic range and the remainder were nominally more accurate in the normal range. The difference in MARD between the normal and anemic range did not correlate with overall MARD.

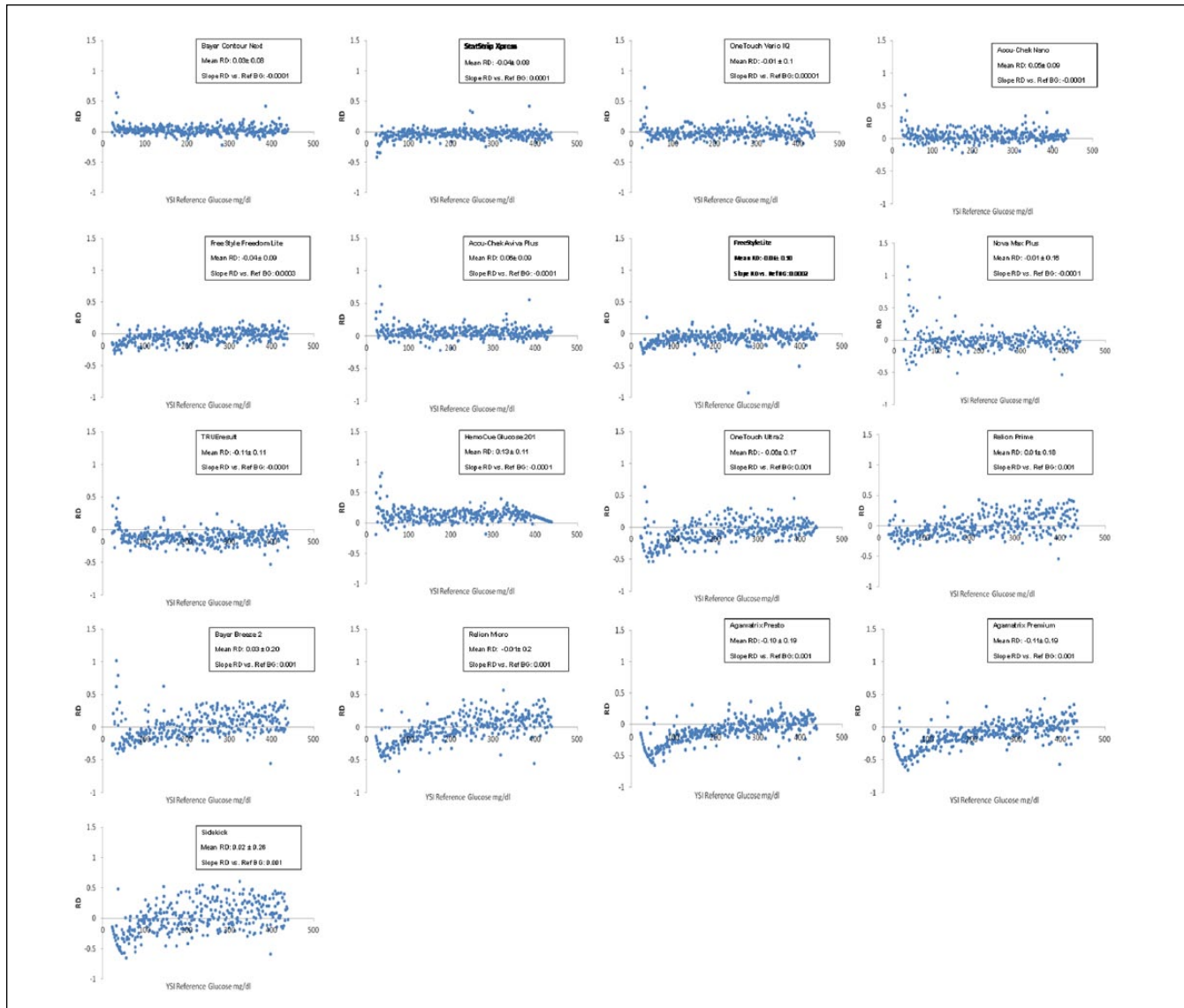
### Meter Accuracy Versus Approval Date and Cost

There was a positive correlation between meter accuracy and FDA decision date (Figure 3A). A more recent approval date correlated with lower MARD ( $P < .01$ ). There was no significant correlation between the cost of strips and the accuracy of the glucose meters (Figure 3B).

