Accuracy Evaluation of Five Blood Glucose Monitoring Systems Obtained from the Pharmacy: A European Multicenter Study with 453 Subjects

Cornelius Tack, M.D.,¹ Harald Pohlmeier, M.D.,² Thomas Behnke, M.D.,³ Volkmar Schmid, Ph.D.,⁴
Marco Grenningloh, Ph.D.,⁴ Thomas Forst, M.D.,⁵ and Andreas Pfützner, M.D., Ph.D.⁵

Abstract

Background: This multicenter study was conducted to evaluate the performance of five recently introduced blood glucose (BG) monitoring (BGM) devices under daily routine conditions in comparison with the YSI (Yellow Springs, OH) 2300 Stat Plus glucose analyzer.

Methods: Five hundred one diabetes patients with experience in self-monitoring of BG were randomized to use three of five different BGM devices (FreeStyle Lite[®] [Abbott Diabetes Care Inc., Alameda, CA], FreeStyle Freedom Lite [Abbott Diabetes Care], OneTouch[®] UltraEasy[®] [LifeScan Inc., Milpitas, CA], Accu-Chek[®] Aviva [Roche Diagnostics, Mannheim, Germany], and Contour[®] [Bayer Vital GmbH, Leverkusen, Germany]) in a daily routine setting. All devices and strips were purchased from local regular distribution sources (pharmacies, four strip lots per device). The patients performed the finger prick and the glucose measurement on their own. In parallel, a healthcare professional performed the glucose assessment with the reference method (YSI 2300 Stat Plus). The primary objective was the comparison of the mean absolute relative differences (MARD). Secondary objectives were compliance with the International Organization for Standardization (ISO) accuracy criteria under these routine conditions and Clarke and Parkes Error Grid analyses.

Results: MARD ranged from 4.9% (FreeStyle Lite) to 9.7% (OneTouch UltraEasy). The ISO 15197:2003 requirements were fulfilled by the FreeStyle Lite (98.8%), FreeStyle Freedom Lite (97.5%), and Accu-Chek Aviva (97.0%), but not by the Contour (92.4%) and OneTouch UltraEasy (91.1%). The number of values in Zone A of the Clarke Error Grid analysis was highest for the FreeStyle Lite (98.8%) and lowest for the OneTouch Ultra Easy (90.4%).

Conclusions: FreeStyle Lite, FreeStyle Freedom Lite, and Accu-Chek Aviva performed very well in this study with devices and strips purchased through regular distribution channels, with the FreeStyle Lite achieving the lowest MARD in this investigation.

Introduction

BLOOD GLUCOSE (BG) CONTROL is a critical issue in diabetes treatment, especially when applying intensive insulin regimens, to avoid occurrence of hyperglycemia or hypoglycemia and their respective long-term consequences. The important role of accurate BG measurement in combination with a pandemically increasing number of diabetes cases worldwide¹ has resulted in an increasing number of BG monitoring (BGM) systems, which have been developed and launched globally.

To ensure appropriate function of the devices, standardization tests were introduced for quality assurance purposes, which also allow for comparable testing of BGM devices. Many clinical trials were conducted²⁻⁴ to test compliance with accuracy requirements stipulated by International Organization for Standardization (ISO) 15197:2003.⁵ However, most investigations only addressed some aspects of accuracy and its consequence on treatment and in a special and controlled laboratory setting for regulatory approval. In consequence, information about compliance of BGM systems with the specifications of the new draft of ISO 15197 (released in 2010)⁶

¹Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.

²Zentrum für Diabetes und Gefäßerkrankungen, Münster, Germany.

³Diabetes Zentrum Neuwied, Neuwied, Germany.

⁴IKFE-CRO GmbH, Mainz, Germany.

⁵IKFE, Mainz, Germany.

is lacking, as well as performance data of devices under daily routine conditions.

