Nonadjunctive Use of Continuous Glucose Monitors for Insulin Dosing: Is It Safe?

Journal of Diabetes Science and Technology 2017, Vol. 11(4) 833-838 © 2017 Diabetes Technology Society Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/1932296816688303 journals.sagepub.com/home/dst



Alan R. Shapiro, MD, MSPH¹

Abstract

With the increasing accuracy of continuous glucose monitors (CGM) have come calls for the Food and Drug Administration (FDA) to label these devices as safe for nonadjunctive dosing of insulin. However, there is evidence that these devices are subject to sporadic, unpredictable, large errors. A text analysis of reports to the FDA MAUDE database since 2015 reveals over 25 000 complaints of CGM sensor inaccuracy, with instances directly leading to serious outcomes. These new data were not considered at a recent FDA Advisory Panel meeting that voted to approve Dexcom G5 relabeling for nonadjunctive use. Social media is another source of surveillance data providing evidence of large CGM inaccuracies in real-world use. We need to improve safety procedures, not remove them. CGMs offer unique information and alerts for managing diabetes, but the issue is not whether they are better than other approaches to monitoring glucose, but how they can be best used in conjunction with devices that offer the confirmatory readings needed for patient safety.

Keywords

CGM accuracy, continuous glucose monitor, MAUDE, nonadjunctive claim, SMBG, text mining

Under current labeling, a continuous glucose monitoring system (CGM) is approved only for adjunctive use in making treatment decisions; any reading from a CGM should be confirmed, typically with a self-monitoring of blood glucose (SMBG) sample of capillary blood obtained by fingerstick. As CGMs have become more accurate, there have been calls to have them relabeled as safe for nonadjunctive use in dosing insulin.¹ At a recent meeting, based on the data and arguments presented, an FDA Advisory Panel voted 8-2 in favor of a proposal that the Dexcom G5 CGM could be considered safe for use in insulin dosing without a requirement for backup confirmation.²

There is, however, considerable documented and anecdotal evidence, not discussed at the meeting, that CGMs can frequently provide users with inaccurate readings. Much discussion at the FDA panel meeting addressed whether the G5 CGM is as good as or better than SMBGs for dosing insulin. But the question whether a CGM is on average better than an SMBG does not address whether a CGM, if subject to sporadic large errors, should be regarded as reliably safe to use without confirmation.

Test strip-based blood glucose meters were never formally approved for nonadjunctive insulin dosing, but in this case, the Dexcom G5 is specifically seeking that approval. The hardware and software components affecting accuracy are considered identical in the most recent G4 Platinum and G5 devices. Accordingly, this article examines data specifically on Dexcom G4/G5 accuracy. However, there is no

attempt to compare the accuracy of this CGM with that of any other CGM or SMBG device, and no such comparative inferences should be drawn from the data presented.

Case Reports of Device Inaccuracy From the FDA MAUDE Database

The FDA Manufacturer and User Facility Device Experience (MAUDE) was established as a surveillance tool for monitoring case reports of problems and safety issues with devices. Although MAUDE has been criticized for containing many incomplete, inconclusive, or even irrelevant cases,^{3,4} it can nevertheless serve as a valuable source of case reports that, in turn, can act as sentinel events in detecting potential device problems. Surveillance systems typically seek to balance the requirement for high sensitivity with the need for high data quality. The less than optimal data quality frequently found in MAUDE means that consideration of the biological plausibility, credibility, and relevance of each report needs to be weighed with the consistency and volume of reports in deciding which signals should be prioritized and rigorously evaluated. To this end, in addition to structured data, the database

¹Department of Medicine, NYU School of Medicine, New York, NY, USA

Corresponding Author:

Alan R. Shapiro, MD, MSPH, 5 Pheasant Run Rd, Pleasantville, NY 10570, USA.

