DIABETES TECHNOLOGY & THERAPEUTICS
Volume 12, Number 3, 2010

DOI: 10.1089/dia.2009.0128

System Accuracy Evaluation of 27 Blood Glucose Monitoring Systems According to DIN EN ISO 15197

Guido Freckmann, M.D.,[1] Annette Baumstark, Ph.D.,[1] Nina Jendrike, M.D.,[1] Eva Zschornack, M.D.,[1] Serge Kocher, Ph.D.,[2] Jacques Tshiananga, M.P.H.,[2] Frank Heister, Ph.D.,[2] and Cornelia Haug, M.D.[1]

Abstract

Background: Blood glucose (BG) monitoring systems enable diabetes patients to effectively control and adjust their therapy. BG monitoring systems with a Conformité Européenne (CE) label should meet the standard DIN EN ISO 15197:2003: ≥95% of the BG results shall fall within ±15 mg/dL of the reference method at BG concentrations <75 mg/dL and within ±20% at BG concentrations ≥75 mg/dL. We intended to verify if BG monitoring systems with a CE label fulfill these minimum accuracy requirements.

Methods: We evaluated 27 BG monitoring systems from 18 manufacturers for system accuracy according to DIN EN ISO 15197:2003. Twenty-four systems were compared with the glucose oxidase reaction (YSI 2300 glucose analyzer [YSI Life Sciences, Yellow Springs, OH]) and three systems with the hexokinase reaction (Hitachi 917 [Roche Diagnostics GmbH, Mannheim, Germany]). Duplicate measurements of 100 blood samples with a defined distribution of BG concentrations from 20 mg/dL to 600 mg/dL from ≥100 subjects were included in the evaluation.

Results: Sixteen of the 27 BG monitoring systems fulfilled the minimum accuracy requirements of the standard, i.e., ≥95% of their results showed the minimum acceptable accuracy. Overall, the mean percentage of results showing the minimum acceptable accuracy was 95.2 ± 5.2%, ranging from 80.0% to 100.0%.

Conclusions: More than 40% of the evaluated BG monitoring systems did not fulfill the minimum accuracy requirements of DIN EN ISO 15197:2003. As inaccurate BG monitoring systems bear the risk of false treatment decisions by the diabetes patient and subsequent possible severe health injury, manufacturers should regularly and effectively check the quality of BG meters and BG test strips.

Introduction

Several studies have demonstrated the importance of tight blood glucose control for diabetes patients, e.g., the Diabetes Control and Complications Trial[1] and the United Kingdom Prospective Diabetes Study.[2] Self-monitoring of blood glucose (SMBG) enables diabetes patients to effectively control their blood glucose (BG) levels. The clinical benefits of SMBG in type 1 diabetes patients are widely accepted.[3] In type 2 diabetes patients, clinical, epidemiological, and economic evidence supporting SMBG is accumulating steadily.[4–9] Guidelines from the American Diabetes Association[10] and the International Diabetes Federation[11] recommend the use of SMBG as an integral component of diabetes therapy.

The standard DIN EN ISO 15197:2003[12] specifies requirements for SMBG devices, e.g., with regard to system performance, accuracy, and precision. Additionally, it specifies procedures for the verification and validation of the performance by the intended users, i.e., the diabetes patients. The minimum acceptable system accuracy requirements are based on the medical requirements for glucose monitoring: "≥95% of the individual glucose results shall fall within ±15 mg/dL of the results of the manufacturer's measurement procedure at glucose concentrations <75 mg/dL and within ±20% at glucose concentrations ≥75 mg/dL."[12]

In Europe, manufacturers of SMBG devices have to provide evidence of conformity with the standard DIN EN ISO 15197:2003 in order to get the Conformité Européenne (CE) label for their products. The manufacturers submit the accuracy data required by the standard to a so-called notified body of the regulatory authority. The CE label is granted based on the manufacturer's accuracy data as well as an assessment of the manufacturer's quality system.

In this study we intended to investigate if 27 SMBG devices with a CE label from 18 manufacturers meet the accuracy requirements requested by DIN EN ISO 15197.

[1]Institute for Diabetes-Technology at the University of Ulm, Ulm, Germany.
[2]Institute for Medical Informatics and Biostatistics, Basel, Switzerland.

Subjects and Methods

The study was performed in 2008 (January–August) at the Institute for Diabetes-Technology in Ulm, Germany. The local ethics committee approved the study protocol. All subjects gave written informed consent. The methodology applied in this study is described in detail in DIN EN ISO 15197:2003.[12] Deviations from this standard are described in this section.

Study population

The study population comprised adult patients (>18 years old) with diabetes mellitus type 1 and type 2, as well as subjects without diabetes. Exclusion criteria were (1) pregnancy or lactation period for female subjects, (2) severe acute disease, and (3) severe chronic disease endangering the subject due to the study. Interruption criteria for individual subjects were (1) retraction of written informed consent and (2) incidences or adverse events interfering with the study continuation. Testing of BG monitoring system accuracy was performed in five test series with a duration of at least 10 days. For each tested system 100 subjects were included.

BG monitoring systems

The following 27 BG monitoring systems, consisting of a BG meter and BG test strips each, were compared: Accu-Chek® Active (Roche Diagnostics, Indianapolis, IN), Accu-Chek® Aviva (Roche Diagnostics), Ascensia Contour® (Bayer HealthCare LLC, Mishawaka, IN), Bayer Contour™ TS (Bayer Consumer Care AG, Basel, Switzerland), Beurer GL 30 (Beurer GmbH & Co. KG, Ulm, Germany), Bionime Rightest™ GM101 (Bionime Corp., Dali City, Taiwan), Bionime Rightest GM300 (Bionime Corp.), Clever Chek® TD-4222 (Taidoc Technology Corp., San-Chung, Taiwan), Finetest™ (Infopia Co., Ltd., Anyang, Republic of Korea), Finetest Auto-coding™ (Infopia Co., Ltd.), FineTouch™ (Terumo Corp., Tokyo, Japan), Fora® TD-4227 (Taidoc Technology Corp.), FreeStyle Freedom™ (Abbott Diabetes Care Inc., Alameda, CA), FreeStyle Lite® (Abbott Diabetes Care Inc.), GlucoCard™ X-Meter (Arkray, Inc., Kyoto, Japan), Glucofix® mio (Menarini Diagnostics S.r.l., Florence, Italy), GlucoHexal® (Allmedicus Co., Ltd., Seoul, Republic of Korea), Gluco-Test TD-4209 (Taidoc Technology Corp.), IME-DC BG monitoring system (IME-DC, Oberkotzau, Germany), OneTouch® Ultra®2 (LifeScan Inc., Milpitas, CA), OneTouch UltraEasy™ (LifeScan Inc.), Optium™ Xceed™ (E) bought in Spain (MediSense, Abingdon, UK) and Optium Xceed (F) bought in France (Abbott Diabetes Care Ltd., Witney, UK), SensoCardPlus (77 Elektronika Kft., Budapest, Hungary), SmartLAB® sprint (HMM Diagnostics GmbH, Dossenheim, Germany), Stada Glucocheck® (Home Diagnostics, Inc., Fort Lauderdale, FL), and Wellion® Linus (AgaMatrix, Salem, NH). The BG monitoring systems displayed either whole blood BG values or plasma equivalent BG values in mg/dL.