Therefore, the goal of the present study was to compare the performance of five recently introduced BGM devices (Accu-Chek® Aviva [Roche Diagnostics, Mannheim, Germany], Contour® [Bayer Vital GmbH, Leverkusen, Germany], FreeStyle Freedom Lite® [Abbott Diabetes Care Inc., Alameda, CA], FreeStyle Lite [Abbott Diabetes Care], and OneTouch® UltraEasy® [LifeScan Inc., Milpitas, CA]) directly obtained from the shelf of a German pharmacy with regard to the criteria as set forth in the respective ISO guidelines, when measurements are conducted not in a controlled environment setting but by recently trained patients familiar with the BGM procedures in a clinical setting. Main objectives of the study were the comparison of the mean absolute relative differences (MARDs) and the compliance with the accuracy criteria described in ISO 15197⁵ and the new draft (not yet approved) from 2010.⁶

Materials and Methods

Study design and Good Clinical Practice conformity

This trial was a randomized, prospective, open, multicenter study that was conducted by an independent contract research organization (IKFE-CRO, Mainz, Germany). The trial consisted of one visit and was conducted at two investigational sites in Germany and one study site in The Netherlands. It was approved by the responsible ethics committees in both countries and conducted in accordance with the principles of the Declaration of Helsinki and the applicable guidelines for Good Clinical Practice, whichever provided the greater protection of the individual. Prior to study participation informed consent was obtained from all patients in writing.

BGM systems

The five BGM systems investigated were as follows: (1) FreeStyle Lite, (2) FreeStyle Freedom Lite, (3) OneTouch UltraEasy, (4) Accu-Chek Aviva, and (5) Contour. The YSI (Yellow Springs, OH) 2300 glucose analyzer (glucose oxidase method) was chosen as the reference method because it is accepted by the regulatory agencies and was available at all study sites.

Study population and study conduct

Patients with type 1 or type 2 diabetes mellitus were enrolled into this trial. Further inclusion criteria were as follows: age between 18 to 75 years, stable BG level at the investigator's discretion, intensive insulin therapy, and experience in self-measurement of BG with multiple daily measurements. Exclusion criteria were as follows: unstable glycemic control (investigator's discretion), lack of legal capacity, pregnancy, blood donation within the last 30 days, and reasons precluding successful study participation (lack of compliance, diagnosis of a progressive/serious disease). The primary objective was the comparison of the MARD values, and secondary objectives were compliance with ISO 15197 and accuracy as assessed by error grid analysis under these clinical conditions with some obvious deviation from the strict ISO protocol (only clinical samples; no laboratory manipulation to very high and very low glucose concentrations; measure-

ments done by the patients themselves, not by healthcare professionals).

Upon enrollment, patients were randomized to use three of the five different BGM systems in a predefined order in each study center using permuted block randomization, stratified by clinical site. This randomization scheme ensured similar usage of the five different BGM systems in a random sequence in each study site. The patients came in fasting and washed their hands prior to performing the measurements. Upon instruction to perform the first BGM, patients had to prick themselves at the fingertip with a commercially available lancing device, to measure BG with two strips in duplicate. All BGM devices and strips were purchased in a local pharmacy (four strip lots per device). Prior to and after the self-test, a healthcare professional collected capillary blood from the fingertip for the laboratory reference measurement with the YSI 2300 Stat Plus glucose analyzer (glucose oxidase method). The YSI 2300 whole blood results were converted to plasma equivalent results, and these results were used for comparison with the test strip results. Patients sequentially repeated all aforementioned steps for the second and third BGM. Hematocrit was measured to assure that hematocrit values of blood samples were within the range specifications of the individual BGM device.

Statistical analysis

Patients were excluded from the analysis if the time delay between the self-measurements and the performance of the YSI measurements exceeded more than 60 min. MARD between the values obtained by an individual BGM versus the YSI reference was calculated for the different BGM systems. MARD is a continuous measure that uses the magnitude of the relative difference for each observation. An analysis of variance was conducted, and the statistical significance of differences between devices was analyzed by unpaired Student's *t* tests. To account for the multiple comparisons, the significance level was adjusted to 0.0051 for each individual difference.