Email: alan.shapiro@med.nyu.edu

contains unstructured text narrative which has the potential for providing details of adverse events and of manufacturers' responses to these events. This text is sufficiently stylized, and employs a sufficiently small vocabulary, that it can be approached as a sublanguage amenable to automated qualitative analysis.⁵

Even a few sentinel events have the potential to reveal important problems. For example, on the basis of just 2 customer complaints, a problem was recently discovered which proactively led to the recall of several lots of a glucagon injection system.⁶ In contrast, examination of the text of complaints in the MAUDE database reveals that there have been 25 966 complaints specifically about G4/G5 inaccuracy that were submitted from the beginning of 2015 through September 2016. These complaints were not considered at the July 21 FDA Advisory Panel meeting.²

MAUDE records, unfortunately, are often incomplete. Of the 25 966 complaints of inaccuracy, 3290 (12%) did not directly provide additional information other than indicating that a user had specifically complained to the manufacturer that the G4 or G5 device was inaccurate and did not report a more serious adverse event. Another 53 apparently duplicated reports were identified. This still left 22 623 complaints that did provide additional information.

In over a third of these remaining complaints, the inaccuracy was explained away as being due to user error, that is, users did not follow the instructions specifically stated in the User's Guide. Analysis of the text of the manufacturer's comments in these cases reveals 16 types of CGM user errors that were viewed by the manufacturer as contributing to inaccurate values in real-world cases (Table 1). When comparing CGMs with SMBG readings, it is often noted that SMBG readings can be subject to error due to inadequate user cleansing of the fingerstick site,⁷ as well as a variety of other circumstances.⁸ Currently, these SMBG readings are used to calibrate CGMs, creating additional sources of potential CGM inaccuracy, which then can be subject to CGM user errors as well.

Some users were admonished for relying on the CGM device for making insulin dosing decisions, as in the following:

Complaint narrative: "Patient contacted Dexcom technical support on mm/dd/2015 to report continuous glucose monitoring (CGM) inaccuracies compared to blood glucose (BG) meter and a hypoglycemic event that occurred on mm/dd/2015.... Prior to eating dinner, patient administered herself insulin and CGM displayed a value of 115 mg/dl. An hour after eating, the CGM alerted her of a high BG level, displaying 359 mg/dl. An additional value from a separate BG meter was not taken at this time for comparison. The patient and her husband treated the high BG level with . . . insulin. Patient's husband then tested the patient's blood glucose via a fingerstick, which read 109 mg/dl. The CGM device displayed 314 mg/dl. The patient's husband drove her to urgent care and upon arrival her BG level was tested by a nurse; the value was 42 mg/dl."

Table I. Causes of Inaccurate CGM Readings Attributed to User Errors.

- I. Use in patients less than 2 years old
- 2. Insertion of sensor in unapproved site: arms, buttocks, back, thigh
- 3. Pediatric patient using a receiver approved for adults; and vice versa
- 4. Use in pregnancy
- 5. Using transmitter past expiration date
- 6. Using sensor past expiration date
- 7. Use of device while taking medications containing acetaminophen
- 8. Improper insertion of transmitter.
- 9. Calibration while error symbol is displayed
- Calibration during a rapid change in glucose, ie, rise or fall >2 mg/dl/minute
- Calibration using a meter to calibrate that is different from the meter used to measure glucose
- Calibration using of alternative blood glucose testing sites other than fingertips
- Delay of greater than 5 minutes in entering calibration values
- 14. Entering calibration values outside the 40-400 mg/dl range
- 15. Calibrating less often that every 12 hours
- 16. NOT calibrating after experiencing inaccuracies

Source: MAUDE Database.

Manufacturer response: "It was reported that the patient based treatment decisions off the CGM values. The G4 platinum continuous glucose monitoring system user's guide states: do not use the Dexcom G4 platinum system for treatment decisions, such as how much insulin you should take. The Dexcom G4 platinum system does not replace a blood glucose meter. Always use the values from your blood glucose meter for treatment decisions. Blood glucose values may differ from sensor glucose readings. Using the sensor glucose readings for treatment decisions could lead to low or high blood glucose value. Further, diabetes mellitus is a known cause of hypoglycemia."

This is effectively the same device, with the same sensors and same algorithms, as is now being considered for nonadjunctive use. The difference would be that the current cautions about relying on device readings for making treatment decisions would be removed.