### Discussion

In this study we found that 17 commercially available glucose meters from 9 manufacturers had widely varying accuracy across a wide range of reference PG values and hemoglobin values. In this study, conducted by an independent research center without support from any meter manufacturer, we selected meters that are commonly used in the diabetes patient population. This laboratory investigation is one of the largest meter comparison studies done to date, with only 2 previous reports comparing more meters in a single study.<sup>12,13</sup> Most of the recently published studies have compared meters that are commonly used in Europe,<sup>9,12-16</sup> but not all of these meters are available in the United States, and some of the meters sold in the United States may be different from their European counterparts with the same or similar names. Previous studies conducted in the United States did not investigate all the devices that we included in our study.<sup>17-21</sup>

We chose to use the MARD in addition to RD plots and the ISO numerical criteria to assess accuracy. MARD calculation provides more detailed evaluation compared to the ISO criteria, which are binary classifications. For the meters



**Figure 2.** Relative differences (RD) plots (relative difference of meter plasma glucose minus reference plasma glucose versus reference plasma glucose). Meters are ordered by increasing overall MARD.

that were previously investigated, our results are generally similar to those of other studies.<sup>9,18,20</sup> However, the MARDs calculated in our study are higher than the MARDs for the same devices in some previous studies,<sup>22,23</sup> which might be due to the inclusion in our study of more samples in the hypoglycemic range and of samples with a wide range of hemoglobins, as well as the use of manipulated samples.

The accuracy of blood glucose monitoring systems depends on many factors, including the strip enzyme, the manufacturing consistency of the strips, the algorithms used to produce PG results, temperature, humidity, altitude, interfering substances, sample source, collection method, and hematocrit level.<sup>24,25</sup> We chose to assess accuracy in a laboratory setting, which eliminated sources of variability such as

temperature humidity and altitude, and which allowed us to present the same sample to each of 17 meters at the same time while ensuring that we always presented an adequate sample to each meter. Furthermore, it allowed us to test the meters with samples that had reference PG values spread evenly over the entire PG range of the meters. Our study showed glucose meters performed best at normal PG concentrations, and less well in the hypoglycemic range, consistent with the findings of several previous studies.<sup>26-28</sup>

The conditions of our study were more challenging than point-of-care glucose meters are likely to encounter in typical outpatient settings because we included many hypoglycemic samples and samples with a wide range of hematocrit values. We evaluate whether the meter meters met the

**Table 3.** Accuracy Based on ISO Numerical Criteria.

Glucose meter	Met ISO 2003 criteria	ISO 2003 (%)	95% CI (%)	Met ISO 2013 criteria	ISO 2013 (%)	95% CI (%)
Contour Next	Yes	98.3	[96.2, 99.2]	Yes	96.8	[94.3, 98.2]
StatStrip Xpress	Yes	98.8	[97.0, 99.6]	Yes	96.5	[94.0, 98.0]
OneTouch VerioIQ	Yes	97.7	[95.4, 98.8]	No	91.9	[88.5, 94.4]
Accu-Chek Nano	Yes	97.1	[94.7, 98.4]	No	91.0	[87.5, 93.6]
FreeStyle Freedom Lite	Yes	97.1	[94.7, 98.4]	No	92.2	[88.8, 94.6]
Accu-Chek Aviva Plus	Yes	96.0	[93.3, 97.6]	No	91.0	[87.5, 93.6]
FreeStyle Lite	Yes	98.0	[95.8, 99.0]	No	91.9	[88.5, 94.4]
Nova Max	No	92.5	[89.2, 94.8]	No	88.2	[84.3, 91.2]
TRUEresult	No	82.4	[78.0, 86.0]	No	68.2	[63.1, 72.9]
HemoCue Glucose 201	No	85.3	[81.1, 88.6]	No	67.1	[61.9, 71.8]
OneTouch Ultra2	No	80.6	[76.1, 84.5]	No	70.2	[65.2, 74.8]
ReliOn Prime	No	75.4	[70.6, 79.7]	No	62.1	[56.9, 67.1]
BREEZE ®2	No	74.0	[69.1, 78.4]	No	62.4	[57.2, 67.4]
ReliOn Micro	No	72.3	[67.3, 76.7]	No	61.9	[56.6, 66.8]
AgaMatrix PRESTO	No	74.6	[69.7, 78.9]	No	63.3	[58.1, 68.2]
AgaMatrix JAZZ	No	72.5	[67.6, 77.0]	No	60.4	[55.1, 65.5]
SideKick	No	58.4	[53.1, 63.5]	No	49.1	[43.9, 54.4]