Consistency of the measurements of individual BGM was analyzed by regression analysis. To obtain slopes and intercepts, regression analyses according to Passing and Bablok⁷ were conducted. The correlation coefficient (*r*) was calculated using the standard regression method of Pearson. To obtain the mean absolute percentage residual, the regression line was used to calculate a predicted value for an individual BGM system measurement from the corresponding YSI 2300 reference value. The difference between the actually measured value and the predicted value yields the absolute percentage residual.

Although the study was not conducted in strict compliance with the ISO 15197 guidelines, the number and percentage of measurements within ± 15 mg/dL (for glucose concentrations < 75 mg/dL) and $\pm 20\%$ (for glucose concentrations ≥ 75 mg/dL) were evaluated in order to examine compliance with accuracy criteria of ISO 15197:2003 (95% of the measurements have to be in the acceptance range).⁵ Accuracy compliance according to the new draft of the ISO 15197 is defined by percentage of results within ± 15 mg/dL (for glucose concentrations < 100 mg/dL) and within $\pm 15\%$ (for glucose concentrations ≥ 100 mg/dL) and was also assessed and interpreted (acceptance criteria, 95% within these predefined ranges).⁶ All

statistical analyses were performed using SAS[®] version 9.2 for Windows (SAS Institute, Cary, NC). Tests for the clinical significance of deviating BGM system measurements were performed by means of error grid analyses according to Clarke et al.⁸ and Parkes et al.⁹

Results

In total, 501 subjects were enrolled and participated in this study. Forty-eight subjects were excluded from the final analysis of the BGM results for the following reasons: (1) deviation between YSI 2300 duplicate measurements of more than 4% for glucose ≥ 100 mg/dL or 4 mg/dL for < 100 mg/dL, (2) delay of YSI measurements versus BGM system measurements by more than 1 h, (3) hematocrit values outside the range of the system specifications of the corresponding BGM device, (4) results outside the dynamic range of the BGM device, and (5) failure to meet all inclusion/exclusion criteria. Although reason 2 was the major cause of exclusion, 82% of the measurements were completed within 15 min (92.6% within 30 min). Data from the remaining 453 patients were included in the statistical analysis.

Demographics and baseline characteristics

The number of patients was comparable for the three study sites—Münster, Germany ($n=160$; 35.3%), Neuwied, Germany ($n=144$; 31.8%), and Nijmegen, the Netherlands ($n=149$; 32.9%)—with an equal distribution of type 1 (49.4%) and type 2 (50.6%) patients (98.2% white, 57.6% male, age 53 ± 15 years old [mean $\pm$ SD], glycated hemoglobin $7.8 \pm 1.35\%$ [mean $\pm$ SD]). The level of education of most patients was secondary school (60.2%), followed by university (16.4%), trade school (13.7%), and primary school (8.4%). Prior to this study, most patients (248 of 453; 54.7%) routinely conducted four daily BG measurements. Substantial numbers of patients were conducting five (67 of 453; 14.8%), three (58 of 453; 12.8%), and six (38 of 453; 8.4%) daily measurements, whereas only a few patients measured more frequently (seven times a day [23 of 453; 5.1%]) or less frequently (twice a day [18 of 453; 4.0%] or once a day [one of 453; 0.2%]). Randomization resulted in comparable use of the different devices (FreeStyle Lite, 240 times; FreeStyle Freedom Lite, 244 times; Accu-Chek Aviva, 252 times; Contour, 255 times; and OneTouch UltraEasy, 246 times).

Bias and regression analysis

Mean percentage bias was lowest for the Contour (-0.2%) and the two FreeStyle devices (1% each) and highest for the

Accu-Chek Aviva (5.3%) and OneTouch UltraEasy (4.6%; Table 1). The MARD values shown in Table 1, however, demonstrate that the lowest MARD was observed for the FreeStyle devices (FreeStyle Lite, 4.9%; FreeStyle Freedom Lite, 5.5%), and the highest MARD was seen with Contour (9.0%, $P < 0.0005$ vs. both FreeStyle devices) and OneTouch UltraEasy (9.7%, $P < 0.0005$). The 95% confidence intervals (CIs) and SDs increased from the FreeStyle Lite (95% CI, 4.1 to 5.7; SD, $\pm 7.5\%$) to the OneTouch UltraEasy (95% CI, 8.9 to 10.4; SD, ± 12.1) (Table 2). MARD of the Accu-Chek Aviva was also significantly lower than that of the Contour and OneTouch UltraEasy.