Reference methods

Reference measurements were performed with two different methods for all BG monitoring systems: (1) glucose oxidase (YSI 2300 STAT Plus glucose analyzer, YSI Life Sciences, Yellow Springs, OH), and (2) hexokinase (Hitachi 917, Roche Diagnostics GmbH, Mannheim, Germany). The accuracy of the glucose oxidase method was verified measuring NERL Glucose Standards (Thermo Fisher Scientific, East Providence, RI), verified against National Institute of Standards and Technology (NIST) (Gaithersburg, MD) reference material. The accuracy of the hexokinase method was verified measuring NIST Standard Reference Material 965a. Measurements with reference method 1 were performed at the Institute for Diabetes-Technology; measurements with reference method 2 were performed at Roche Diagnostics GmbH. The reference method used for evaluation of an individual BG monitoring system was chosen according to the manufacturer's device labeling. If no reference method was specified in the manufacturer's device labeling, reference method 1 was used. If the specified reference method was not available, method 1 or 2 were chosen depending on the chemistry of the specified reference method. Reference method 1 is the reference method for all BG monitoring systems with the exception of Finetest, Accu-Chek Active, and Accu-Chek Aviva. Reference method 2 is the reference method for Finetest, Accu-Chek Active, and Accu-Chek Aviva. Measurements with both reference methods were performed from whole blood, and both reference methods showed whole blood BG values in mg/dL. For plasma-calibrated systems, results from reference method 1 were converted from whole blood BG values to plasma equivalent BG values as follows: plasma equivalent BG value (in mg/dL) = whole blood BG value (in mg/dL)/[1 − (0.0024× hematocrit value [in %])],[13] with the exception of Optium Xceed, for which plasma equivalent BG value (in mg/dL) = 1.12× whole blood BG value (in mg/dL). Results from reference method 2 were converted from whole blood BG values to plasma equivalent BG values as follows: plasma equivalent BG value (in mg/dL) = 1.11×whole blood BG value (in mg/dL).

Study targets

The standard ISO 15197 specifies that the BG concentrations of the blood samples shall be distributed as follows: 5% <50 mg/dL, 15% ≥50 to <80 mg/dL, 20% ≥80 to <120 mg/dL, 30% ≥120 to <200 mg/dL, 15% ≥200 to <300 mg/dL, 10% ≥300 to <400 mg/dL, and 5% ≥400 mg/dL. The blood samples are distributed into the different concentration categories based on the mean reference BG values of reference method 1. The drift between the first and second reference measurement must be ≤4 mg/dL at BG concentrations ≤100 mg/dL and ≤4% at BG concentrations >100 mg/dL. Unaltered whole blood samples were used at BG concentrations of 50 mg/dL to 400 mg/dL. If sufficient numbers of unaltered whole blood samples with BG concentrations <50 mg/dL were not available, additional samples were prepared as follows: the blood samples were collected in lithium heparin test tubes, incubated at room temperature to allow for glycolysis, and gently mixed before testing. If sufficient numbers of unaltered whole blood samples with BG concentrations >400 mg/dL were not available, additional samples were prepared as follows: the blood samples were collected in lithium heparin test tubes, supplemented with concentrated glucose solution (40% glucose in 0.9% NaCl), and gently mixed before testing. The resulting hematocrit value was calculated and checked to be within the range specified by the manufacturer.

Test procedure

The tests were performed in a laboratory setting with controlled room temperature (23 ± 5°C) and humidity. The

tests were performed by clinical personnel trained to the limitations of the BG monitoring systems, the manufacturer's device labeling, the safety practices, and the test protocol. For each BG monitoring system two individual BG meters were used during the entire test procedure. Suitable control procedures were performed daily prior to each evaluation. Only in case of failure was the BG meter replaced. For each BG monitoring system one lot of BG test strips was used. Test strips were taken from at least seven different packages or vials [three packages with 150 single packaged test strips for Optium Xceed (F)], and the vials were changed after about 10 subjects. Measurements were performed on at least 10 days. Two different BG monitoring systems were tested in parallel.

Test protocol

The steps in the testing procedure were as follows: (1) A fresh capillary blood sample was obtained from the subject by skin puncture. (2) BG concentration was measured with reference method 1 in duplicate. (3) A 20-μL aliquot was removed and deproteinized. (4) BG concentration was measured with BG meter 1 of BG monitoring system 1. (5) BG concentration was measured with BG meter 2 of BG monitoring system 1. (6) BG concentration was measured with BG meter 1 of BG monitoring system 2. (7) BG concentration was measured with BG meter 2 of BG monitoring system 2. (8) BG concentration was measured with reference method 1 in duplicate. (9) A 20-μL aliquot was removed and deproteinized. Steps (1)–(9) were repeated until at least 100 capillary blood samples from 100 subjects had been collected (hematocrit value within manufacturer's range, distribution of BG concentrations as described). The 20-μL aliquots from steps (3) and (9) were deproteinized in test tubes containing 400 μL of 0.33 mol/L perchloric acid. The test tubes were centrifuged, and the supernatants were transferred to fresh test tubes and stored at -20°C for later triplicate testing with reference method 2.

Statistical analyses

Data were excluded from statistical analysis if (1) a handling error occurred, (2) no reference value was available, (3) a technical error was documented, (4) the data set was not complete, (5) the hematocrit value was outside the validated range of the respective BG monitoring system, (6) the maximum number of samples in a given BG concentration category was already reached, (7) the temperature was outside the range $23 \pm 5^{\circ}$C, (8) the humidity was outside the range specified by the manufacturer, or (9) the drift between the first and second reference measurement was >4 mg/dL at BG concentrations $\leq$100 mg/dL or >4% at BG concentrations >100 mg/dL. Data of 100 subjects were included in the system accuracy evaluation. Calculations were performed in mmol/L with a conversion factor of 18.02.

The BG monitoring system results were compared to the mean result of the two duplicate reference measurements, performed immediately before and after the measurements with the BG monitoring system. At BG concentrations <75 mg/dL the absolute and relative number of BG monitoring system results within $\pm$15 mg/dL, $\pm$10 mg/dL, and $\pm$5 mg/dL of the reference were tabulated. At BG concentrations $\geq$75 mg/dL the absolute and relative number of BG monitoring system results within $\pm$20%, $\pm$15%, $\pm$10%, and $\pm$5% of the reference were tabulated. In order to assess the overall accuracy of a BG monitoring system the percentage of BG monitoring system results within $\pm$15 mg/dL at BG concentrations <75 mg/dL was added to the percentage of results within $\pm$20% at BG concentrations $\geq$75 mg/dL. The differences between individual BG monitoring system results and the reference results were plotted against the average of the BG monitoring system and the reference results in Bland-Altman plots.[14] The individual BG monitoring system results were plotted against the reference results in Clarke error grids.[15] The absolute and relative numbers of BG monitoring system results within error zones A, B, C, D, and E were tabulated.