Magnitude of Inaccuracies

Most cases records in the MAUDE database did not directly state, or provide sufficient detail to calculate, the size of the inaccuracy being reported. However, because the %20/20 discrepancies between CGM and SMBG values proposed in the International Organization for Standardization (ISO) 15197:2003 standards have been used at Dexcom for determining reportable inaccuracies, presumably the discrepancies reported to the MAUDE database were those greater than the 20% limits.



Figure 1. Cumulative proportions of the size of reported differences between concurrent CGM and BGM readings. 55% of reported differences were 100 mg/dl or more. Data points show the magnitudes of the reported discrepancies, many with multiple occurrences, plotted against their cumulative proportions.

In 415 complaints, there was sufficient detail to examine the size of the reported inaccuracies. Figure 1 provides a descriptive summary of the magnitude of the discrepancies between CGM and patient or hospital blood glucose meter. Examination of the cumulative percentages of the discrepancies reveals that over 95% were equal to or greater than 30 mg/dl, and 55% were equal to or greater than 100 mg/dl. Dosing, or not dosing, off of discrepancies of this size can, and did, lead to serious outcomes. In most of the complaints of inaccuracy, users noted a large discrepancy between their CGM reading and either their confirmatory fingerstick reading or with how they were feeling. They contacted the manufacturer to inquire whether their device was faulty, and/or to express their concern that such a large discrepancy could have potentially been dangerous. Aware of the discrepancy, they did not make treatment decisions based solely on the readings and did not report injury as a result of the inaccuracy. But there were more than **Complaint narrative:** "Patient contacted Dexcom on mm/ dd/2015 to report continuous glucose monitoring (CGM) inaccuracies and an adverse event that occurred on mm/dd/ 2015. Patient stated that the CGM device was reading 255 mg/ dl, with double arrows indicating rising glucose, and she treated herself with a double dose of insulin. Patient did not confirm the CGM reading against a BG meter prior to giving herself the medication. Patient then took another reading and the CGM displayed 254 mg/dl while the BG meter displayed 70 mg/dl. Patient lost consciousness 15 minutes after medicating with insulin. Patient's sister found her and administered glucagon."

Complaint narrative: "Patient's mother contacted Dexcom technical support on mm/dd/2015 to report continuous glucose monitoring (CGM) inaccuracies compared to blood glucose (BG) meter and diabetic ketoacidosis... Patient was not feeling well and was vomiting. Patient's mother saw that the CGM was displaying 170 mg/dl and a fingerstick was taken, which read 350 mg/dl. Patient's mother drove the patient to the hospital, where they were admitted into the intensive care unit (ICU), due to hyperglycemia that resulted in the patient going into diabetic ketoacidosis."

Complaint narrative: "A friend of the patient's mother reported that the patient experienced inaccurate blood glucose values; the CGM device was reading 80 mg/dl but the patient's blood glucose level was reportedly much lower than this. It was further stated that the patient was treated with glucagon but the medication did not reach the patient in time and the patient passed away."

Manufacturer response: "It remains unknown to Dexcom whether the official cause of death was determined to be hypoglycemia. No additional patient or event information was reported. Diabetes mellitus is a known cause of hypoglycemia and death."

Nontraditional Real-World Evidence of Device Inaccuracy

Not everyone communicates with the FDA or the manufacturer when they have a problem with a device. Social media can be another source of surveillance data.^{9,10} After the July 21 Advisory Panel vote was announced, while some patients with diabetes applauded the decision, there were others who protested along the lines of "This is great but we have had many numbers on the CGM not even close to his actual finger stick reading. . . . I love the CGM but I truly don't feel comfortable dosing insulin on my child yet"¹¹ or "As an endocrinologist and type 1 diabetes patient myself, while I absolutely think Dexcom is a very successful tool in diabetes management, by no means do I think it accurate enough to replace fingersticks when making treatment decision on insulin dosing."¹² Some individuals wrote blogs¹³ and some posted pictures of CGM-SMBG discrepancies.¹⁴