Regression analyses were conducted for all BGM systems versus the YSI 2300 reference method. The parameters shown in Table 3 demonstrate very good correlations of all BGM devices with the YSI 2300 measurements.

Accuracy according to ISO 15197:2003 and its new draft of 2010

This study was not done in strict compliance with ISO 15197, which requires a specific distribution of the glucose values in predefined ranges. Also, no artificial manipulation of samples was done to achieve very high or very low glucose levels. However, to study accuracy of BG measurements in this routine clinical setting according to ISO 15197:2003, BGM readings were plotted against the results of the YSI reference measurements as shown in Figure 1. Table 4 shows the number of readings within different ranges of deviation for each BGM device (< 5 , < 10 , and < 15 mg/dL). All measurements < 75 mg/dL with FreeStyle Lite were within ± 15 mg/dL; by contrast, only 83% of the measurements with OneTouch UltraEasy were within this acceptance range. Accuracy of BG readings ≥ 75 mg/dL was analyzed for four ranges (within $\pm 5\%$, $\pm 10\%$, $\pm 15\%$, and $\pm 20\%$). FreeStyle Lite performed best with the highest percentage of values within all four categories (Table 4); the BGM system with the lowest number of readings within these four categories was the OneTouch UltraEasy. Combining the data obtained for ± 15 mg/dL and $\pm 20\%$ yields the values required for assessment of the ISO 15197:2003 requirements. It is evident from Table 4 that both FreeStyle devices as well as the Accu-Chek Aviva complied with the ISO requirements, in contrast to the Contour and OneTouch UltraEasy, which yielded less than 95% values within the acceptance ranges.

Evaluation of the data according to the new draft of the ISO 15197⁶ resulted in comparable results (Table 5) with one major exception: only the combined values (within $\pm 15\%$ and ± 15 mg/dL) of the FreeStyle Lite and FreeStyle Freedom Lite

TABLE 1. MEAN ABSOLUTE DIFFERENCE AND MEAN ABSOLUTE RELATIVE DIFFERENCE BETWEEN MEASUREMENTS WITH THE DIFFERENT BLOOD GLUCOSE METERS AND THE YSI 2300

Blood glucose meter	Mean bias		MARD			Subjects (n)	Measurements (n)
	mg/dL	%	%	95% CI	SD (%)		
FreeStyle Lite FreeStyle	0.8	1.0	4.9	4.1 to 5.7	7.5	240	480
Freedom Lite	0.3	1.0	5.5	4.8 to 6.3	8.5	244	488
Accu-Chek Aviva	7.2	5.3	6.8	6.1 to 7.6	7.5	252	504
Contour	-1.2	-0.2	9.0	8.3 to 9.8	12.0	255	510
OneTouch UltraEasy	6.3	4.6	9.7	8.9 to 10.4	12.1	246	492

CI, confidence interval; MARD, mean absolute relative difference; n, number.

TABLE 2. ANALYSIS OF VARIANCE OF THE DIFFERENCE IN MEAN ABSOLUTE BIAS ON A PERCENTAGE BASIS

Blood glucose meter comparison	Difference in mean absolute bias (%)	Confidence interval	P value
FreeStyle Lite vs. FreeStyle Freedom Lite	-0.7	(-2.2, 0.9)	0.2470
FreeStyle Lite vs. Accu-Chek Aviva	-2.0	(-3.5, -0.4)	0.0005*
FreeStyle Lite vs. Contour	-4.2	(-5.7, -2.6)	<0.0001*
FreeStyle Lite vs. OneTouch UltraEasy	-4.8	(-6.4, -3.2)	<0.0001*
FreeStyle Freedom Lite vs. Accu-Chek Aviva	-1.3	(-2.9, 0.3)	0.0191
FreeStyle Freedom Lite vs. Contour	-3.5	(-5.1, -1.9)	<0.0001*
FreeStyle Freedom Lite vs. OneTouch UltraEasy	-4.1	(-5.7, -2.5)	<0.0001*
Accu-Chek Aviva vs. Contour	-2.2	(-3.7, -0.6)	<0.0001*
Accu-Chek Aviva vs. OneTouch UltraEasy	-2.8	(-4.4, -1.2)	<0.0001*
Contour vs. OneTouch UltraEasy	-0.6	(-2.2, 0.9)	0.2658