Results

The accuracy of 27 BG monitoring systems was assessed by comparison with a reference method. For each BG monitoring system 200 results were obtained from 100 subjects according to the standard DIN EN ISO 15197:2003.[12]

Table 1 shows the percentage of BG monitoring system results within $\pm$15 mg/dL, $\pm$10 mg/dL, and $\pm$5 mg/dL of the reference at BG concentrations <75 mg/dL and the percentage of BG monitoring system results within $\pm$20%, $\pm$15%, $\pm$10%, and $\pm$5% of the reference at BG concentrations $\geq$75 mg/dL. At BG concentrations <75 mg/dL 55.6% of the BG monitoring systems have $\geq$95% of their individual results within $\pm$15 mg/dL of the reference method, 7.4% within $\pm$10 mg/dL, and 0.0% within $\pm$5 mg/dL. At BG concentrations $\geq$75 mg/dL 77.8% of the BG monitoring systems have $\geq$95% of their results within $\pm$20% of the reference method, 37.0% within $\pm$15%, 3.7% within $\pm$10%, and 0.0% within $\pm$5%.

For BG system accuracy determination according to the standard, at least 95% of the individual BG monitoring system results shall fall within $\pm$15 mg/dL of the reference at BG concentrations <75 mg/dL and within $\pm$20% of the reference at BG concentrations $\geq$75 mg/dL. The number of acceptable results at BG concentrations <75 mg/dL is added to the number of acceptable results at BG concentrations $\geq$75 mg/dL to determine the total number of acceptable results. Sixteen BG monitoring systems (59.3%) fulfilled the minimum accuracy requirements of the standard; 11 systems (40.7%) did not. Overall, the mean percentage of results showing the minimum acceptable accuracy calculated from all 27 tested BG monitoring systems was 95.2 ± 5.2%, ranging from 80.0% to 100.0%.

The relative differences in percentages between an individual result of a BG monitoring system and the corresponding result of the reference method were plotted against the average of the result of the BG monitoring system and the corresponding result of the reference method in Bland-Altman plots (Fig. 1). For easier comparison of the BG monitoring systems the plot area had a standardized range from -40% to $+40$%. For some BG monitoring systems not all 200 data points fitted into the standardized plot area. The mean bias of the BG monitoring systems compared to the reference method was 0.2 ± 5.1% and ranged from -7.8% to $+13.1$%. The mean lower 95% limit of agreement of the BG monitoring systems compared to the reference method was -19.5 ± 9.3% and ranged from -40.6% to -3.2%. The mean upper 95% limit of agreement of the BG monitoring systems compared to the

Table 1. BG Monitoring System Accuracy Results

	ISO 15197		*BG concentration <75 mg/dL*			*BG concentration ≥75 mg/dL*			
BG monitoring system	*Requirements fulfilled*[a]	*Percentage*	*±15 mg/dL*	*±10 mg/dL*	*±5 mg/dL*	*±20%*	*±15%*	*±10%*	*±5%*
Accu-Chek Active	**Yes**	**100.0%** (200/200)	**100%** (30/30)	93% (28/30)	47% (14/30)	**100%** (170/170)	**98%** (166/170)	88% (149/170)	52% (88/170)
Accu-Chek Aviva	**Yes**	**100.0%** (200/200)	**100%** (36/36)	83% (30/36)	69% (25/36)	**100%** (164/164)	**99%** (162/164)	**95%** (156/164)	71% (117/164)
Ascensia Contour	**Yes**	**98.5%** (197/200)	**100%** (36/36)	86% (31/36)	64% (23/36)	**98%** (161/164)	91% (149/164)	74% (121/164)	44% (72/164)
Bayer Contour TS	No	90.0% (180/200)	**98%** (39/40)	83% (33/40)	43% (17/40)	88% (141/160)	70% (112/160)	45% (72/160)	18% (29/160)
Beurer GL 30[b]	No	91.3% (179/196)	61% (22/36)	44% (16/36)	22% (8/36)	**98%** (157/160)	90% (144/160)	68% (108/160)	37% (59/160)
Bionime Rightest GM101	**Yes**	**100.0%** (200/200)	**100%** (40/40)	**100%** (40/40)	80% (32/40)	**100%** (160/160)	**99%** (158/160)	93% (148/160)	61% (97/160)
Bionime Rightest GM300	**Yes**	**100.0%** (200/200)	**100%** (40/40)	93% (37/40)	75% (30/40)	**100%** (160/160)	**95%** (152/160)	84% (135/160)	53% (85/160)
Clever Chek TD-4222	No	90.5% (181/200)	71% (24/34)	56% (19/34)	21% (7/34)	**95%** (157/166)	84% (139/166)	68% (113/166)	37% (62/166)
Finetest	**Yes**	**97.5%** (195/200)	**100%** (32/32)	91% (29/32)	63% (20/32)	**97%** (163/168)	92% (155/168)	65% (110/168)	27% (45/168)
Finetest Auto-coding	No	94.5% (189/200)	89% (34/38)	63% (24/38)	24% (9/38)	**96%** (155/162)	90% (145/162)	71% (115/162)	39% (63/162)
FineTouch	No	90.0% (180/200)	94% (34/36)	72% (26/36)	31% (11/36)	89% (146/164)	82% (134/164)	71% (116/164)	41% (68/164)
Fora TD-4227	No	89.0% (178/200)	84% (32/38)	50% (19/38)	16% (6/38)	90% (146/162)	78% (126/162)	60% (97/162)	30% (48/162)
FreeStyle Freedom	**Yes**	**100.0%** (200/200)	**100%** (38/38)	92% (35/38)	66% (25/38)	**100%** (162/162)	**100%** (162/162)	94% (152/162)	69% (111/162)
FreeStyle Lite	**Yes**	**100.0%** (200/200)	**100%** (36/36)	89% (32/36)	47% (17/36)	**100%** (164/164)	**99%** (162/164)	87% (142/164)	53% (87/164)
GlucoCard-X-Meter	**Yes**	**99.0%** (198/200)	**95%** (36/38)	92% (35/38)	74% (28/38)	**100%** (162/162)	93% (150/162)	82% (133/162)	52% (84/162)
Glucofix mio	No	93.5% (187/200)	80% (32/40)	50% (20/40)	25% (10/40)	**97%** (155/160)	94% (150/160)	76% (121/160)	41% (65/160)
GlucoHexal	No	80.0% (160/200)	70% (21/30)	30% (9/30)	7% (2/30)	82% (139/170)	65% (111/170)	43% (73/170)	16% (27/170)
Gluco-Test TD-4209	No	94.5% (189/200)	91% (31/34)	79% (27/34)	50% (17/34)	**95%** (158/166)	88% (146/166)	68% (113/166)	31% (52/166)
IME-DC BG meter	No	88.5% (177/200)	76% (26/34)	47% (16/34)	26% (9/34)	91% (151/166)	87% (144/166)	70% (117/166)	37% (62/166)
OneTouch Ultra 2	**Yes**	**100.0%** (200/200)	**100%** (36/36)	**97%** (35/36)	69% (25/36)	**100%** (164/164)	**96%** (158/164)	76% (124/164)	43% (71/164)
OneTouch Ultra Easy	**Yes**	**99.0%** (198/200)	**100%** (36/36)	94% (34/36)	75% (27/36)	**99%** (162/164)	91% (150/164)	71% (116/164)	30% (50/164)
Optium Xceed (E)	**Yes**	**98.5%** (197/200)	**98%** (39/40)	93% (37/40)	43% (17/40)	**99%** (158/160)	93% (149/160)	78% (125/160)	44% (71/160)
Optium Xceed (F)	**Yes**	**99.0%** (198/200)	**98%** (39/40)	90% (36/40)	63% (25/40)	**99%** (159/160)	**95%** (152/160)	77% (123/160)	46% (74/160)
SensoCardPlus	**Yes**	**99.0%** (198/200)	**97%** (33/34)	68% (23/34)	26% (9/34)	**99%** (165/166)	**98%** (162/166)	92% (152/166)	58% (97/166)
SmartLAB sprint	**Yes**	**96.0%** (192/200)	89% (34/38)	55% (21/38)	13% (5/38)	**98%** (158/162)	**96%** (155/162)	89% (144/162)	58% (94/162)
Stada Glucocheck	No	88.0% (176/200)	71% (27/38)	58% (22/38)	24% (9/38)	92% (149/162)	82% (133/162)	66% (107/162)	43% (69/162)
Wellion Linus	**Yes**	**95.0%** (190/200)	79% (30/38)	50% (19/38)	8% (3/38)	**99%** (160/162)	93% (151/162)	80% (130/162)	49% (80/162)

