Closing the loop, squaring the circle: Studies on insulin delivery, glucose monitoring and the artificial pancreas
Luijf, Y.M.

Citation for published version (APA):
Luijf, Y. M. (2013). Closing the loop, squaring the circle: Studies on insulin delivery, glucose monitoring and the artificial pancreas

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 8

Accuracy and reliability of current CGM systems;  
a direct comparison

Y.M. Luijf, J.K. Mader, W. Doll, T. Pieber, J. Place, E. Renard, D. Bruttomesso, A. Filippi,  
A. Avogaro, S. Arnolds, C. Benesch, L. Heinemann and J.H. DeVries

On behalf of the AP@home consortium

Diabetes Technol Ther. 2013 May 7. (Epub ahead of print)
Objective: To assess accuracy and reliability of three CGM systems.

Research design and methods: We studied the Animas Vibe with Dexcom 4th generation version A sensor (G4A), the Abbott Freestyle Navigator I (NAV) and the Medtronic Paradigm with Enlite sensor (ENL), in 20 patients with type 1 diabetes mellitus. All systems were investigated both in the Clinical Research Center (CRC) and at home. In the CRC, patients received a meal with a delayed and increased insulin dose to induce a postprandial glucose peak and nadir. Hereafter, randomization determined which two of the three systems would be worn at home until the end-of-functioning, attempting use beyond manufacturer specified lifetime. Patients performed at least 5 reference fingersticks per day. An analysis of variance was performed on all data points ≥15 min apart.

Results: Overall Mean Absolute Average Difference (MARD) (SD) measured at the CRC was 16.5 (14.3) % for NAV and 16.4 (15.6) % for ENL, outperforming G4A [20.5 (18.2) %; P<0.001]. Overall MARD when assessed at home was 14.5 (16.7) % for NAV and 16.5 (18.8) % for G4A, outperforming ENL [18.9 (23.6) %; P=0.006]. Median time until end-of-functioning was similar: 10.0 (1.0) days for G4A, 8.0 (3.5) days for NAV and 8.0 (1.5) days for ENL (P=0.119).

Conclusions: In the CRC, G4A was less accurate than NAV and ENL sensors, which seemed comparable. However, at home, ENL was less accurate than NAV and G4A. Moreover, CGM systems often show sufficient accuracy to be used beyond manufacturer specified lifetime.
INTRODUCTION

Continuous Glucose Monitoring (CGM) systems are available for patients with diabetes for more than 10 years now. Most CGM systems assess glucose in the subcutaneous interstitial fluid employing the glucose-oxidase approach. This provides minimal discomfort to patients and allows CGM usage at home (1). While improvement in HbA1c with the use of CGM has been demonstrated (2), accuracy of glucose measurement is not yet good enough to allow patients to fully trust their CGM glucose readings, to such an extent that CGM manufacturers still recommend patients to use capillary blood glucose measurements before any treatment decisions are made (e.g. insulin dosing). The less than desirable accuracy of CGM systems also hampers endeavors to automate insulin administration by means of artificial pancreas systems (AP-systems), in which the CGM data are used by an algorithm to automatically determine and administer the appropriate amount of insulin needed to establish and maintain euglycaemia. Obviously, AP-systems require highly accurate CGM data. The limited accuracy of CGM systems is, in part, caused by the compartment in which CGM systems measure glucose. Both a delay related to the measurement technique and the existence of a lag time between changes in blood glucose and interstitial fluid glucose are challenging CGM accuracy (3). Additionally, it has been shown that accuracy of needle-type CGM systems can be poor in the hypoglycaemic range (4). Therefore, it is of utmost importance to assess the accuracy of CGM systems in a pertinent manner. So far however, there is no reference procedure to evaluate the accuracy and reliability of CGM systems that are introduced into the market. Most pre-market analyses of CGM systems have been done by means of Clarke Error Grid Analysis and assessment of correlation between CGM glucose values and reference blood glucose values collected during non standardized and/or non specified conditions (5-7).

The primary aim of this study was to assess the accuracy of the three widely used needle-type CGM systems in a way that includes both a highly standardized assessment within a clinical research center (CRC) as well as real-life usage at home. In addition, we assessed CGM accuracy within manufacturer specified lifetime (MSL) and also beyond as it is possible to “re-activate” the CGM systems by tricking the
CGM receiver into activating the old sensor as if a new sensor has been placed. It appears that patients do this routinely due to the poor reimbursement of CGM systems in many countries (8). Especially patients who pay out of their own pocket for their CGM system try to use this expensive equipment as long as possible.

METHODS

This was a multinational, randomized, open-label trial. Main inclusion criteria included a diagnosis of type 1 diabetes since at least 6 months, a body mass index (BMI) <35 kg/m² and an HbA1c level <10% (86 mmol/mol) at time of inclusion. Main exclusion criteria included pregnancy and use of medication which impacts glucose metabolism other than those used to treat their diabetes. Drugs that may impair enzymatic measurement of glucose by the sensors also had to be omitted during the investigation procedures (e.g. acetaminophen).

