

# Evolution of Analytical Performance in Portable Glucose Meters in the Last Decade

PHILIP BÖHME, MD  
MICHELE FLORIOT, MD  
MARIE-AUDE SIRVEAUX, MD  
DANIELLE DURAIN, RN

OLIVIER ZIEGLER, MD, PHD  
PIERRE DROUIN, MD  
BRUNO GUERCI, MD, PHD

**OBJECTIVE** — To assess and compare the technical accuracy of portable glucose meters during the last decade.

**RESEARCH DESIGN AND METHODS** — One-thousand preprandial (pre) and postprandial (post) capillary whole-blood glucose values measured with meters owned mainly by diabetic patients were compared with a single laboratory method yearly from 1989 to 1999. A total of 21,950 capillary measurements and their corresponding laboratory reference values were analyzed at our clinic.

**RESULTS** — The lowest mean absolute difference was found in 1989 (pre:  $2 \pm 22$  mg/dl, post:  $9 \pm 31$  mg/dl) (mean  $\pm$  SD). The highest mean absolute difference was observed in 1993 (pre:  $31 \pm 33$  mg/dl) and 1996 (post:  $50 \pm 35$  mg/dl). The highest mean relative deviation was observed in 1990 (pre: 16.4%) and 1996 (post: 20.6%). The highest percentage of readings that were within a 5% deviation limit were observed in 1998 (pre: 44.5%) and in 1997 (post: 36.7%). Based on blood glucose levels within  $\pm 5$  and  $\pm 10\%$  of laboratory values, the technical accuracy of meters was similar for 1989 and 1999 ( $P = 0.27$  and  $0.52$ , respectively). The percentage of pre values in zone A of Clarke's error grid analysis was  $>90\%$  in 1989, 1997, 1998, and 1999.

**CONCLUSIONS** — The analytical performance of glucose meters decreased between 1990 and 1996 but was restored between 1997 and 1999. Nevertheless, our data suggest that the technical accuracy of glucose meters has not significantly improved during the last decade. Complementary studies taking into account the preanalytical improvements of the recent meters, as well as their calibration method, appear necessary.

*Diabetes Care* 26:1170–1175, 2003

Self-monitoring of blood glucose is now considered to be an important tool in the management of insulin-treated diabetes (1). Its importance has increased since Diabetes Control and Complications Trial results have shown that proper control of glucose levels is associated with a reduced risk of microvascular disease in type 1 diabetes (2). Portable glucose meters are widely used

in type 2 diabetes, although their contribution to blood glucose improvement has not been established (3). In the last two decades, results obtained with glucose meters appear to have become more reproducible and accurate. The meters have also become smaller, faster for blood glucose analysis, and easier to use (4,5). Although there are some studies concerning user error (6,7), information relating to

the evolution of analytical error appears limited (8). To determine whether the analytical accuracy has changed in relation to the global improvement of meter performances, we have prospectively compared, at our clinic and under medical control, 1,000 preprandial and postprandial capillary blood glucose measures using various portable meters on a yearly basis between 1989 and 1999 in comparison to a single reference method (RM).

## RESEARCH DESIGN AND METHODS

In our Department of Diabetes, most of the diabetic patients (mainly type 1 and type 2 diabetes) benefit from an annual check-up during which each patient's device is evaluated in a standard manner. We also conduct a yearly check on the devices employed in our inpatient and outpatient clinics for diabetes. Other examinations, including an electrocardiogram, leg Doppler, and fluorescein angiography are also carried out on the same day. To evaluate the technical accuracy of each home blood glucose meter (HB), pre- and postlunch (1.5 h after the meal) meter readings were compared with the results of a RM. All capillary blood samples were taken at room temperature ( $\sim 20^\circ\text{C}$ ) after verification, calibration, and washing of the device by one of the experienced nurses working at our unit. Each month, these nurses were trained to use newly released devices coming onto the market. All patients were asked to wash and dry their hands with warm water, after which a finger prick was made with a standard autolancet, and a suitable drop of blood was applied to the test strip according to the manufacturer's instructions. Additional blood derived from the same capillary sample was collected into a microtube containing fluoride and immediately brought to the local laboratory, where a glucose oxidase method was used to determine the plasma glucose concentration. Laboratory analysis was performed after centrifugation (5 min at 18,000g) in a microcentrifuge. Glucose in the supernatant fluid was assayed with a reagent

From the Service de Diabétologie, Maladies Métaboliques & Maladies de la Nutrition, Hôpital Jeanne d'Arc, CHU Nancy, France.