*Statistically significant difference.

achieved the ISO requirements of $\geq 95\%$ of the measured values within the acceptable range in comparison with the reference device.

Error grid analyses

Results of the Clarke and Parkes Error Grid analysis are graphically displayed for the FreeStyle Lite in Figure 2. Table 6 shows a numerical comparison of these two approaches for the individual BGM systems. Using the Clarke Error Grid analysis, the frequency of values falling into Zone A was highest for the FreeStyle Lite (98.8%) and lowest for the OneTouch UltraEasy (90.4%). The FreeStyle Lite yielded only values in Zones A and B, whereas the other four BGM systems produced few values in Zone D (OneTouch UltraEasy, 1.6%; FreeStyle Freedom Lite, 1.2%; Accu-Chek Aviva, 0.4%; and Contour, 0.4%). Similar results were obtained by means of the Parkes Error Grid analysis.

Adverse events

Four adverse events occurred during this study. All were mild events of hypoglycemia, which were not related to the study devices.

Discussion and Conclusion

The study was designed to minimize selection bias by purchasing the BGM systems directly from the pharmacy. Four lots of test strips were tested with each BGM system in the study to ensure representative performance would be obtained. To evaluate actual patient use, all BGM system tests

were done by the diabetes patients, the intended users, and not by trained healthcare professionals. Subjects participating in this study were between 18 and 75 years, had different levels of education, and conducted one to seven or more daily BG measurements and thus are a representative group of potential users.

MARD is a continuous measure that makes use of the magnitude of the percentage bias of each observation. Therefore MARD is correlated with the ISO accuracy measurement. Dichotomization involved in the calculation of percentage within can be useful to understand clinical impact (e.g., Consensus Error Grid analysis), but also leads to loss of information. As performance improves, a single MARD measure is sensitive to differences in accuracy and precision; percentage within becomes insensitive to changes in accuracy and precision. Therefore, additional calculations with tighter criteria are required. Evaluation of the average difference between individual BGM systems and the YSI 2300 reference yielded results that correspond to those obtained regarding analysis of the ISO 15197 compliance, as is reflected by the lowest MARD values of the FreeStyle Lite and FreeStyle Freedom Lite. The highest MARD values by the Contour and OneTouch UltraEasy are in line with the results from analysis of ISO compliance. The only device that achieved 100% of values within Zones A and B using both types of error grid analyses when operated by end users was the FreeStyle Lite. Mean absolute percentage residual was also lowest for the FreeStyle Lite.

Under these daily routine conditions and without investigating very high or very low values, the accuracy specifications defined by ISO 15197:2003 were met by three BGM

TABLE 3. REGRESSION ANALYSIS OF THE VALUES OF THE INDICATED BLOOD GLUCOSE METERS VERSUS THE YSI 2300 REFERENCE METHOD

Blood glucose meter	Slope	Intercept	95% CI		r	Mean absolute % residual	Mean paired % CV	Subjects (n)	Measurements (n)
			Slope	Intercept					
FreeStyle Lite	0.98	2.5	0.97 to 0.99	0.6 to 4.2	0.985	4.7	2.5	240	480
FreeStyle Freedom Lite	0.98	3.6	0.97 to 0.99	1.3 to 5.6	0.985	5.3	3.1	244	488
Accu-Chek Aviva	1.01	4.8	1.00 to 1.03	2.9 to 6.5	0.989	5.1	2.9	252	504
Contour	0.98	0.6	0.96 to 1.01	-2.2 to 3.7	0.966	9.1	5.1	255	510
OneTouch UltraEasy	1.04	0.5	1.01 to 1.06	-2.8 to 3.6	0.974	8.7	4.5	246	492

CI, confidence interval; CV, coefficient of variation; r, correlation coefficient.