Meters are listed in order of overall increasing MARD. Samples were not collected according to the ISO methods, so failure to meet the criteria in this study does not mean that the meters are not compliant.

**Table 4.** Accuracy Comparison in Specimens With Low and Normal Hemoglobin

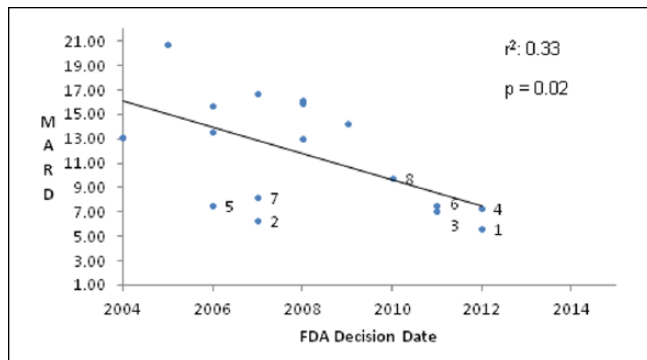
Glucose meter	Hb ≤ 12		Hb > 12		P value	Difference in MARDs (%) <sup>a</sup>
	MARD	SD	MARD	SD		
Contour Next	5.5	7.1	5.9	5.5	.6	-6.4
StatStrip Xpress	5.8	5.4	7.0	6.8	.06	-17.7
OneTouch VerioIQ	6.3	7.1	8.0	6.4	.03	-20.8
Accu-Chek Nano	8.3	7.8	6.1	6.0	.004	36.5
FreeStyle Freedom Lite	6.2	5.0	9.3	7.4	.001	-33.4
Accu-Chek Aviva Plus	8.4	8.4	6.6	7.1	.04	27.5
FreeStyle Lite	6.3	5.5	10.8	9.9	<.001	-42.0
Nova Max	8.9	13.5	10.8	11.4	.2	-17.5
TRUEresult	13.8	8.7	12.1	8.0	.06	14.4
HemoCue Glucose 201	15.7	10.7	10.0	7.7	<.001	57.0
OneTouch Ultra2	12.8	11.6	14.7	11.4	.1	-13.0
ReliOn Prime	15.7	11.1	12.5	8.7	.004	25.1
BREEZE ®2	17.9	13.4	13.0	10.4	<.001	38.1
ReliOn Micro	17.2	12.3	14.3	11.8	.03	20.1
AgaMatrix PRESTO	14.4	12.9	18.4	14.5	.007	-21.5
AgaMatrix JAZZ	15.0	13.0	19.0	14.7	.008	-21.1
SideKick	22.9	17.8	18.0	14.5	.007	27.0

Meters are listed in order of overall increasing MARD.

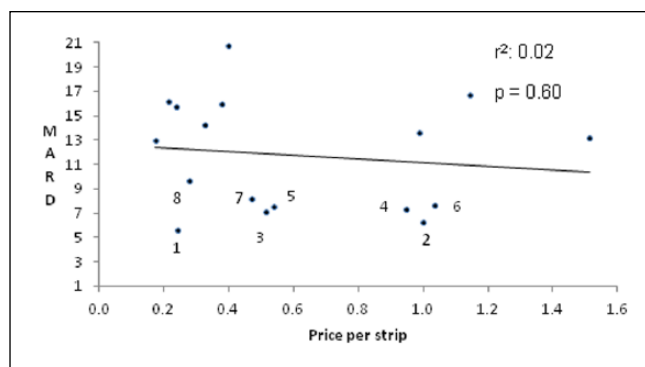
<sup>a</sup>MARD in Hb ≤ 12 - MARD in Hb > 12 / MARD in Hb > 12% × 100.