Address correspondence and reprint requests to Dr. Philip Böhme, Service de Diabétologie, Maladies Métaboliques & Maladies de la Nutrition, Hôpital Jeanne d'Arc, Centre Hospitalo-Universitaire de Nancy, BP 303, 54201 Toul cedex, France. E-mail: p.bohme@chu-nancy.fr.

Received for publication 22 August 2002 and accepted in revised form 15 January 2003.

**Abbreviations:** EGA, error grid analysis; HB, home blood glucose meter; RM, reference method.

A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.

**Table 1—Yearly absolute means and relative deviations of the results obtained with glucose devices with respect to the reference value (RM – HB), percentages of values within  $\pm 5\%$  deviation from the reference value, and percentage of the total samples in zone A of the EGA**

Year of study	Samples (n)	Preprandial glycemicias				Postprandial glycemicias			
		Mean $\pm$ SD (mg/dl)	Relative deviation (%)	Values within $\pm 5\%$ (%)	Values in zone A (%)	Mean $\pm$ SD (mg/dl)	Relative deviation (%)	Values within $\pm 5\%$ (%)	Values in zone A (%)
1989	997	2 $\pm$ 22	11.8	27.0*	90.7	9 $\pm$ 31	10.3	32.0	91.2*
1990	998	22 $\pm$ 27	16.4	15.5	73.3	32 $\pm$ 32	14.7	16.3	74.4
1991	992	24 $\pm$ 25	14.8	15.6	74.5	36 $\pm$ 32	14.9	14.1	74.4
1992	998	24 $\pm$ 29	14.5	17.7	74.8	35 $\pm$ 34	15.3	13.8	70.2
1993	999	31 $\pm$ 33	15.5	14.7	71.0	50 $\pm$ 40	19.4	6.8	51.4
1994	997	24 $\pm$ 30	14.4	17.0	76.7	49 $\pm$ 37	19.9	5.2	49.5
1995	994	24 $\pm$ 29	15.2	17.6	73.4	43 $\pm$ 36	18.2	9.2	57.7
1996	1,000	29 $\pm$ 27	16.1	13.2	67.5	50 $\pm$ 35	20.6	4.6	48.3
1997	1,000	3 $\pm$ 16	7.8	41.3	98.4	13 $\pm$ 21	8.0	36.7	96.9
1998	1,000	7 $\pm$ 14	7.0	44.5	98.4	14 $\pm$ 15	8.0	32.4	98.0
1999	1,000	15 $\pm$ 14	9.9	24.9*	93.1	20 $\pm$ 19	10.2	25.2	92.2*
Total	10,975	19 $\pm$ 24	13.2	22.6	81.1	32 $\pm$ 30	14.3	17.8	73.1

\* $\chi^2$  test between 1989 and 1999,  $P > 0.05$ .

based on glucose oxidase and peroxidase (Glucose enzymatique PAP 7500; BioMérieux SA, Marcy-L'étoile, France) in an analyzer (Lab System). Glucose concentration was calculated by using a molar absorptivity (505 nm). The same analyzer was used over the entire study period. The coefficient of variation of this technique was  $<3\%$  during the overall study period.

All results, comprising preprandial (pre) and postprandial (post) readings for all patients, based on their glucose meter and the corresponding laboratory micro-method measurements, were recorded. In this study, 1,000 pre and post samples from 1 January of each year between 1989 and 1999 were used for the statistical analysis. Measurements that were found to be "low" or "high" by the metering devices during the pre or post period were excluded from the evaluation.