Values are percentages (valid *n*/total *n*). Values ≥95% are in bold type; remaining values are <95%.

[a]BG results within ±15 mg/dL at BG concentrations <75 mg/dL and within ±20% at BG concentrations ≥75 mg/dL.

[b]The manufacturer's device labeling states that this BG monitoring system is calibrated to whole blood. Results of the analysis versus the whole blood reference method are as follows: ISO 15197 requirements fulfilled, no 73.5% (147/200); BG concentration <75 mg/dL, ±15 mg/dL 43% (17/40), ±10 mg/dL 38% (15/40), and ±5 mg/dL 25% (10/40); and BG concentration ≥75 mg/dL, ±20% 81% (130/160), ±15% 72% (115/160), ±10% 53% (85/160), and ±5% 29% (47/160). According to personal communication with the manufacturer (Beurer GmbH & Co. KG) "the system is plasma calibrated and had been declared as whole blood calibrated by mistake." Results of the analysis versus the plasma reference method are shown here.

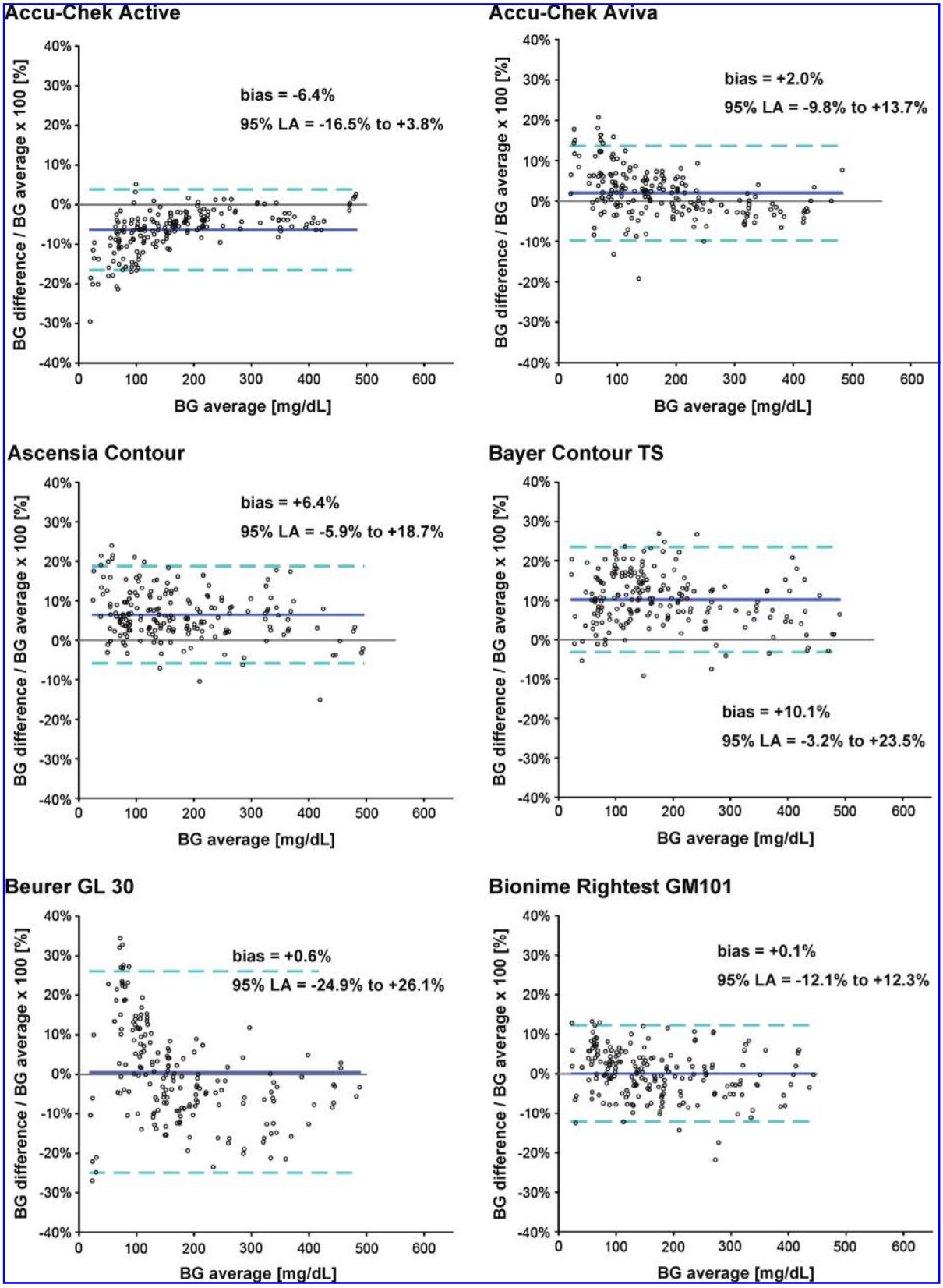

FIG. 1. (*continued*).