accuracy criteria stipulated by the ISO 15197:2003 and ISO 15197:2013 in the setting of our study. There were several differences between our testing procedure and that specified by ISO 15197, which specifies testing fresh capillary blood samples, duplicating the reference BG measurements, testing 2 meters for each system, and using only samples with

hematocrit values within allowed range specified by the manufacturer of the meter. In the context, of our study only 7 glucose meters met the numerical accuracy criteria for ISO 2003, and only 2 of those met the more stringent criteria of ISO 2013. Because our testing protocol was different from that specified by the ISO standards our results do not mean



**Figure 3A.** Relationship between FDA decision date and overall MARD.



**Figure 3B.** Relationship between MARD and retail strip price. Meters with MARDs <10: 1. Contour Next; 2. Nova Xpress; 3. OneTouch VerioIQ; 4. Accu-Chek Nano; 5. FreeStyle Freedom Lite; 6. Accu-Chek Aviva Plus; 7. FreeStyle Lite; 8. Nova Nova Max.

that these meters wouldn't meet the criteria when tested according to the ISO protocol. We found that some of the previously tested meters performed more poorly in our test than in previous study, consistent with our conditions being more challenging.<sup>12,13,15</sup> It is likely that some of the meters that had higher MARDs in this study would perform better under actual use conditions. Nonetheless, this study provides information on the relative robustness of the meters to challenging conditions, and it would be reasonable to choose meters that had higher accuracy over those with lower accuracy in the context of this study.

Chronic diseases (eg, kidney disease), high altitude, medications and life style modifications (eg, smoking and exercise) can alter hematocrit,<sup>29</sup> and anemia is common in patients with diabetes.<sup>30</sup> Abnormally low hematocrit therefore has the potential to be a source of measurement bias for some patients, and should be addressed in testing. Consistent with our results, previous studies have shown that some devices are more affected by hematocrit levels than others, and that low hematocrit values (<35%) resulted in overestimates of laboratory glucose levels for some devices.<sup>22,31-33</sup> In our study, we

measured whole blood glucose using the YSI and reported PG using a conversion based on hematocrit. Point-of-care meters that do not have a way of measuring or estimating hemoglobin must apply a fixed mathematical offset to report a "plasma calibrated" result that assumes a normal hematocrit level. This is not expected to be as accurate as a correction based on measured hematocrit or hemoglobin. Therefore, it is somewhat surprising that only 8 of the 17 meters had a higher MARD when tested with anemic blood samples compared to those tested with samples containing normal hemoglobin. This may be related to improved meter technology.<sup>22,31,32,34,35</sup> All 7 meters that met the numerical criteria for ISO 2003 in the context of this study, and had the lowest MARDs, utilize GDH as the strip enzyme. Systems that utilize GDH are less affected by common interfering conditions such as hypoxia<sup>24,36</sup> and by nonglucose sugars. The correlation between FDA approval dates suggests that new technologies have improved the potential for meter accuracy.

Our results showed that glucose meter accuracy was not significantly correlated with retail strip cost. A limitation of our analysis is that costs to the payer are typically lower than the retail pricing, and we did not have access to these data. The market share data from Nielsen may not be fully representative, but based on the Nielsen data the meters we tested account for more than 90% of the market. Although some strips that performed well were moderately priced, many strips with low retail cost performed very poorly in our study, so cost should probably not be the sole criterion for insurers when choosing which meters to cover in their benefit plans.