At the public hearing at the July 21 FDA panel meeting, a medical device specialist showed 39 representative slides from CGM users spontaneously providing alerts, complaints, or inquiries to other users on social media sites about their experiences with potentially catastrophic intermittent large CGM inaccuracies. The patient representative to the meeting volunteered "Frankly I have quite a bit of inaccuracy with the device from time to time. I don't understand why, and I'd love to be able to get some guidance as to why that's happening."²

Taken together, these diverse surveillance sources are providing evidence that a non-negligible proportion of CGM users have been confronting erroneous values, and have been concerned about the potentially catastrophic results had they used these values without confirmation for insulin dosing.

Implications for Measures of Device Accuracy

To the extent that a blood glucose measuring device is subject to unpredictable, sporadic, and substantial errors, the most common measures of accuracy will not provide an adequate summary of device performance.

The mean absolute relative deviation (MARD) provides a concise summary of average device error and has been used extensively to evaluate the accuracy of CGM and SMBG devices. However in a situation where there is a suspicion of intermittent large errors, there is a need to focus on the large errors that occur sporadically using the device, not to hide these events in an overall summary average.

The Dexcom G5 is said to have a MARD less than 10% which has been used as evidence for the suitability of its nonadjunctive use for insulin dosing. In the present case, the concern is that there is a subset of bad sensors or, circumstances, or, perhaps, a set of patient characteristics causing CGMs to be less accurate. What matters is getting a clear picture of the right tail of a skewed or even multimodal distribution of errors, more than masking these errors in an overall average. One can easily imagine a sensor with a slightly larger MARD being preferable to one that was usually slightly more accurate but occasionally had catastrophic errors.

Another often used way of summarizing accuracy is to describe the frequency with which errors of various sizes occur. One can describe the distribution of errors in statements such as "errors greater than 30% occur only 2% of the time." The clinical data submitted to portray the accuracy of the Dexcom device in adults came from a study of only 51

patients, 44 with type 1 diabetes and 7 with type 2 diabetes, observed for one day in a controlled setting.¹⁵ A separate study of 79 patients was used to characterize the decreased accuracy in pediatric patients. Such small sample sizes might not even detect one instance of an infrequently occurring large error; they are clearly far too small for exploring whether there is a subset of patients, sensors, or conditions that lead sporadically to inaccurate CGM measurements.

But in fact, averaging errors or even looking at the frequency of errors of different sizes may be the wrong model. Given the potential consequences of an error, each major error would be better treated as an information-rich sentinel event requiring a substantive explanation, not treated as just another data point. Rather than being treated as mere outliers and further obscured by summarizing with, for example, a median absolute relative deviation (medARD), these less frequent but important cases should be investigated with the goal of exploring how future such errors can be anticipated or prevented.

As the sample of real-world complaints taken from MAUDE illustrate, the consequences of insulin dosing decisions based on inaccurate values can be disruptive or even catastrophic. Documenting that the G5 is accurate enough 95% or even 99% of the time is akin to an airline assuring its passengers that its steering mechanism is accurate 99% of the time. In the end, the precise frequency of large errors may not matter as much as the recognition that all large errors are unacceptable in insulin dosing.

Evidence for the Elimination of Safety Checks in Insulin Dosing

Given the evidence that the G4/G5 is subject to infrequent large inaccuracies that can lead to insulin dosing errors, one might ask what new clinical evidence was used to support the case for nonadjunctive labeling for the device. Remarkably, the only clinical data provided were from the same small studies previously used in gaining approval for adjunctive use of the device.

Instead of clinical evidence, data from these studies were used to provide an error distribution for the CGM in simulation studies. But it is precisely the as yet unknown frequency and magnitude of inaccuracies in the device that is in question. Once a unimodal error distribution with a 9% MARD for device errors is accepted as a valid description of reality, the answer to questions about the simulated accuracy of the device have already been determined.