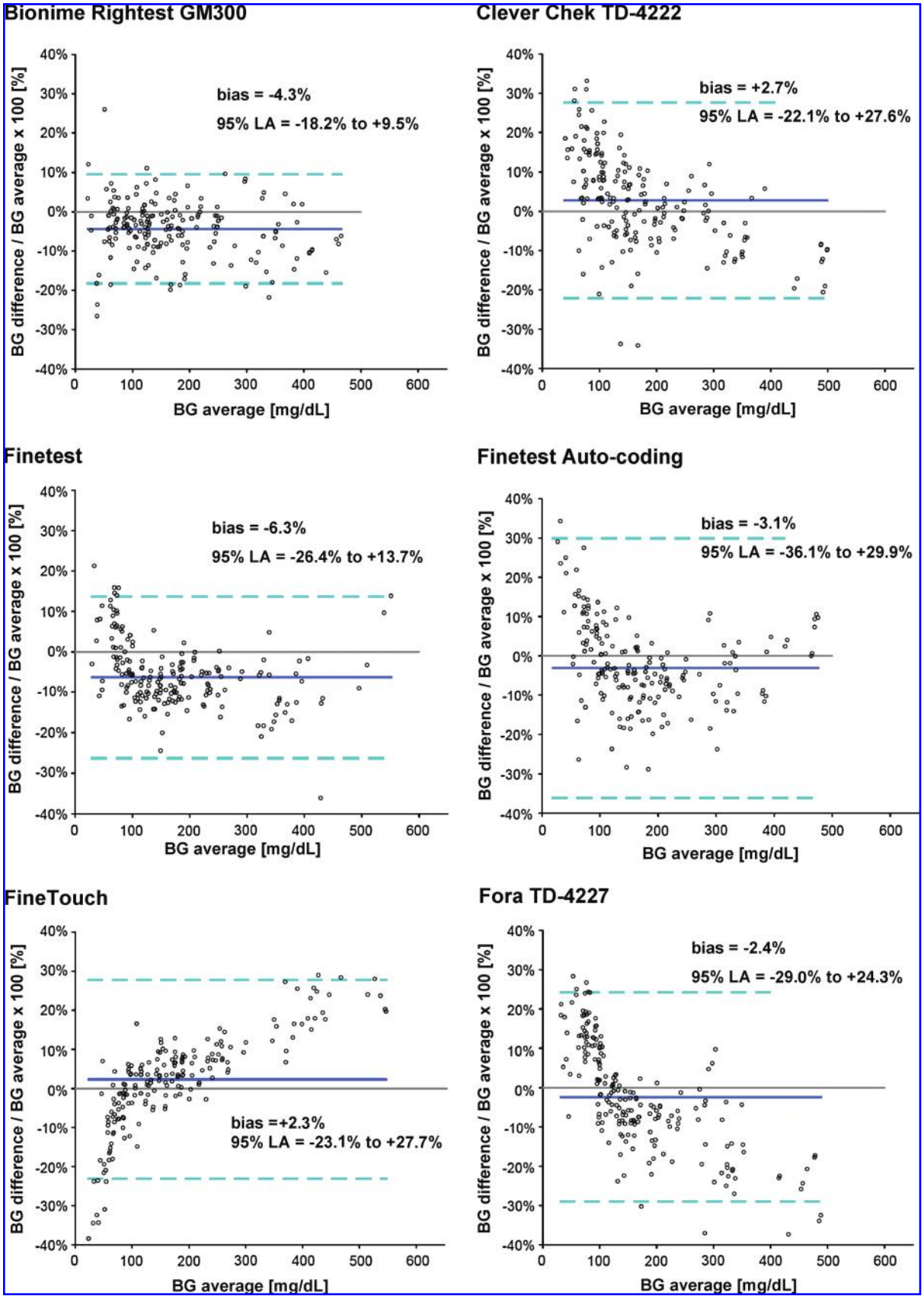

FIG. 1. (*continued*).

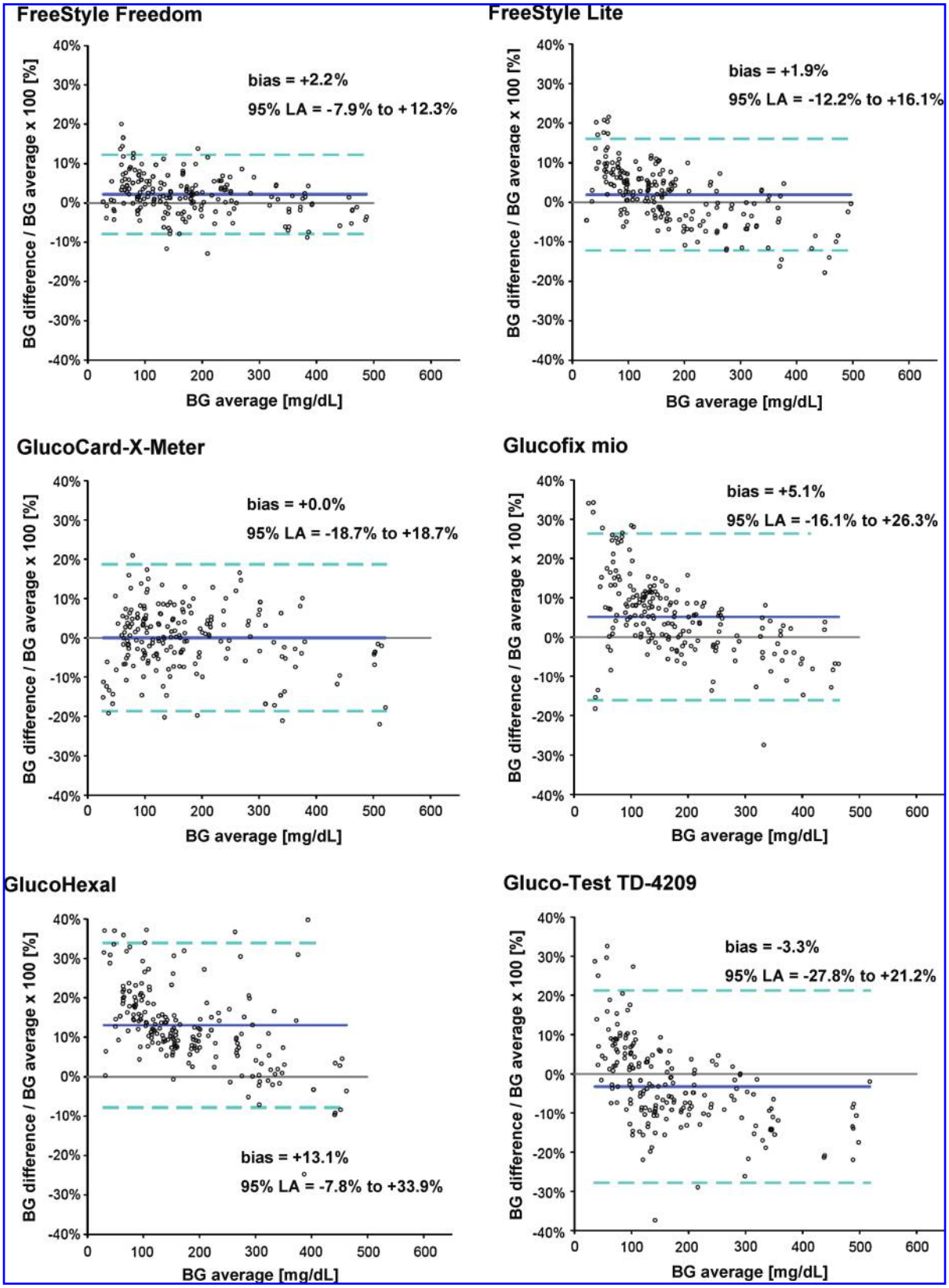

FIG. 1. (*continued*).