Our study has several limitations. This study was performed in a laboratory by trained research staff, the test protocol assessed technical accuracy and not clinical accuracy, and the mix of samples was likely more challenging than would be encountered in most clinical situations. Therefore, our results may not be representative of the performance of these systems in the hands of patients with diabetes. We excluded several sources of error including physical errors, for example altitude, humidity, and temperature, and patient related errors including contaminated fingers.<sup>37,38</sup> The study was not performed according to the protocol suggested by the ISO 15197 guideline, as our primary endpoint was MARD with the YSI 2300 as the reference, and the samples were not collected in accordance with the ISO guidelines. We used the YSI as the reference method because most glucose meters in the United States are calibrated using this glucose oxidase based method. The use of different reference methods (oxidase-based vs hexokinase-based) can lead to differences in the apparent accuracy of glucose meters.<sup>39,40</sup> We used a conversion, based on hemoglobin, to convert reference glucose measurements on whole blood to plasma equivalents instead of direct measurement of PG. Since not all the device packages included quality control solutions and most of meters do not recommend regular testing with control solution, we did not perform control solution testing.

We used discarded blood samples that were deidentified. We did not have information on possible medications and other patient conditions that could interfere with blood glucose measurements. Since many samples came from hospitalized patients the array of possible interfering substances may have been larger than would be encountered in the outpatient setting. We added glucose or incubated blood samples to produce samples with very high and very low glucose levels, respectively. A previous study reported that correlation with the reference method was adversely impacted by inclusion of samples that were spiked.<sup>41</sup> However, since all meters tested the same samples any such effects should have impacted all of the meters equally. Meters using glucose oxidase chemistry were concentrated in the bottom half of our accuracy rankings. Systems using glucose oxidase can be affected by hypoxia and hyperoxia.<sup>42</sup> Samples had access to room air during incubation and were mixed by inversion prior to each measurement, thereby oxygenating the blood and ensuring that our samples were not hypoxic. In fact, samples may have been hyperoxic relative to capillary samples, which would be expected to lower the apparent plasma glucose.<sup>42</sup> However, there was no consistent pattern of a negative bias of meter measurements among the meters using glucose oxidase chemistry, and our reference method, the YSI, uses glucose oxidase chemistry. Together, these considerations make it unlikely that hyperoxia can account for the relatively poor performance of the meters using glucose oxidase chemistry in this study.

## Conclusion

In this study the accuracy of commercially available glucose meters varied widely. These data on the relative accuracy of point-of-care glucose meters and their robustness to challenging samples may be useful in informing the choice of a glucose meter.

## Abbreviations

ANOVA, analysis of variance; BG, blood glucose; FAD, flavin adenine dinucleotide; FDA, Food and Drug Administration; GDH, glucose dehydrogenase; GOx, glucose oxidase; Hb, hemoglobin; ISO, International Organization for Standardization; MARD, mean absolute relative difference; Mut Q, mutant variant of quinoprotein; PG, plasma glucose; PL, private label; PQQ, pyrroloquinoline-quinone; RD, relative difference; SD, standard deviation.

## Acknowledgment

We thank Hui Zheng, PhD for statistical advice.

## Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: After the results of the study were obtained, SJR requested and received a donation of glucose meters and strips from Nova Biomedical for several studies of automated glucose control. The

Nova Biomedical meters and strips used for this study were not donated, but were purchased at retail from a distributor. SJR has received support-in-kind (loaned equipment) for another study from Abbott Diabetes Care. None of the other authors have conflicts to report.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Supported by grant R01DK097657 from the National Institute of Diabetes and Digestive and Kidney Diseases to SJR and training grant 2T32DK007028-41.