Twenty patients were included in this trial, four patients each in the five participating clinical centers of the AP@home consortium: Amsterdam, the Netherlands; Graz, Austria; Montpellier, France; Neuss, Germany and Padua, Italy. All patients completed an inclusion visit during which informed consent was obtained and they received training in the use of the CGM systems. The CGM systems used in this trial included the Abbott Freestyle Navigator I (NAV) (Abbott Diabetes Care, Alameda, CA, USA), the Medtronic Paradigm Veo system with Enlite sensor (ENL) (Medtronic, Northridge, CA, USA) and the Animas Vibe with Dexcom G4 version A sensor (G4A) (a collaboration between Animas Corp., West Chester, PA, USA and Dexcom Inc., San Diego, CA, USA). All systems were calibrated according to manufacturer’s specification with finger stick measurements using an Abbott Freestyle blood glucose meter (Abbott Diabetes Care, Alameda, CA, USA). This meter was chosen as the calibration meter for all three systems since NAV can only be calibrated using its built-in blood glucose meter which only accepts Freestyle strips. The other two CGM systems allow calibration against glucose values from any blood glucose meter. Patients were encouraged to avoid calibration directly after meals or strenuous activity and patients were also asked to perform self measurement of blood glucose (SMBG) at least 5 times per day (pre-/postprandial and before the night), in addition to the SMBG’s needed for (re)
calibration, which were not included in the accuracy analysis. After placement of the sensors of all three CGM systems in the abdominal region, patients left the CRC and returned the following day to enter the CRC phase of the trial (Figure 1).

![Figure 1: Study procedures during CRC admission day (adapted from Wentholt et al. (4)).](image)

Patients arrived at the CRC at 08:00 and sampling for reference blood glucose levels started immediately using the YSI 2300 STAT PLUS glucose-lactate analyzer (YSI Incorporated, Yellow Springs, OH). Patients were then served a breakfast at 08:15, but their mealtime insulin dose was deliberately delayed until 08:45, at which time a prandial insulin dose was given accompanied by an additional 5 U insulin. This procedure aimed to induce maximum excursions in postprandial glucose levels followed by a decline to low blood glucose levels. Blood was sampled every 15 min except between 08:00 and 08:15, 09:00 and 09:30 and 10:00 and 10:30 when blood was sampled every 5 min to capture baseline, peak glucose and nadir glucose levels, respectively. The study day ended at noon, at which time one CGM system was randomly selected to be removed as continuing the home phase with all three systems was considered to be too cumbersome for patients. Thus, patients would return home with two remaining CGM systems. Patients were instructed on how to reactivate the systems after MSL had ended in order to assess longevity and accuracy of the sensor readings beyond MSL. When both remaining CGM systems reached end-of-functioning, patients returned to the CRC and CGM data were downloaded. “end-of-functioning” was defined as an average mean absolute relative difference (MARD) between CGM readings and SMBG results of >25% or >20 mg/dL in the
hypoglycaemic range (<70 mg/dL) on two consecutive days. Patients were provided with instructions on how to calculate the MARD and these calculations were reviewed by study personnel. MARD calculations performed by patients were exclusively used to assess if CGM systems met the criteria which defined their end-of-functioning. Upon completion of the trial, reference glucose meter measurements were matched with the corresponding CGM results and MARD was calculated. Only CGM-reference pairs which were more than 15 min apart from each other were used for the analysis to prevent interdependency of data points (9). The maximum duration of CGM system use was recorded. Primary outcome measure was overall MARD assessed at the CRC and at home. Secondary outcome measures included MARD in the lower glycaemic (<100 mg/dL), euglycaemic (100-200 mg/dL) and hyperglycaemic (>200 mg/dL) range (according to reference measurements). MARD per day of use and duration of CGM system usage were also assessed. An analysis of variance was performed to assess differences in accuracy between CGM systems. A Kaplan-Meier analysis was performed to assess survival of CGM systems. Outcomes are given in mean (SD) or median (IQR), as appropriate. Statistical analysis was performed in PASW Statistics 18.0 (IBM Corp., Armonk, NY, USA).

RESULTS

Technical data
CRC data from three patients from the NAV CGM device could not be uploaded and was excluded from the analysis.

CRC Phase
Overall MARD (SD) at the CRC on day 1 was 16.5 (14.3) % for NAV (number of data pairs=272), 20.5 (18.2) % for G4A (n=306) and 16.4 (15.6) % for ENL (n=306) (overall P<0.001) with both NAV and ENL performing better than G4A (P<0.001). There was no difference between NAV and ENL (P=0.791). When looking at CRC data pairs in the low glycaemic range, NAV and ENL also showed lower MARD values than G4A: 17.4 (15.2) % (n=60) and 21.5 (24.0) % (n=64) versus 26.6 (25.7) % (n=64), respectively; overall P=0.005, with no significant difference between NAV and ENL (P=0.135). In
the euglycaemic range, NAV and ENL also showed lower MARD values than G4A: 18.1 (17.5) % (n=115) and 16.4 (14.2) % (n=126) vs. 21.0 (16.3) % (n=126); overall P=0.006, with no significant difference between NAV and ENL (P=0.683). MARD values in the hyperglycaemic range were not significantly different between CGM systems: 14.0 (8.2) % (n=97) for NAV, 16.6 (13.8) % (n=116) for G4A and 13.6 (9.4) % (n=116) for ENL, overall P=0.140.