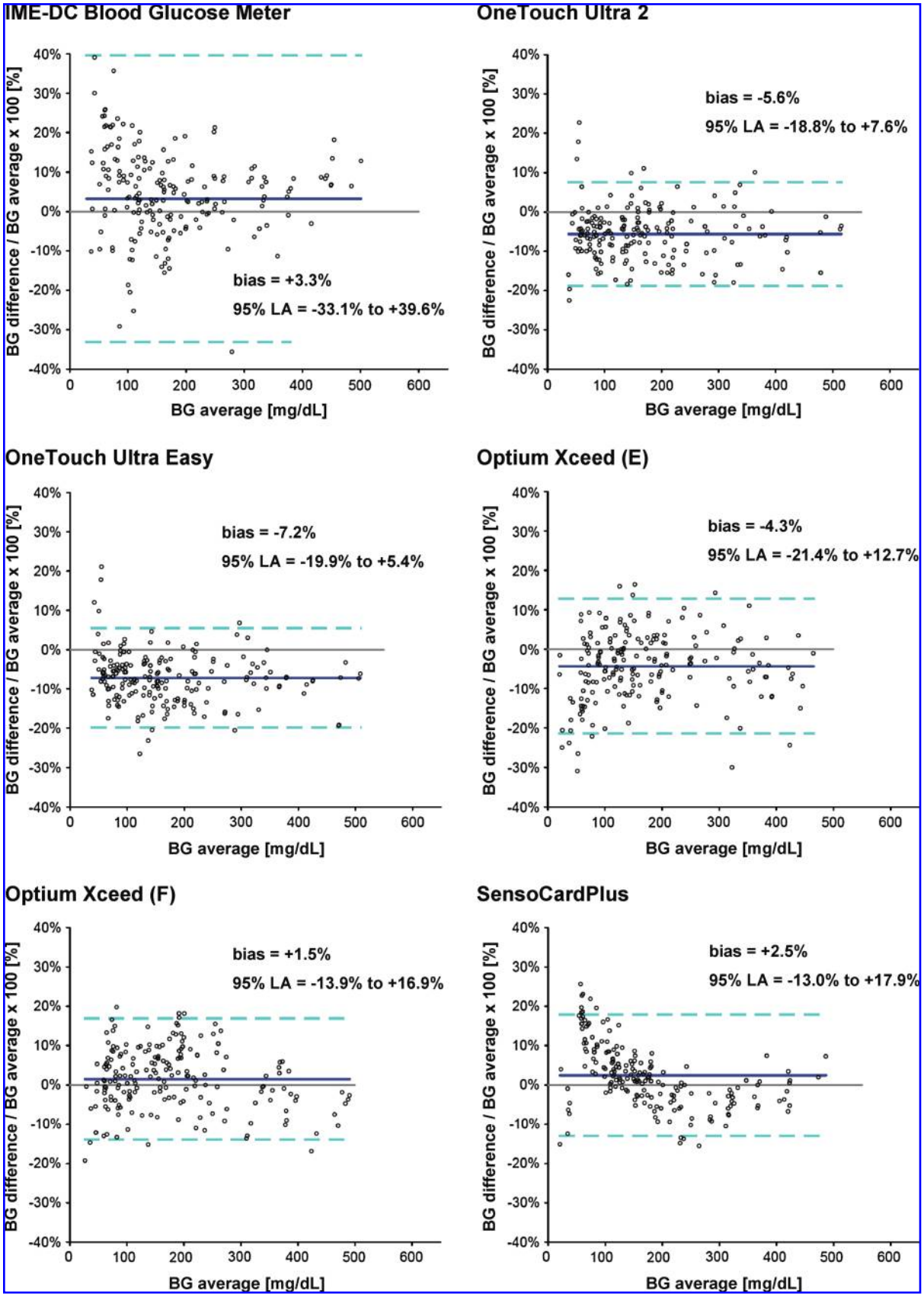

FIG. 1. (*continued*).