The two simulation studies offered were designed to compare the performance of CGMs with a representative SMBG, the Bayer Contour Next. The implied logic was that since the simulated CGM was found, in most circumstances, to be at least as good as, or better than, the simulated SMBG device, and since SMBG devices are currently being used for making diabetes treatment decisions, then, logically, CGMs should also be labeled as usable for making treatment decisions without the need for backup confirmation. But, as mentioned earlier, whether CGMs are better, or provide more information than SMBGs was not the issue posed to the Panel or the one facing the FDA. The issue at hand is not which device is better. Rather, the question is, "Given the considerable evidence of erratic readings in some CGM sensors, can the device be deemed safe to use without the need for a confirmatory safety check?"

The value of maintaining safety checks in the blood glucose testing process can be illustrated without complex simulations. In the study of 51 adult patients used in the simulation studies, only 2% of the readings (in which YSI reference values were greater than 80 mg/dl) were more than 30% in error. Although 2% is a low frequency, a user who consulted a CGM for making treatment decisions without using a safety check say, 6 times a day, would have only a $(1-.02)^{6*30} = 2.6\%$ chance per month of NOT blindly making a decision on an unacceptably erroneous value.

Contrast the numbers if, say, an SMBG is used as confirmatory backup. The chances of both devices being wrong simultaneously would be far smaller than using either alone. If the second device is within \pm 30%, say, 99% of the time, a user would now have a far better(1-(.02*.01))^{6*30} = 96.4% chance per month of NOT making a decision unknowingly based on a wrong value.

In the absence of confirmatory testing, the primary protection from bad CGM data values comes from the proposed instruction "Use a BG meter any time symptoms or expectations do not match sensor glucose readings." Unfortunately, the persons with diabetes who are most in need of CGMs are the roughly 20-30% of patients with reduced glycemic awareness, the ones for whom this advice is least helpful. This counseling may also be insufficient for many pediatric users, or for those following the CGM results of others wirelessly.

It is conceivable that the individuals who have intact glycemic awareness and are able to successfully follow this advice are the ones best able to successfully negotiate the hazards of occasional erroneous CGM values. A recently completed clinical trial, REPLACE-BG,¹⁶ compared the safety of nonadjunctive CGM use with adjunctive CGM USE. It is hoped that the forthcoming results will demonstrate the benefits potentially obtainable from the alerts, trends, and detailed information provided by CGMs. But the study was limited to patients without significant hypoglycemic unawareness so may still leave open the question of safety in the group of patients most in need of the device, and simultaneously most vulnerable to any device inaccuracies.

Blood glucose meters are at the moment the best confirmatory devices available, but fingersticks are notoriously uncomfortable and inconvenient. Other currently investigated approaches that might serve to provide CGMs with a confirmatory backup reading include CGMs with dual sensors,¹⁷ various optical methods, saliva glucose tests, flash glucose monitors and many others.¹⁸ Castle et al¹⁹ provide empirical evidence for the accuracy benefits of using multiple sensors. Use of such approaches in coordination with CGMs would be a vastly safer path to follow than labeling an intermittently erroneous device safe for insulin dosing.

Conclusion

Is it just a subset of CGM users who are subject to aberrant results? Currently, there are no data to answer this question. How frequent are these errors in the real-world population of users? The MAUDE database suggests the problem is remarkably frequent.

Until the causes of these sporadic, large errors have been identified, it would be inappropriate to urge users to regard the device as safe for insulin dosing without confirmation. Safety checks are a standard procedure for risk reduction in many fields. Anesthesiologists, for example, use hemodynamic monitoring on each case even though hemodynamic crises typically do not occur. We should be exploring ways to improve safety procedures, not remove them. CGMs offer unique information and alerts for managing diabetes, but the issue is not whether they are better than other approaches to monitoring glucose, but how they can be best used in conjunction with devices that offer the confirmatory readings needed for patient safety.

Abbreviations

CGM, continuous glucose monitor; FDA, Food and Drug Administration; ISO, International Organization for Standardization; MARD, mean absolute relative deviation; MAUDE, Manufacturer and User Facility Device Experience; medARD, median absolute relative deviation; SMBG, self-monitoring of blood glucose.