### Statistical methods

The intraclass correlation coefficient, as outlined by Landis and Koch (9), was applied to examine the agreement between HB and the RM for each year. Results are shown as a coefficient and the 95% CI.

To assess agreement between the meters and the RM, we used the method of residuals, referred to as the method of Bland and Altman (10). Briefly, the differences between the RM and the HB (residuals) were calculated first, followed by the mean and SD of the differences. The mean  $\pm$  1.96 SD represented the 95% CI.

Based on the American Diabetes As-

sociation (ADA) consensus statement, we used the recommended performance goal of glucose meters: a total analytic error of  $<10\%$  (present ADA and French criteria) or  $<5\%$  (future ADA criteria) (1,11).

Statistical methods included the  $\chi^2$  test to compare frequencies, and significance was implied at  $P < 0.05$ .

Clinical relevance was also analyzed for each year. To achieve this, we used the error grid analysis (EGA) method of Clarke et al. (12). The EGA separates a typical scatter plot into five zones of clinical significance. The presence and severity of a treatment error based on blood glucose assay being evaluated define the zones. Zone A represents the absence of treatment error, zone B represents cases where the two methods disagree by  $>20\%$  but still do not lead to a treatment error, and zones C, D, and E represent increasingly large and potentially harmful discrepancies between the evaluation and RMs.

Statistical analyses were performed using the Statview computer program (Statview V; BrainPower, Calabasas, CA).

**RESULTS**— A total of 21,950 of the 22,000 previous meters readings were analyzed because some of the measurements were indicated as "low" or "high" by the devices. The mean total pre glucose value was  $142 \pm 62$  mg/dl (range 25–502) with the HB and  $160 \pm 75$  mg/dl (range 10–668) with the RM. The mean total post value was  $200 \pm 69$  mg/dl (range 25–

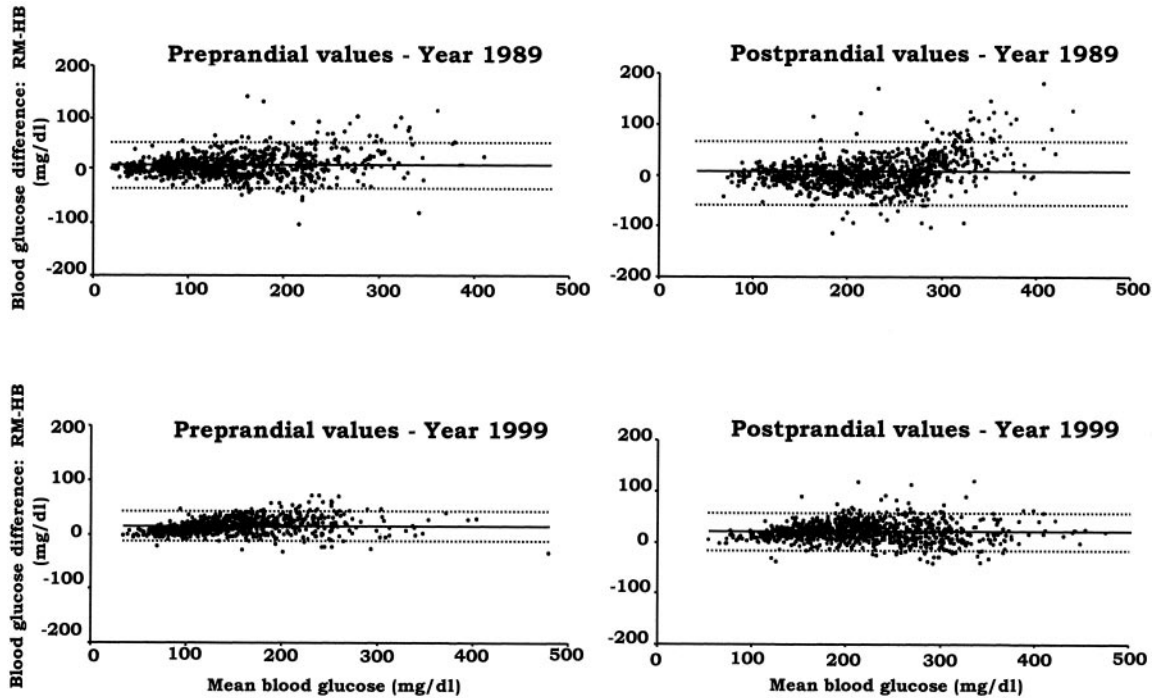
583) with the HB and  $232 \pm 82$  mg/dl (range 10–698) with the RM.