FIG. 1. Measurement bias of capillary blood glucose obtained with hand-held blood glucose monitoring systems versus the laboratory reference device YSI 2300: (A) FreeStyle Lite, (B) FreeStyle Freedom Lite, (C) Accu-Chek Aviva, (D) Contour, and (E) OneTouch UltraEasy. The solid lines represent ± 15 mg/dL for glucose concentrations < 75 mg/dL and $\pm 20\%$ for glucose concentrations ≥ 75 mg/dL.

TABLE 4. ACCURACY AND ISO 15197:2003 COMPLIANCE OF THE INDIVIDUAL BLOOD GLUCOSE METERS
VERSUS THE YSI REFERENCE DEVICE

Blood glucose meter	Compliance with ISO accuracy requirements	Both blood glucose levels within				Blood glucose <75 mg/dL within			Blood glucose ≥75 mg/dL within			
		±5%/ 5 mg/dL	±10%/ 10 mg/dL	±15%/ 15 mg/dL	±20%/ 15 mg/dL	±5 mg/dL	±10 mg/dL	±15 mg/dL	±5%	±10%	±15%	±20%
FreeStyle Lite	Yes	325/480 (68%)	446/480 (93%)	469/480 (98%)	474/480 (99%)	30/48 (63%)	45/48 (94%)	48/48 (100%)	295/432 (68%)	401/432 (93%)	421/432 (97%)	426/432 (99%)
FreeStyle Freedom Lite	Yes	303/488 (62%)	437/488 (90%)	465/488 (95%)	476/488 (98%)	30/52 (58%)	45/52 (87%)	47/52 (90%)	273/436 (63%)	392/436 (90%)	418/436 (96%)	429/436 (98%)
Accu-Chek Aviva	Yes	237/504 (47%)	403/504 (80%)	470/504 (93%)	489/504 (97%)	13/40 (33%)	32/40 (80%)	38/40 (95%)	224/464 (48%)	371/464 (80%)	432/464 (93%)	451/464 (97%)
Contour	No	196/510 (38%)	339/510 (66%)	433/510 (85%)	471/510 (92%)	20/44 (45%)	31/44 (70%)	41/44 (93%)	176/466 (38%)	308/466 (66%)	392/466 (84%)	430/466 (92%)
OneTouch UltraEasy	No	175/492 (36%)	313/492 (64%)	405/492 (82%)	448/492 (91%)	18/42 (43%)	27/42 (64%)	35/42 (83%)	157/450 (35%)	286/450 (64%)	370/450 (82%)	413/450 (92%)

Data are number of within measurements versus the total number of measurements (respective percentage).
ISO, International Organization for Standardization.

TABLE 5. ACCURACY AND ISO 15197 FULFILLMENT OF THE INDIVIDUAL BLOOD GLUCOSE MONITORING SYSTEMS USING THE GLUCOSE THRESHOLD
OF 100 MG/DL REQUIRED ACCORDING TO THE NEW DRAFT OF ISO 15197 OF 2010