Acknowledgment

The author would like to thank Marion Greif for her editorial assistance.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Edelman SV. Regulation catches up to reality: nonadjunctive use of continuous glucose monitoring data. J Diabetes Sci Technol. 2017;11(1):160-164.
- FoodandDrugAdministration.July21,2016MeetingoftheClinical Chemistry and Clinical Toxicology Devices Panel of the Medical Devices Advisory Committee. Available at: http://www.fda.gov/ downloads/AdvisoryCommittees/CommitteesMeetingMaterials/ MedicalDevices/MedicalDevicesAdvisoryCommittee/

ClinicalChemistryandClinicalToxicologyDevicesPanel/ UCM517251.pdf. Accessed September 5, 2016.

- Rajan PV, Kramer DB, Kesselheim AS. Medical device postapproval safety monitoring where does the United States stand? *Circ Cardiovasc Qual Outcomes*. 2015;8(1):124-131.
- Greenfield LJ. Perspective from an academic on postmarket surveillance. In: Brown SL, Bright RA, Tavris DR, eds. *Medical Device Epidemiology and Surveillance*. New York, NY: John Wiley; 2007:159-169.
- Grishman R, Kittredge R. Analyzing Language in Restricted Domains: Sublanguage Description and Processing. New York, NY: Psychology Press; 2014.
- 6. US Food and Drug Administration. Novo Nordisk Inc. issues voluntary nationwide recall of six batches of GlucaGen® HypoKit® (glucagon [rDNA origin] for injection) due to detached needles on the syringe in the kit. Available at: http://www.fda.gov/ Safety/Recalls/ucm519872.htm. Accessed September 16, 2016.
- Lima J, Mesquita D, Santos K, et al. The impact of not washing hands on the result of capillary glycemia. *J Clin Mol Endocrinol*. 2016;1(1:2). Available at: http://clinical-and-molecular-endocrinology.imedpub.com/.
- Erbach M, Freckmann G, Hinzmann R, et al. Interferences and limitations in blood glucose self-testing: an overview of the current knowledge. *J Diabetes Sci Technol*. 2016;10(5):1161-1168.
- Leaman R, Wojtulewicz L, Sullivan R, Skariah A, Yang J, Gonzalez G. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In: *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing 2010 Jul 15*. New York, NY: Association for Computational Linguistics; 2010:117-125.
- Sarker A, Ginn R, Nikfarjam A, et al. Utilizing social media data for pharmacovigilance: A review. *J Biomed Inform*. 2015;54:202-212.
- https://www.facebook.com/diaTribeNews/posts/ 10154383772647103#. Accessed August 18, 2016.
- http://www.medscape.com/viewarticle/866492#vp_2. Accessed September 16, 2016
- https://verylightnosugar.com/2016/07/18/whole-milk-withthe-disclaimers/. Accessed August 18, 2016
- Goffe T. Dosing with Dexcom: a bad idea (for now). Available at: http://www.thomasgoffe.com/uploads/2/9/6/6/ 29665655/1468468685.png. Accessed August 18, 2016.
- Bailey TS, Chang A, Christiansen M. Clinical accuracy of a continuous glucose monitoring system with an advanced algorithm. *J Diabetes Sci Technol*. 2015;9(2):209-214.
- https://clinicaltrials.gov/ct2/show/NCT02258373?term= REPLACE-BG&rank=1.
- Sharifi A, Varsavsky A, Ulloa J, et al. Redundancy in glucose sensing enhanced accuracy and reliability of an electrochemical redundant sensor for continuous glucose monitoring. J Diabetes Sci Technol. 2016;10(3):669-678.
- Nery E, Kundys M, Jeleń P, Jönsson-Niedziółka M. Electrochemical Glucose Sensing: Is There Still Room for Improvement? Anal. *Chem.*, 2016;88(23): 11271–11282
- Castle JR, Pitts A, Hanavan K, et al. The accuracy benefit of multiple amperometric glucose sensors in people with type 1 diabetes. *Diabetes Care*. 2012;35:706-710.