Intraclass correlation coefficients for every year were all  $>0.80$ . This coefficient reached 0.95 or higher for the following years: 1997 (0.965, 95% CI 0.964–0.972), 1998 (0.976, 0.973–0.979), and 1999 (0.973, 0.969–0.976) for the pre values; and 1997 (0.963, 0.954–0.967), 1998 (0.978, 0.975–0.980), and 1999 (0.964, 0.959–0.980) for the post values.

The differences between RM and HB are shown for each year in Table 1. The mean difference for the entire period of the study was  $19 \pm 24$  mg/dl for the pre values and  $32 \pm 30$  mg/dl for the post values. For the pre values, the highest mean difference was observed in 1993 ( $31 \pm 33$  mg/dl). In the post state, the highest mean difference was observed in 1996 ( $50 \pm 35$  mg/dl).

Our analysis exhibited considerable variations from the RM, as judged by calculations of absolute variation, with maximum deviations seen in 1990 (pre: 16.4%) and 1996 (post: 20.6%) (Table 1). Conversely, minimal absolute variations were observed in 1997 (post: 8.0%) and in 1998 (pre: 7.0%; post: 8.0%).

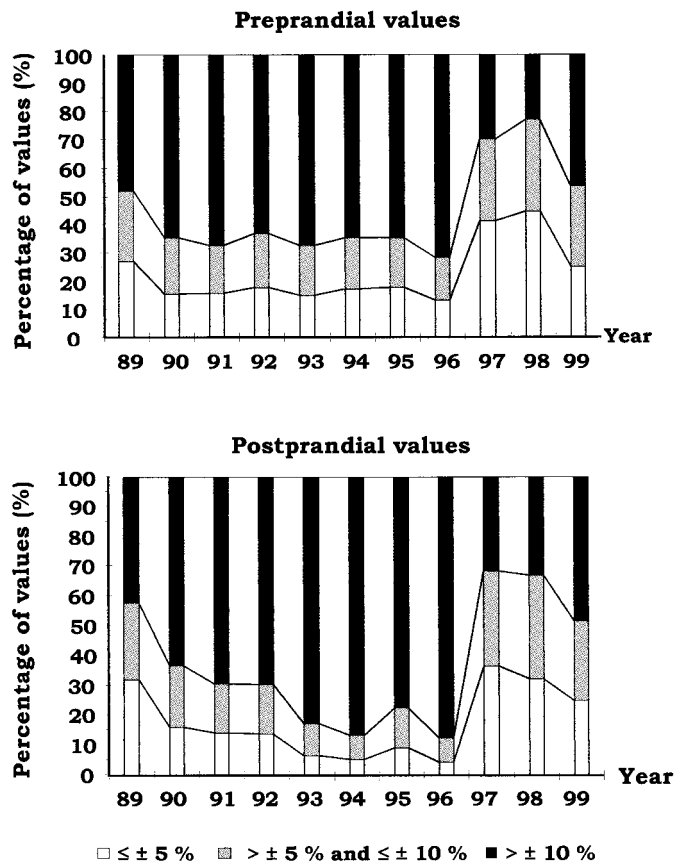
Bland and Altman (10) graphic presentations from 1989 and 1999 (pre and post values) are shown in Fig. 1. The dispersion of plots in 1989 was higher than that in 1999. This observation means that the maximum difference was higher for 1989. Conversely, the mean global difference was closer to zero in 1989, indicating the smaller mean difference between



**Figure 1**—Bland and Altman graphic presentations from 1989 and 1999 for preprandial and postprandial values. Black lines represent mean of the differences. Dashed lines represent mean  $\pm$  1.96 SD.