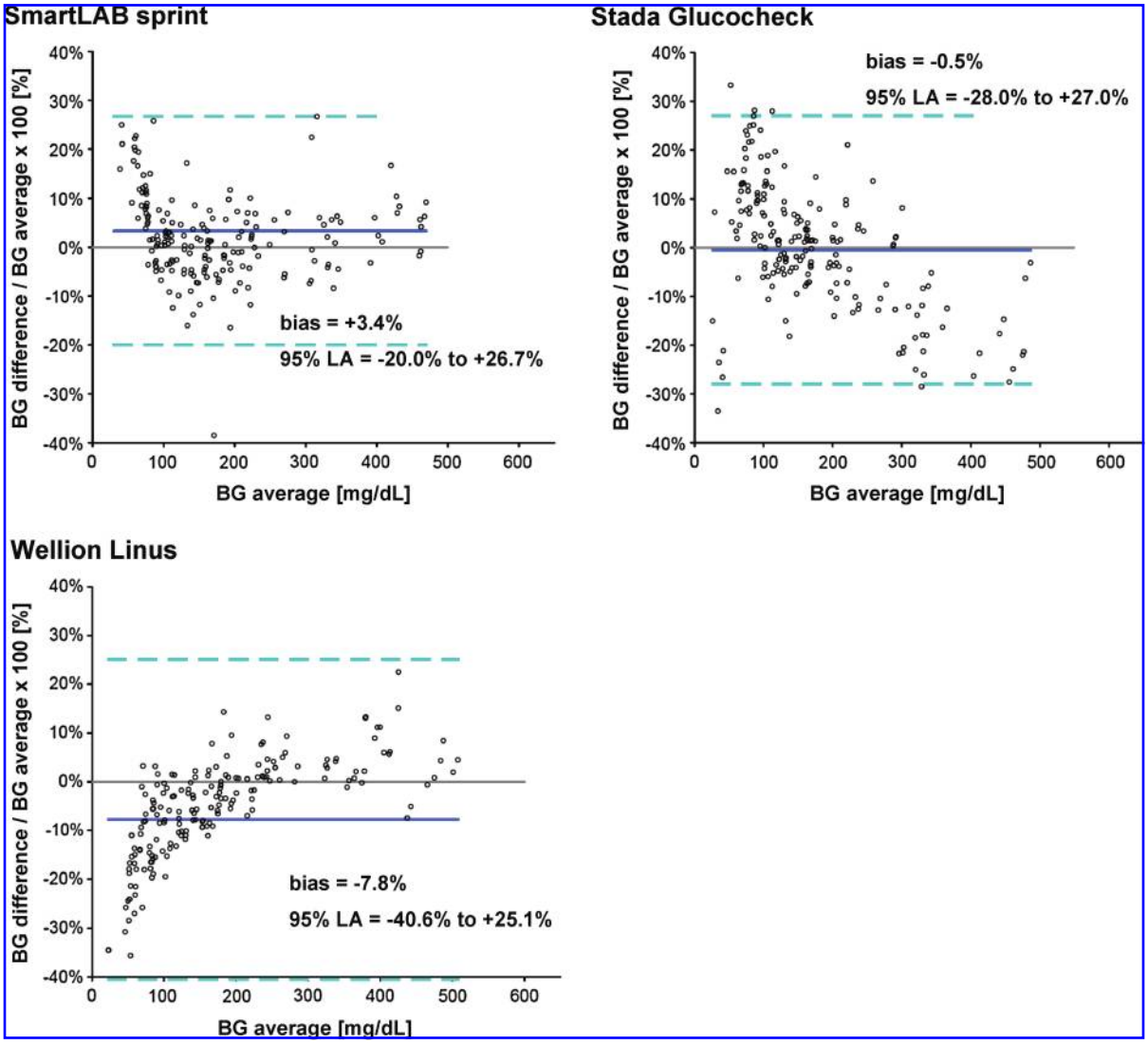

FIG. 1. Bland-Altman plots of 27 BG monitoring systems. The relative differences in percentages between an individual result of a BG monitoring system and the corresponding result of the reference method were plotted against the average of the result of the BG monitoring system and the corresponding result of the reference method. For some BG monitoring systems not all 200 data points fit into the standardized plot area from −40% to +40%. The solid line at 0% difference is the line of identity, the solid black line is the bias, and the dashed gray lines are the 95% limits of agreement (LA). Color images available online at www.liebertonline.com/dia.

reference method was $+19.9 \pm 8.9\%$ and ranged from +3.8% to +39.6%.

The individual BG monitoring system results were plotted against the reference results in Clarke error grids. The absolute and relative number of BG monitoring system results within error zones A–E were tabulated (Table 2), where A represents "clinically accurate," B represents "benign reading error," C represents "overcorrection," D represents "failure to detect," and E represents "errors." Five BG monitoring systems (18.5%) had 100.0% of their results in error zone A, i.e., only results that are "clinically accurate." Thirteen BG monitoring systems (48.1%) had 100.0% of their results in error zones A and B, i.e., only results that are "clinically accurate" or "benign reading errors." Fourteen BG monitoring systems (51.9%) additionally had results in error zones C, D, or E, i.e., results that are "overcorrections," "failures to detect," or "errors," respectively.

Discussion

This study aimed to investigate if 27 CE-labeled BG monitoring systems from 18 different manufacturers meet the requirements defined in the standard DIN EN ISO 15197. The evaluation was performed in a laboratory by trained medical personnel under controlled conditions. The test procedures, protocols, and statistical analyses of the results were as described in DIN EN ISO 15197. Sixteen of the 27 BG monitoring systems fulfilled the minimum accuracy requirements of the standard, i.e., they had ≥95% of their individual BG results within ±15 mg/dL of the reference method at BG concentrations <75 mg/dL and within ±20% of the reference method at BG concentrations ≥75 mg/dL. If the minimum accuracy requirements were as stringent as proposed by the American Diabetes Association, i.e., ≥95% of the individual BG results

Table 2. Clarke Error Grid Analysis

BG monitoring system	Clarke error grid zone				
	A (clinically accurate)	*B (benign reading error)*	*C (overcorrection)*	*D (failure to detect)*	*E (errors)*
Accu-Chek Active	100.0% (200)	0% (0)	0% (0)	0% (0)	0% (0)
Accu-Chek Aviva	99.5% (199)	0% (0)	0% (0)	**0.5%** (1)	0% (0)
Ascensia Contour	98.5% (197)	1.5% (3)	0% (0)	0% (0)	0% (0)
Bayer Contour TS	90.0% (180)	10.0% (20)	0% (0)	0% (0)	0% (0)
Beurer GL 30[a]	89.8% (176)	4.6% (9)	0% (0)	**5.6%** (11)	0% (0)
Bionime Rightest GM101	100.0% (200)	0% (0)	0% (0)	0% (0)	0% (0)
Bionime Rightest GM300	100.0% (200)	0% (0)	0% (0)	0% (0)	0% (0)
Clever Chek TD-4222	92.0% (184)	5.0% (10)	0% (0)	**3.0%** (6)	0% (0)
Finetest	97.5% (195)	2.5% (5)	0% (0)	0% (0)	0% (0)
Finetest Auto-coding	95.5% (191)	4.0% (8)	0% (0)	**0.5%** (1)	0% (0)
FineTouch[b]	93.8% (182)	4.6% (9)	**1.6%** (3)	0% (0)	0% (0)
Fora TD-4227[c]	87.9% (174)	9.1% (18)	0% (0)	**3.0%** (6)	0% (0)
FreeStyle Freedom	100.0% (200)	0% (0)	0% (0)	0% (0)	0% (0)
FreeStyle Lite	99.5% (199)	0% (0)	0% (0)	**0.5%** (1)	0% (0)
GlucoCard-X-Meter	99.0% (198)	0.5% (1)	0% (0)	**0.5%** (1)	0% (0)
Glucofix mio	93.0% (186)	4.0% (8)	0% (0)	**3.0%** (6)	0% (0)
GlucoHexal	78.0% (156)	14.5% (29)	**1.0%** (2)	**6.5%** (13)	0% (0)
Gluco-Test TD-4209	96.0% (192)	4.0% (8)	0% (0)	0% (0)	0% (0)
IME-DC BG meter	90.0% (180)	7.5% (15)	0% (0)	**2.0%** (4)	**0.5%** (1)
OneTouch Ultra 2	100.0% (200)	0% (0)	0% (0)	0% (0)	0% (0)
OneTouch Ultra Easy	99.0% (198)	1.0% (2)	0% (0)	0% (0)	0% (0)
Optium Xceed (E)	99.0% (198)	1.0% (2)	0% (0)	0% (0)	0% (0)
Optium Xceed (F)	99.0% (198)	1.0% (2)	0% (0)	0% (0)	0% (0)
SensoCardPlus	99.0% (198)	0.5% (1)	0% (0)	**0.5%** (1)	0% (0)
SmartLAB sprint	97.0% (194)	2.5% (5)	0% (0)	**0.5%** (1)	0% (0)
Stada Glucocheck	88.5% (177)	9.5% (19)	0% (0)	**2.0%** (4)	0% (0)
Wellion Linus	99.0% (198)	1.0% (2)	0% (0)	0% (0)	0% (0)

Values are percentages (valid *n*). Clinically critical values in error zones C–E >0% are in bold; remaining values are in clinically uncritical error zones A + B.

[a]Only 196 BG values ≤550 mg/dL were included into the analysis.

[b]Only 194 BG values ≤550 mg/dL were included into the analysis.

[c]Only 198 BG values ≤550 mg/dL were included into the analysis.

should be within ±10%[16] or even ±5%[17] of the reference method instead of ±20%, none of the BG monitoring systems would conform to these stricter accuracy requirements.