## Supplemental Material

The supplementary material is available at <http://dst.sagepub.com/supplemental>

## References

1. 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 mellitus. *N Engl J Med.* 1993;329:977-986.
2. Nathan DM, Cleary PA, Backlund JY, et al. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med.* 2005;353:2643-2653.
3. Kiechle F. Blood glucose: measurement in the point-of-care setting. *Laboratory Medicine.* 2000;31:282.
4. Gandhi GY, Kovalaske M, Kudva Y, et al. Efficacy of continuous glucose monitoring in improving glycemic control and reducing hypoglycemia: a systematic review and meta-analysis of randomized trials. *J Diabetes Sci Technol.* 2011;5:952-965.
5. Hoeks LB, Greven WL, de Valk HW. Real-time continuous glucose monitoring system for treatment of diabetes: a systematic review. *Diabet Med.* 2011;28:386-394.
6. Centers for Disease Control and Prevention. *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014.* Atlanta, GA: US Department of Health and Human Services; 2014.
7. Hughes MD. The business of self-monitoring of blood glucose: a market profile. *J Diabetes Sci Technol.* 2009;3:1219-1223.
8. IFCC Scientific Division Working Group on Selective Electrodes. Recommendation on reporting results for blood glucose (from an IFCC stage 1 document). 2001. Available at: <http://www.ifcc.org/ejifcc/vol12no4/vol12no4a4.htm>.
9. Sutherland HLRT. Technical and clinical accuracy of five blood glucose meters: clinical impact assessment using error grid analysis and insulin sliding scales. *J Clin Pathol.* 2015;68:e3.
10. International Organization for Standardization. In vitro diagnostic test systems—requirements for blood-glucose monitoring systems for self-testing in managing diabetes Mellitus. EN ISO 15197:2003 (E).
11. International Organization for Standardization. In vitro diagnostic test systems—requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus. ISO 15197:2013 (E).
12. Freckmann G, Baumstark A, Jendrike N, et al. System accuracy evaluation of 27 blood glucose monitoring systems according to DIN EN ISO 15197. *Diabetes Technol Ther.* 2010;12:221-231.