Blood glucose meter	Fulfillment of ISO accuracy requirements	All blood glucose levels within				Blood glucose <100 mg/dL within			Blood glucose ≥100 mg/dL within		
		±5%/5 mg/dL	±10%/10 mg/dL	±15%/15 mg/dL	±5 mg/dL	±10 mg/dL	±15 mg/dL	±5 %	±10 %	±15 %	±15 %
FreeStyle Lite	Yes	330/480 (69%)	447/480 (93%)	471/480 (98%)	70/104 (67%)	96/104 (92%)	102/104 (98%)	260/376 (69%)	351/376 (93%)	369/376 (98%)	369/376 (98%)
FreeStyle	Yes	305/488 (63%)	437/488 (90%)	466/488 (95%)	63/104 (61%)	95/104 (91%)	98/104 (94%)	242/384 (63%)	342/384 (89%)	368/384 (96%)	368/384 (96%)
Freedom Lite	No	238/504 (47%)	404/504 (80%)	475/504 (94%)	39/94 (41%)	68/94 (72%)	90/94 (96%)	199/410 (49%)	336/410 (82%)	385/410 (94%)	385/410 (94%)
Accu-Chek Aviva	No	197/510 (39%)	346/510 (68%)	439/510 (86%)	40/120 (33%)	71/120 (59%)	98/120 (82%)	157/390 (40%)	275/390 (71%)	341/390 (87%)	341/390 (87%)
Contour	No	177/492 (36%)	313/492 (64%)	408/492 (83%)	33/94 (35%)	59/94 (63%)	78/94 (83%)	144/398 (36%)	254/398 (64%)	330/398 (83%)	330/398 (83%)
OneTouch UltraEasy	No										

Data are number of within measurements versus total number of measurements (respective percentage).
ISO, International Organization for Standardization.

FIG. 2. Analysis of clinical significance of measuring inaccuracies of individual blood glucose monitoring systems by (A) Clarke and (B) Parkes Error Grid analyses. Results are shown for the FreeStyle Lite device.

devices tested (FreeStyle Lite, FreeStyle Freedom Lite, and Accu-Chek Aviva). These results are in line with former results^{1,3,10,11} showing compliance with ISO 15197:2003 requirements by these devices in a more controlled test environment. Under more controlled conditions, the OneTouch UltraEasy and Contour also fulfilled the specifications,^{2,10} whereas they failed to do so in our study.

Applying the more strict requirements defined by the new draft of ISO 15197 revealed the FreeStyle Lite and FreeStyle Freedom Lite as the only devices complying with these potential new standard specifications even in a more clinical routine setting. Regardless of the version of ISO 15197 applied, the Contour and OneTouch UltraEasy failed the accuracy criteria in this routine study setting.

Our study has several limitations, which need to be considered before drawing conclusions with clinical relevance.

First of all, and as mentioned above, the investigation was not performed in strict compliance with the protocol suggested by the ISO 15197 guideline, as our primary objective was MARD in comparison with the YSI 2300. Artificially prepared very high and very low samples were not tested, as we were interested in the clinical accuracy when measurements were performed by patients. In turn, we collected many more data points than suggested by ISO. Second, the Accu-Chek device is based on a hexokinase method, whereas we used the glucose oxidase-based YSI analyzer as the reference and comparator method. This may have contributed to the differences observed between the FreeStyle devices and Accu-Chek Aviva with respect to MARD, as negative biases of approximately 3–8% have been described previously between YSI and hexokinase-based reference methods.^{12,13} As such, different reference methods might introduce a deviation of the

TABLE 6. STRATIFICATION OF PAIRED VALUES WITHIN THE INDICATED ZONES ACCORDING TO THE CLARKE AND PARKES ERROR GRID ANALYSES

Method, blood glucose meter	Error grid zone										Total n
	A		B		C		D		E		
	n	%	n	%	n	%	n	%	n	%	
Clarke											
FreeStyle Lite	474	98.8	6	1.3	0	0.0	0	0.0	0	0.0	480
FreeStyle Freedom Lite	475	97.3	7	1.4	0	0.0	6	1.2	0	0.0	488
Accu-Chek Aviva	487	96.6	15	3.0	0	0.0	2	0.4	0	0.0	504
Contour	471	92.4	37	7.3	0	0.0	2	0.4	0	0.0	510
OneTouch Ultra Easy	445	90.4	39	7.9	0	0.0	8	1.6	0	0.0	492
Parkes											
FreeStyle Lite	477	99.4	3	0.6	0	0.0	0	0.0	0	0.0	480
FreeStyle Freedom Lite	480	98.4	7	1.4	1	0.2	0	0.0	0	0.0	488
Accu-Chek Aviva	498	98.8	6	1.2	0	0.0	0	0.0	0	0.0	504
Contour	486	95.3	24	4.7	0	0.0	0	0.0	0	0.0	510
OneTouch UltraEasy	463	94.1	29	5.9	0	0.0	0	0.0	0	0.0	492

values that is not caused by the device itself. This effect might have contributed to the relative high mean bias (Table 1) of the Accu-Chek Aviva (hexokinase) in comparison with the YSI (glucose oxidase)-calibrated BG meters and subsequently might influence the interpretation of the results. However, these limitations do not apply to the other comparator meters.