HB and RM reported for this year for both pre and post values (2 and 9 mg/dl, respectively) compared with the results from 1999 (15 and 20 mg/dl) (Table 1). For the other years, as well as 1989 and 1999, the dispersion of plots was not constant and increased with the glucose level (data not shown).

The percentage of pre and post values within  $\pm 5$  and  $\pm 10\%$  are shown for each year in Table 1 and Fig. 2. For pre values, the three most favorable percentages for blood glucose values within a  $\pm 5\%$  deviation from the reference values were found in 1998 (44.5%), 1997 (41.3%), and 1989 (27.0%). With respect to the  $\pm 5$  and  $\pm 10\%$  levels in the pre state, there was no significant difference between the percentage for 1989 and 1999 ( $P = 0.27$  and  $0.52$ , respectively). In the post state, the best outcome was obtained in 1997 (36.7%), followed by 1998 (32.4%) and 1989 (32.0%). The least favorable results were obtained in 1996, with 13.2% pre and 4.6% post values within the 5% deviation range. The largest number of values within  $\pm 10\%$  of the RM was observed in 1998, both for the pre (77.1%) and post values (69.0%). Nevertheless, 1996 was the year with least favorable results because only 15.3% of



**Figure 2**—Yearly percentages of values within 5 and 10% and  $>10\%$  deviation from the reference value.

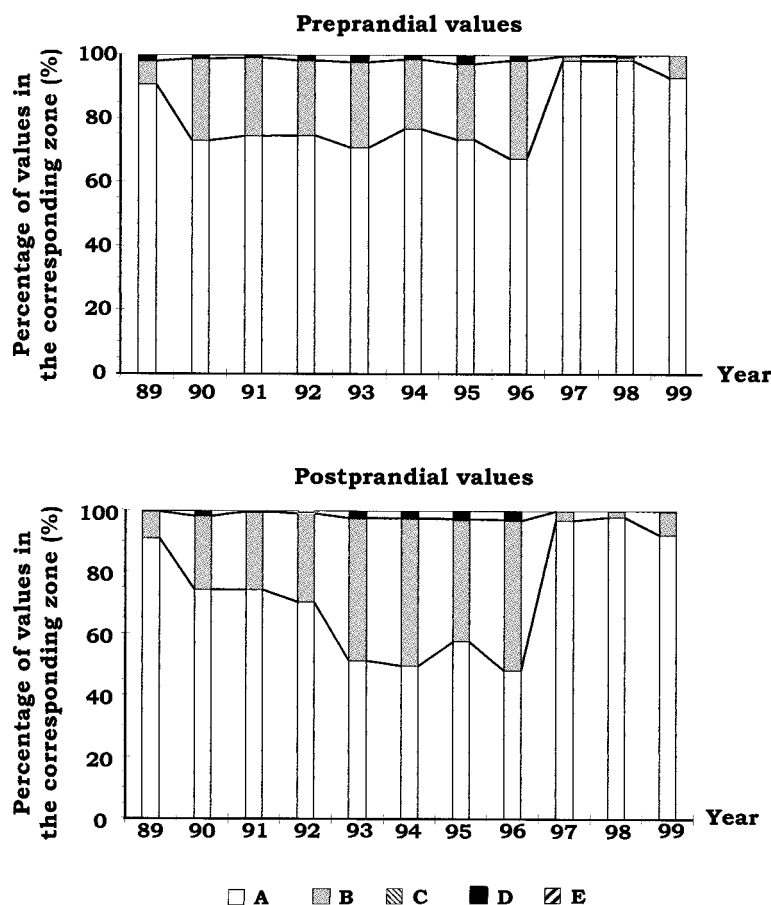


Figure 3—Yearly percentages of values in corresponding zones of the EGA.

the pre and 12.6% of the post values were within the 10% level.