There are a number of factors influencing the outcome of a BG monitoring system in this kind of investigation. First, lot-to-lot differences of a given BG monitoring system are not detected by assessing only one lot. Thus it could be useful to analyze several lots in order to evaluate a BG monitoring system. Second, the overall errors observed in a method comparison always include the error of the reference method, i.e., there is a significant contribution of the reference method to the inaccuracies reported here. The contribution of different reference methods with respect to inaccuracy of the overall result of the BG monitoring systems might be different. For example, Twomey[18] found the YSI 2300 STAT, using the glucose oxidase method, to be negatively biased for plasma glucose when compared with the Olympus (Tokyo, Japan) AU640™ hexokinase method. These differences between reference methods would introduce additional bias that is not due to the BG monitoring systems tested. Third, the conversion factor from whole blood BG values to plasma equivalent BG values is specific for the manufacturer. Not all manufacturers stick to the recommendations of the International Federation of Clinical Chemistry[19] on reporting of BG results. The results may also vary depending on whether whole blood or plasma samples are used for reference measurements. For example, there may be slight differences between plasma equivalent BG concentrations calculated by multiplying whole blood concentrations with a conversion factor—as performed in this investigation—and BG concentrations measured from separated plasma. In order to increase the comparability of BG monitoring system assessments by manufacturers, it could be useful to standardize both the reference method and the blood sample type, i.e., measurement in whole blood or plasma. Fourth, the standard DIN EN ISO 15197 defines two BG concentration categories (<75 mg/dL and ≥75 mg/dL) for assessment of accuracy. However, real-life BG levels are continuous and span a range of values. Ideally, the performance of BG monitoring systems would be equal over the complete range of BG values. However, real BG monitoring systems will, e.g., perform better at normal BG concentrations than at the extremes of the measurement range, or vice versa, or exhibit completely different performance characteristics. To visualize these different measurement characteristics, we used Bland-Altman plots to display both bias and precision over the complete range of BG values assessed. It could also be useful to weight the accuracy of a BG monitoring system in different BG concentration ranges differently in analogy to the zones in Clarke error grid analysis. This would allow for a more detailed comparison of the accuracy of different BG

monitoring systems. Considering all these factors influencing the outcome of BG monitoring systems, the stringent threshold between BG monitoring systems fulfilling or missing the minimum accuracy requirements defined by DIN EN ISO 15197 could possibly lead to misjudgment of individual systems.

In addition to the evaluation according to EN ISO 15197 the data of the 27 BG monitoring systems were analyzed in Clarke error grids in order to assess the clinical relevance of the results. Thirteen BG monitoring systems had 100% of their results in error zones A and B, which are considered as "clinically uncritical"; 14 BG monitoring systems additionally had results in error zones C–E, which are considered as "clinically critical." According to Clarke et al.[15] and Clarke,[20] "clinically critical" results bear the risk of wrong therapy decisions, e.g., inappropriate adjustments to the insulin dose or inappropriate intake of carbohydrates. Depending on the degree and the direction of these inappropriate adjustments the consequences for the diabetes patient can vary from harmless to serious health injuries.

In summary, we found more than 40% of BG monitoring systems with a CE label that did not meet the minimal accuracy requirements of the standard DIN EN ISO 15197. As inaccurate BG monitoring systems bear the risk of false treatment decisions by the diabetes patient and subsequent possible severe health injury, manufacturers should regularly and effectively check the quality of BG meters and test strips.

Acknowledgments

This study was funded by a grant from Roche Diagnostics GmbH, Mannheim, Germany.

Author Disclosure Statement

G.F. received refunding of traveling expenses for congress participation from Roche Diagnostics GmbH, Mannheim, Germany, the sponsor of this study. A.B., N.J., E.Z., S.K., J.T., F.H., and C.H. declare no competing financial interests exist.

References

1. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Research Group. N Engl J Med 1993; 329:977–986.
2. U.K. Prospective Diabetes Study 16. Overview of 6 years' therapy of type II diabetes: a progressive disease. U.K. Prospective Diabetes Study Group. Diabetes 1995;44:1249–1258.
3. Blonde L, Karter AJ: Current evidence regarding the value of self-monitored blood glucose testing. Am J Med 2005;118(9 Suppl 1):20–26.
4. Welschen LM, Bloemendal E, Nijpels G, Dekker JM, Heine RJ, Stalman WA, Bouter LM: Self-monitoring of blood glucose in patients with type 2 diabetes who are not using insulin: a systematic review. Diabetes Care 2005;28:1510–1517.
5. Sarol JN, Nicodemus NA, Tan KM, Grava MB: Self-monitoring of blood glucose as part of a multi-component therapy among non-insulin requiring type 2 diabetes patients: a meta-analysis (1966–2004). Curr Med Res Opin 2005;21:173–183.
6. Jansen JP: Self-monitoring of glucose in type 2 diabetes mellitus: a Bayesian meta-analysis of direct and indirect comparisons. Curr Med Res Opin 2006;22:671–681.
7. McGeoch G, Moore RA: Self-monitoring of blood glucose in type-2 diabetes: what is the evidence? Diabetes Metab Res Rev 2007;23:423–440.
8. Palmer AJ, Dinneen S, Gavin JR, III, Gray A, Herman WH, Karter AJ: Cost-utility analysis in a UK setting of self-monitoring of blood glucose in patients with type 2 diabetes. Curr Med Res Opin 2006;22:861–872.
9. Schnell O, Hummel M, Weber C: Economic and clinical aspects of diabetes regarding self-monitoring of blood glucose. Diabetes Technol Ther 2008;10(Suppl 1):S72–S81.
10. American Diabetes Association: Standards of medical care in diabetes–2009. Diabetes Care 2009;32(Suppl 1):S13–S61.
11. IDF Clinical Guidelines Task Force: Global Guideline for Type 2 Diabetes. Brussels: Brussels: International Diabetes Federation, 2005.
12. DIN EN ISO 15197: In Vitro Diagnostic Test Systems–Requirements for Blood Glucose Monitoring Systems for Self-Testing in Managing Diabetes Mellitus (ISO 15197:2003). European Committee for Standardization, Brussels, 2003.
13. Astles JR, Sedor FA, Toffaletti JG: Evaluation of the YSI 2300 glucose analyzer: algorithm-corrected results are accurate and specific. Clin Biochem 1996;29:27–31.
14. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310.
15. 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.
16. American Diabetes Association: Consensus statement on self-monitoring of blood glucose. Diabetes Care 1987;10: 95–99.
17. Self-monitoring of blood glucose. American Diabetes Association. Diabetes Care 1994;17:81–86.
18. Twomey PJ: Plasma glucose measurement with the Yellow Springs Glucose 2300 STAT and the Olympus AU640. J Clin Pathol 2004;57:752–754.
19. D'Orazio P, Burnett RW, Fogh-Andersen N, Jacobs E, Kuwa K, Kulpmann WR, Larsson L, Lewenstam A, Maas AH, Mager G, Naskalski JW, Okorodudu AO: Approved IFCC recommendation on reporting results for blood glucose (abbreviated). Clin Chem 2005;51:1573–1576.
20. Clarke WL: The original Clarke Error Grid Analysis (EGA). Diabetes Technol Ther 2005;7:776–779.

Address correspondence to:
Serge Kocher, Ph.D.
Institute for Medical Informatics and Biostatistics
Clarastrasse 12
CH-4058 Basel, Switzerland

E-mail: kocher@imib.ch