In summary, using several parameters to assess the reliability of BGM devices, the FreeStyle Freedom Lite and especially the FreeStyle Lite were shown to be the most accurate and reliable BG meters in our clinical setting in comparison with the YSI glucose oxidase reference method in the hands of the end user population, with study materials obtained through regular supply channels.

Acknowledgments

This study was funded by Abbott Diabetes Care. Abbott Diabetes Care was involved in the design of the study and performed the statistical analysis. The IKFE in Mainz was responsible for obtaining all study materials, managed the study, and reviewed the statistical analysis. The manuscript was prepared by the authors. Abbott Diabetes Care was permitted to review and comment on the manuscript, but the final decision on the content was exclusively retained by the authors.

Author Disclosure Statement

C.T., T.F., and A.P. have received travel support, speaker fees, and research grants from Abbott Medisense. H.P., T.B., V.S., and M.G. declare no competing financial interests.

References

1. International Diabetes Federation: Diabetes Atlas, 4th ed. Brussels: International Diabetes Federation, 2009.
2. Kristensen GB, Monsen G, Skeie S, Sandberg S: Standardized evaluation of nine instruments for self-monitoring of blood glucose. *Diabetes Technol Ther* 2008;10:467–477.
3. Freckmann G, Baumstark A, Jendrike N, Zschornack E, Kocher S, Tshiananga J, Heister F, Haug C: System accuracy evaluation of 27 blood glucose monitoring systems according to DIN EN ISO 15197. *Diabetes Technol Ther* 2010;12:221–231.
4. Kuo CY, Hsu CT, Ho CS, Su TE, Wu MH, Wang CJ: Accuracy and precision evaluation of seven self-monitoring blood glucose systems. *Diabetes Technol Ther* 2011;13:596–600.
5. ISO 15197:2003: In Vitro Diagnostic Test Systems—Requirements for Blood-Glucose Monitoring Systems for Self-Testing in Managing Diabetes Mellitus. Geneva: World Health Organization, 2003.
6. ISO 15197:2010: In Vitro Diagnostic Test Systems—Requirements for Blood-Glucose Monitoring Systems for Self-Testing in Managing Diabetes Mellitus (ISO/DIS 15197:2010); German version prEN ISO 15197:2010. www.slideshare.net/saccharun/iso-15197-ge (accessed November 2011).
7. Passing H, Bablok W: A new biometrical procedure for testing the equality of measurements from two different analytical methods. Applications of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem* 1983;21:709–720.
8. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 1987;10:622–628.
9. Parkes JL, Slatin SL, Pardo S, Ginsberg BH: A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care* 2000;23:1143–1148.
10. Ng R, Lock J: Can glucose meters meet tighter accuracy requirements? [abstract]. *Diabetologia* 2010;53(Suppl 1):S418.
11. Lock JP, Brazg R, Bernstein RM, Taylor E, Patel M, Ward J, Alva S, Chen T, Welsh Z, Amor W, Bhogal C, Ng R: Performance of a new test strip for FreeStyle blood glucose monitoring systems. *Diabetes Technol Ther* 2011;13:1–10.
12. Twomey PM: Plasma glucose measurement with the Yellow Springs Glucose 2300 STAT and the Olympus AU640. *J Clin Pathol* 2004;57:752–754.
13. Genter P, Ipp E: Accuracy of plasma glucose measurements in the hypoglycemic range. *Diabetes Care* 1994;17:595–598.

Address correspondence to:
 Prof. Dr. Andreas Pfützner, M.D., Ph.D.
 IKFE
 Parcusstrasse 8
 D-55116 Mainz, Germany
 E-mail: AndreasP@ikfe.de