13. Freckmann G, Schmid C, Baumstark A, Pleus S, Link M, Haug C. System accuracy evaluation of 43 blood glucose monitoring systems for self-monitoring of blood glucose according to DIN EN ISO 15197. *J Diabetes Sci Technol*. 2012;6:1060-1075.
14. Baumstark A, Pleus S, Schmid C, Link M, Haug C, Freckmann G. Lot-to-lot variability of test strips and accuracy assessment of systems for self-monitoring of blood glucose according to ISO 15197. *J Diabetes Sci Technol*. 2012;6:1076-1086.
15. Freckmann G, Baumstark A, Schmid C, Pleus S, Link M, Haug C. Evaluation of 12 blood glucose monitoring systems for self-testing: system accuracy and measurement reproducibility. *Diabetes Technol Ther*. 2014;16:113-122.
16. Pfitzner A, Hengesbach C, Demircik F, Schipper C, Forst T, Musholt PB. Performance of blood glucose meters in compliance with current and future clinical ISO15197 accuracy criteria. *Curr Med Res Opin*. 2014;30:185-190.
17. Brazg RL, Klaff LJ, Parkin CG. Performance variability of seven commonly used self-monitoring of blood glucose systems: clinical considerations for patients and providers. *J Diabetes Sci Technol*. 2013;7:144-152.
18. DeSalvo DJ, Shanmugham S, Ly TT, Wilson DM, Buckingham BA. Accuracy evaluation of blood glucose monitoring systems in children on overnight closed-loop control. *J Diabetes Sci Technol*. 2014;8:969-973.
19. Halldorsdottir S, Warchal-Windham ME, Wallace JF, Pardo S, Parkes JL, Simmons DA. Accuracy evaluation of five blood glucose monitoring systems: the North American comparator trial. *J Diabetes Sci Technol*. 2013;7:1294-1304.
20. Klaff LJ, Brazg R, Hughes K, et al. Accuracy evaluation of Contour Next compared with five blood glucose monitoring systems across a wide range of blood glucose concentrations occurring in a clinical research setting. *Diabetes Technol Ther*. 2015;17:8-15.
21. Rivers SM, Kane MP, Bakst G, Busch RS, Hamilton RA. Precision and accuracy of two blood glucose meters: FreeStyle Flash versus One Touch Ultra. *Am J Health Syst Pharm*. 2006;63:1411-1416.
22. Pfitzner A, Schipper C, Ramljak S, et al. Determination of hematocrit interference in blood samples derived from patients with different blood glucose concentrations. *J Diabetes Sci Technol*. 2013;7:170-178.
23. Tack C, Pohlmeier H, Behnke T, et al. Accuracy evaluation of five blood glucose monitoring systems obtained from the pharmacy: a European multicenter study with 453 subjects. *Diabetes Technol Ther*. 2012;14:330-337.
24. Rebel A, Rice MA, Fahy BG. Accuracy of point-of-care glucose measurements. *J Diabetes Sci Technol*. 2012;6:396-411.
25. Tonyushkina K, Nichols JH. Glucose meters: a review of technical challenges to obtaining accurate results. *J Diabetes Sci Technol*. 2009;3:971-980.
26. Francescato MP, Geat M, Stel G, Cauci S. Accuracy of a portable glucose meter and of a continuous glucose monitoring device used at home by patients with type 1 diabetes. *Clin Chim Acta*. 2012;413:312-318.
27. Freckmann G, Pleus S, Link M, et al. Accuracy evaluation of four blood glucose monitoring systems in unaltered blood samples in the low glycemic range and blood samples in the concentration range defined by ISO 15197. *Diabetes Technol Ther*. 2015;17:625-634.
28. Sonmez A, Yilmaz Z, Uckaya G, et al. The accuracy of home glucose meters in hypoglycemia. *Diabetes Technol Ther*. 2010;12:619-626.
29. Takubo T, Tatsumi N. [Reference values for hematologic laboratory tests and hematologic disorders in the aged]. *Rinsho Byori*. 2000;48:207-216.
30. Thomas MC, MacIsaac RJ, Tsalamandris C, Power D, Jerums G. Unrecognized anemia in patients with diabetes: a cross-sectional survey. *Diabetes Care*. 2003;26:1164-1169.
31. Pfitzner A, Musholt PB, Schipper C, et al. Blood glucose meters employing dynamic electrochemistry are stable against hematocrit interference in a laboratory setting. *J Diabetes Sci Technol*. 2013;7:1530-1537.
32. Ramljak S, Lock JP, Schipper C, et al. Hematocrit interference of blood glucose meters for patient self-measurement. *J Diabetes Sci Technol*. 2013;7:179-189.
33. Tang Z, Lee JH, Louie RF, Kost GJ. Effects of different hematocrit levels on glucose measurements with handheld meters for point-of-care testing. *Arch Pathol Lab Med*. 2000;124:1135-1140.
34. Demircik F, Ramljak S, Hermanns I, Pfitzner A, Pfitzner A. Evaluation of hematocrit interference with MyStar extra and seven competitive devices. *J Diabetes Sci Technol*. 2015;9:262-267.
35. Musholt PB, Schipper C, Thome N, et al. Dynamic electrochemistry corrects for hematocrit interference on blood glucose determinations with patient self-measurement devices. *J Diabetes Sci Technol*. 2011;5:1167-1175.
36. Huang TY, Chang HW, Tsao MF, et al. Evaluation of accuracy of FAD-GDH- and mutant Q-GDH-based blood glucose monitors in multi-patient populations. *Clin Chim Acta*. 2014;433:28-33.
37. Ginsberg BH. Factors affecting blood glucose monitoring: sources of errors in measurement. *J Diabetes Sci Technol*. 2009;3:903-913.
38. Hortensius J, Slingerland RJ, Kleefstra N, et al. Self-monitoring of blood glucose: the use of the first or the second drop of blood. *Diabetes Care*. 2011;34:556-560.
39. Link M, Schmid C, Pleus S, et al. System accuracy evaluation of four systems for self-monitoring of blood glucose following ISO 15197 using a glucose oxidase and a hexokinase-based comparison method. *J Diabetes Sci Technol*. 2015;9:1041-1050.
40. Twomey PJ. Plasma glucose measurement with the Yellow Springs Glucose 2300 STAT and the Olympus AU640. *J Clin Pathol*. 2004;57:752-754.
41. Karon BS, Griesmann L, Scott R, et al. Evaluation of the impact of hematocrit and other interference on the accuracy of hospital-based glucose meters. *Diabetes Technol Ther*. 2008;10:111-120.
42. Tang Z, Louie RF, Lee JH, Lee DM, Miller EE, Kost GJ. Oxygen effects on glucose meter measurements with glucose dehydrogenase- and oxidase-based test strips for point-of-care testing. *Crit Care Med*. 2001;29:1062-1070.