**Home Phase**

Overall MARD (SD) when assessed at home was 14.5 (16.7) % during the 5 days of MSL for NAV (n=329) and 16.5 (18.8) % during the 7 days of MSL for G4A (n=462) (Figure 2). Both CGM systems performed better than ENL (18.9 (23.6) %) (n=312) during its 6 days of MSL (P=0.006). There was no difference between NAV and G4A during their respective MSL (P=0.534). Overall MARD (SD) after MSL was lower for G4A (15.6 (17.7) %) (n=1162) vs. NAV (18.9 (17.0) %; P=0.002) (n=337) and ENL (30.0 (26.0) %; P<0.001) (n=174).

During MSL there is no effect of duration of use on the accuracy of the CGM systems, with MARD not being different between days of use for NAV (P=0.898), G4A (P=0.846) and ENL (P=0.153) (Figure 2A).

G4A sensors displayed occasional extreme longevity with a maximum time until end-of-functioning of 82 days, versus 26 days for NAV and 16 days for ENL (Figure 2B). However, median time until end-of-functioning was not different with 10.0 (1.0) days for G4A, 8.0 (3.5) days for NAV and 8.0 (1.5) days for ENL (P=0.119).
Clarke Error Grid Analysis (CEGA)

The distribution of data pairs in CEGA zones per CGM system during MSL and during the entire CGM system lifetime is presented in Figure 3 and given in Table 1. During MSL there was no significant difference in distribution between the various CEGA zones (P=0.132).

Figure 2: A: MARD per day per CGM system (MARD (SEM)) B: Survival curves of CGM systems
Figure 3: Clarke Error Grid Analysis per CGM system during complete lifetime and MSL
DISCUSSION

To our knowledge, this trial is the first head-to-head comparison of these three CGM systems. In a comprehensive assessment, we looked at accuracy, assessed both under CRC and home conditions, and longevity also beyond MSL. We showed earlier that accuracy of CGM systems assessed at the CRC differs from assessment at home (10). Indeed, a difference can be appreciated with the NAV and ENL CGM systems outperforming the G4A system when assessed on day 1 of use at the CRC, while analysis of the home phase shows superior accuracy for NAV and G4A with a relative underperformance of ENL sensor. One part of the explanation of this difference is that CRC assessment of accuracy occurs during a relatively brief period of several hours with frequent of sampling of reference values, whereas the home phase allows for accuracy assessment over several days. Also, as can be seen in Figure 2, G4A has the largest MARD on day 1 of use, but from day 2 of use onwards its MARD falls in between NAV and ENL until the end of MSL, although not statistically different from the NAV sensor. It appears as if the G4A system needs a longer warm-up period. This point may have been improved in the Dexcom G4 Platinum sensor (G4 version B) which recently acquired regulatory approval in both the US and Europe.

A shortcoming of this study is the insufficient amount of data pairs in the hypoglycaemic area. We used the method described by Wentholt et al. (4) from which delaying and increasing the mealtime dose should have led to a period of minor hypoglycaemia. However, we were less successful in achieving hypoglycaemia in this trial. Future investigations following this procedure will need to use more aggressive insulin dosing in order to maximize the number of data pairs in the hypoglycaemic range.

This study also shows that sensor life can be extended beyond MSL by reactivating the sensors when the sensor session has ended. It should be noted that all current CGM systems are unable to make the distinction between a reactivated sensor and a new sensor, therefore all information about past calibration points and performance of the sensor is likely lost. This however does not negate the fact that a large proportion of CGM systems can be used far beyond MSL, thus improving their cost-efficacy. Unfortunately, there is currently no way for the patient to predict which sensors can be used for longer periods of time and patients need to actively make
their own assessment about the accuracy of each individual sensor beyond MSL, using for example the criteria set out in this trial, rather than making a subjective assessment of sensor accuracy.

In conclusion, we showed that during CRC assessment the Dexcom CGM system was less accurate than the Abbott and Medtronic systems, which seemed comparable. However, during assessment at home the Medtronic system was less accurate than the Abbott and Dexcom systems. In addition, our data showed that sensors can be used beyond manufacturer specified lifetime if accuracy is assessed by the patient on a daily base.

ACKNOWLEDGEMENTS

This research was supported by the European Community Framework Programme 7 (FP7-ICT-2009-4 grant number 247138).
REFERENCE LIST