For the pre values, most of the years studied showed that >70% of the measured capillary blood glucose values fell into zone A of the EGA, except in 1996 (67.5%) (Table 1 and Fig. 3). For the post values, percentages of plots in zone A were not statistically different between 1989 and 1999 ( $P = 0.32$ ). The years with least favorable results included 1993–1996. None of the values obtained from the glucose meters (except for one pre value in 1995) corresponded to zone E. Only a small number of values each year corresponded to zone D (pre: from 0.40% in 1998 to 2.71% in 1995; post: from 0.01% in 1998 to 2.89% in 1996), irrespective of the level of glycemia. Although 1989 was associated with a high percentage of pre values in zone A, 1.91% of plots fell in zone D in the pre state. Conversely, in relation to the post values, the percentage of plots in zone D did not overlap 0.40% from 1997 to 1999. The percentage of post values in zone D re-

mained <0.40% for 1989, 1997, 1998, and 1999 (Fig. 3).

**CONCLUSIONS**— About 22,000 capillary blood glucose measurements, comprising ~1,000 annual pre- and postprandial samples, were analyzed between 1989 and 1999 to compare the results obtained with glucose meters and with a time-stable reference laboratory method. We found that the accuracy of the meters decreased dramatically between 1990 and 1996 and that their performance recovered after 1997.

To assess technical accuracy, we have intentionally not used the Spearman's correlation test because the  $r$  coefficient measures the extent to which two sets of data fit a linear relationship, but not the consistency between data (5). One of the effective procedures for measuring agreement between methods is intraclass correlation (13). Moreover, Bland and Altman's (10) graphic presentation makes it easier to evaluate the accuracy of these methods and to assess the magnitude of

disagreement by examining the distribution of differences.

Estimation of the percentage of HB values that fall within  $\pm 5$  and  $\pm 10\%$  was also appropriate because it allowed the same classification, irrespective of the year and the glucose level (1). Finally, the EGA elaborated by Clarke et al. (12) was a method that directly related to patients in terms of clinical management.

Using these validated comparison methods (9,10,12,14), we showed that the performance of glucose meters from years 1997 to 1999 improved substantially compared with those from 1990 to 1996. In contrast, meter devices in 1989 were almost as accurate as those from 1997 to 1999. With the Bland and Altman approach, our data confirmed the slight superiority of 1989 and 1997 to 1999, as well as demonstrated meter inaccuracy from 1990 to 1996. Moreover, we found that accuracy was significantly better in 1989 than in 1999 with regard to the 5 and 10% deviation from the RM, especially for the preprandial values. EGA analysis showed similar results, except for the number of preprandial, but not postprandial, values that fall in zone D, indicating a slight superiority in the last 3 years of the study.

The present study is the first to report large numbers of glucose monitor readings. Brunner et al. (4) reported a study with 1,794 mean readings, but not over a long period of follow-up.

All the studies relating to accuracy and precision of home blood glucose meters were performed in a cross-sectional manner (4,8,15–18). Our results suggest that the older glucose meters from the late 1980s have similar accuracy to the newer ones. In the early 1990s, Bain et al. (19) showed that three of the main devices exhibited variations from the RM. Others showed that recent meters have better performances than the older ones (8). These authors reported that four new meters yielded nearly 50% of values that were within  $\pm 5\%$ , while two of older meters gave a worse percentage (32.5 and 33.5%) in this area. Conversely, in their study, one of the older meters was more accurate than two from the newer generation.

Devices that were tested in 1989 and purchased before that year had a more stable accuracy than those purchased in 1990. This could be an explanation for the observed decrease in analytical per-

formance. Unfortunately, we did not record all the names of the meters used in this study. Although we have no information on the date of purchase and the trade name for each device, we cannot eliminate the possibility that technical accuracy decreased between 1990 and 1996 due to the aging of meters. However, there is no longitudinal study to demonstrate that the analytical performance of meters decreased with time or frequency of use. This would be interesting with regard to the reliability of self-monitoring frequency with HbA<sub>1c</sub> improvement (2,10,20).

Our study includes other limitations. First, we compared capillary whole blood with capillary plasma using the glucose RM. It is known that capillary whole-blood glucose is lower than the venous plasma equivalent (21). Thus, during the last decade, the majority of glucose meters were calibrated based on a capillary whole-glucose technique, such as YSI (22). However, some manufacturers, like Lifescan or Abbott/Medisense, recently changed their calibration method (from capillary whole blood to capillary plasma) during the study period in France (after 1995 and 1998, respectively) and it was difficult to take this modification into account during the study. This can nevertheless explain the relative improvement noted in the last years of the study. We also did not consider biological interference, such as the hematocrit level (23,24), glucose levels (25), bilirubinemia (26), or fluorescein (27). All these factors should modify the values of glucose meters. However, the interest of the present data are that the RM remained the same throughout the period of study and the preanalytical errors were reduced to the minimum. User errors were avoided because the trained nurses of our unit conducted all the steps of meter management, including the washing and calibration of the device and glucose measurement.

Good correlation and agreement between the datasets did not permit evaluation of clinical accuracy of the data. Therefore, EGA, based on treatment goals, was performed to evaluate the usefulness and clinical accuracy of the glucose determinations (12). In this work, most of the data were in the acceptance area (A and B). However, and for each year, only a few values fell in zone D, which involved inaccurate therapeutic correction. Moreover, preprandial glycemia

led more often to false therapeutic adjustment in 1989 than in 1999 because of the frequency of values that fell in zone D. Such data indicate the only advantage of the newer devices compared with the older ones, although this was not proven for higher levels of glycemia.

Finally, results from none of the years studied met the ADA's accuracy goal of being within 5% of the laboratory reference value. This was not surprising and remains consistent with the most recent studies analyzing the performance of glucose meters (5,17,25,28).

Despite this modest but not significant improvement in the analytical area, we agree that the more recent meters allow minimization of user errors and, therefore, overall accuracy. Multiple factors that can interfere with glucose analysis, such as improper application, timing, and removal of excess blood, have been eliminated by advances in technology. Moreover, some authors showed that recent meters should reduce the total error, i.e., analytical and especially user error (7).

In summary, our data showed that the accuracy of glucose meters was significantly decreased between 1990 and 1996 but that their performance recovered and slightly improved from 1997 to 1999 in comparison to the late 1980s. This improvement remained insufficient with regard to most of the official goals of meter performance. However, since our study includes some inaccuracies, similar longitudinal surveys taking into account the technological progress, as well as the calibration method of the newer glucose meters, appear necessary.

**Acknowledgments**— This article is dedicated to Professor Pierre Drouin, deceased 21 October 2002.

#### References

1. American Diabetes Association: Clinical practice recommendations 1996. *Diabetes Care* 19 (Suppl. 1):S1–S118, 1996
2. The Diabetes Control and Complications Trial Research Group: The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 329:977–986, 1993
3. Faas A, Schellevis FG, Van Eijk JT: The efficacy of self-monitoring of blood glucose in NIDDM subjects: a criteria-based literature review. *Diabetes Care* 20:1482–1486, 1997
4. Brunner GA, Ellmerer M, Sendlhofer G, Wutte A, Trajanoski Z, Schaupp L, Quehenberger F, Wach P, Krejs GJ, Pieber TR: Validation of home blood glucose meters with respect to clinical and analytical approaches. *Diabetes Care* 21:585–590, 1998
5. Poirier JY, Le Prieur N, Campion L, Guilhem I, Allannic H, Maugeudre D: Clinical and statistical evaluation of self-monitoring blood glucose meters. *Diabetes Care* 21:1919–1924, 1998
6. Ziegler O, Guerci B, Drouin P: Blood glucose self-monitoring in 1993 [article in French]. *Diabetes Metab* 18:473–475, 1992
7. Lehmann R, Kayrooz S, Greuter H, Spinas GA: Clinical and technical evaluation of a new self-monitoring blood glucose meter: assessment of analytical and user error. *Diabetes Res Clin Pract* 53:121–128, 2001
8. Weitgasser R, Gappmayer B, Pichler M: Newer portable glucose meters: analytical improvement compared with previous generation devices? *Clin Chem* 45:1821–1825, 1999
9. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 33:159–174, 1977
10. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310, 1986
11. Charbonnel B, Boivineau C, Chopinet P, Daninos JM, Drouin P, Durain D, Guyon F, Passa P: Blood glucose self-monitoring in diabetes: recommendations of ALFEDIAM (French Language Association for the Study of Diabetes and Metabolic Diseases) [article in French]. *Diabetes Metab* 21:285–289, 1995
12. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 10:622–628, 1987
13. Landis JR, Miller ME, Davis CS, Koch GG: Some general methods for the analysis of categorical data in longitudinal studies. *Stat Med* 7:109–137, 1988
14. Koschinsky T, Dannehl K, Gries FA: New approach to technical and clinical evaluation of devices for self-monitoring of blood glucose. *Diabetes Care* 11:619–629, 1988
15. Johnson RN, Baker JR: Error detection and measurement in glucose monitors. *Clin Chim Acta* 307:61–67, 2001
16. Chan JC, Wong RY, Cheung CK, Lam P, Chow CC, Yeung VT, Kan EC, Loo KM, Mong MY, Cockram CS: Accuracy, precision and user-acceptability of self blood glucose monitoring machines. *Diabetes Res Clin Pract* 36:91–104, 1997
17. Fanghanel G, Sanchez-Reyes L, Morales

- M, Torres E, Chavira J, Sotres D, Valles V: Comparative accuracy of glucose monitors. *Arch Med Res* 29:325–329, 1998
18. Johnson RN, Baker JR: Analytical error of home glucose monitors: a comparison of 18 systems. *Ann Clin Biochem* 36:72–79, 1999
  19. Bain OF, Brown KD, Sacher RA, McPherson RA: A hospital-wide blind control program for bedside glucose meters. *Arch Pathol Lab Med* 113:1370–1375, 1989
  20. Ziegler O, Kolopp M, Got I, Genton P, Debry G, Drouin P: Reliability of self-monitoring of blood glucose by CSII-treated patients with type I diabetes. *Diabetes Care* 12:184–188, 1989
  21. Chmielewski SA: Advances and strategies for glucose monitoring. *Am J Clin Pathol* 104 (4 Suppl. 1):S59–S71, 1995
  22. Johnson RN, Baker JR: Accuracy of devices used for self-monitoring of blood glucose. *Ann Clin Biochem* 35:68–74, 1998
  23. Phillipou G, Seaborn CJ, Hooper J, Phillips PJ: Capillary blood glucose measurements in hospital inpatients using portable glucose meters. *Aust N Z J Med* 23:667–671, 1993
  24. Smith EA, Kilpatrick ES: Intra-operative blood glucose measurements: the effect of haematocrit on glucose test strips. *Anaesthesia* 49:129–132, 1994
  25. Moses R, Schier G, Matthews J, Davis W: The accuracy of home glucose meters for the glucose range anticipated in pregnancy. *Aust N Z J Obstet Gynaecol* 37:282–286, 1997
  26. Jain R, Myers TF, Kahn SE, Zeller WP: How accurate is glucose analysis in the presence of multiple interfering substances in the neonate? (glucose analysis and interfering substances). *J Clin Lab Anal* 10:13–16, 1996
  27. Antcliff RJ, Hugkulstone CE: The effect of fluorescein on the accuracy of blood glucose determination by glucose meters. *Acta Ophthalmol Scand* 77:197–199, 1999
  28. Stenger P, Allen ME, Lisius L: Accuracy of blood glucose meters in pregnant subjects with diabetes (Letter). *Diabetes Care* 19: 268–269, 1996