An Analysis of 2019 FDA Adverse Events for Two Insulin Pumps and Two Continuous Glucose Monitors

Jan S. Krouwer, PhD[1] iD

Abstract

US Food and Drug Administration adverse event data for 2019 were analyzed for two insulin pumps and two continuous glucose monitors (CGMs). The analyses were selective—they were guided by the text described in the adverse events. They included (1) percent using auto mode for the Medtronic 670G pump, (2) distributions of hyper and hypo glucose values for Medtronic and Tandem pumps, (3) a Parkes error grid for Dexcom CGM vs glucose meter when the complaint was inaccuracy, and (4) the most frequent events for Abbott Freestyle. We found that for the 670G pump, there were more hypo events when auto mode was on than when auto mode was off. With Dexcom CGMs, users complained about inaccurate result when most results were in the B zone. With the Abbott Freestyle, the most frequent adverse event was an allergic skin reaction.

Keywords

continuous glucose monitor, adverse event, MAUDE, insulin pump, Parkes glucose meter error grid

Introduction

The US Food and Drug Administration (FDA) Manufacturer and User Facility Device Experience (MAUDE) adverse event database is a useful but unfortunately underutilized source of performance data for diabetes devices. This study shows how it can be used when analyzing adverse events from 2019 for two insulin pumps and two continuous glucose monitors (CGMs). The analysis was selective—it was guided by the text described in the adverse events and was not intended to be comprehensive.

The analysis included (but was not limited to) the following:

- For the Medtronic 670G insulin pump, we determined the proportion of pump users using auto mode and categorized the adverse events that occurred when auto mode was enabled vs when auto mode was disabled.
- For the Tandem t:slim X2 Basal-IQ pump, we determined the distribution of hyperglycemic (hyper) and hypoglycemic (hypo) glucose values and the most frequent complaints.
- For the Dexcom CGMs, for the events where inaccuracy was the complaint, we graphed the results of CGM vs blood glucose meter (BGM) in a Parkes error grid.
- For the Abbott Freestyle, we determined the most frequent complaints.

These analyses illustrate examples of the information that one can find in the adverse event database. They are not intended to compare performance among devices since there is no rate information, nor are they intended to judge the performance of the devices.

Limitations

Unlike events from controlled studies, adverse events are from real-world use but are *unverified*. Moreover, there is no way to determine adverse event rates since device usage is not public and one does not know whether all events that should be reported have been reported or whether events have been properly classified. Complaint causes are rarely discovered as most devices are not returned to the manufacturer. There was no information about the time between CGM and BGM results in the method comparison. Also, the Parkes error grid used for the method comparison was not published for that purpose.

[1]Krouwer Consulting, Sherborn, MA, USA

Corresponding Author:
Jan S. Krouwer, PhD, Krouwer Consulting, 26 Parks Drive, Sherborn, MA 01770, USA.
Email: jan.krouwer@comcast.net

Methods

Each adverse event is a series of records starting with a phone call complaint summarized by the manufacturer, followed by one or more records describing the manufacturer's response to the complaint.

The MAUDE database has been previously described.[1] Text files were downloaded from the FDA website.[2] Using Microsoft Access software and structured query language (SQL) queries, I created a dataset of Medtronic 670G pumps and a dataset of Tandem t:slim X2 Basal-IQ pumps (the Control-IQ Tandem pumps were not available in 2019).

All datasets contained the fields: (1) MDR_REPORT_KEY, a unique number that was used to join tables and to identify events, (2) EVENT_TYPE (M=malfunction, IN=injury, or D=death), (3) BRAND_NAME, the brand, (4) MANUFACTURER_D_NAME, the manufacturer, (5) REPORT_SOURCE_CODE (user or manufacturer), and (6) FOI_TEXT (freedom of information), which was the text associated with the complaint or the manufacturer's response.

The final 670G dataset ($N=5823$) was obtained by querying the FOI_TEXT field for the word "AUTO." The word "AUTO" appeared because Medtronic personnel asked about the auto mode status of the pump during phone calls from pump user's complaints. Thus, this dataset contains adverse events for pumps capable of a closed-loop system. Both pump datasets were exported to Excel.

For the 670G pump data we created a uniformly distributed 10% sample of the Excel records by generating a list of sorted MDR_REPORT_KEYs. For each event in this sample, we recorded the event code, whether auto mode was being used (Y=yes, N=no, U=unknown), if the user had been hospitalized or received emergency medical treatment (Y=yes, N=no), and the blood glucose result (H=high, L=low, N=normal) and blood glucose value if available. High blood glucose values were values over 180 mg/dL, and low levels were values below 70 mg/dL. Writing Visual Basic for Applications (vba) programs helped with the analysis of the Excel data.

Similarly, we narrowed the dataset of CGMs to a dataset of Dexcom CGMs (G4, G5, and G6) and to a dataset of Abbott Freestyle CGMs by querying the field BRAND_NAME. For the Dexcom dataset, the final dataset ($N=7707$) was obtained by querying the FOI_TEXT field for both terms "INACCURACY" and "METER READING." The word "INACCURACY" appeared as part of the complaint during the phone call. The word "METER READING" indicated that CGM and glucose meter readings might be available. The dataset was then exported to Excel. A vba program gathered Dexcom CGM vs BGM data ($N=5436$). An R routine was used to prepare plots and to calculate points in Parkes error grid zones.[3]

To make the data analysis more manageable for this study, we divided the analysis into two parts. For the Medtronic 670G pump, we examined a 10% sample of records. For the Tandem pump, Dexcom, and Abbott CGMs, we wrote vba programs to query each dataset.

Results—Medtronic 670G

Table 1 shows that the auto mode was more often not in use, the pump users were not often hospitalized, there were more hyper (hyperglycemic) than hypo (hypoglycemic) results, and most events were classified as injury. When auto mode status was known (85%), 55% of users were in manual mode and 44% of users were in auto mode.

Figure 1 (left) shows a box plot of hyper (high), hypo (low), and normal glucose results (when glucose results were reported, which was 97% of the time). The median high result was 460 mg/dL with the median low result 46 mg/dL.

Table 2 shows that when auto mode was on, there were more hypo events than when auto mode was off.

Results—Tandem T:slim X2 Basal-IQ

Figure 1 (right) shows a box plot of hyper (high), hypo (low), and normal glucose results (when glucose results were reported, which was 2% of the time). The median high result was 380 mg/dL with the median low result 45 mg/dL.

Table 3 shows the most frequent complaints for the Tandem pump.

Results—Dexcom G4, G5, and G6

Figure 2 shows a Parkes error grid plot of the method comparison values where Y=CGM and X=BGM.

Table 4 shows the number and percentage in each error grid zone.

Results—Abbott Freestyle

Table 5 shows the most frequent complaints for the Abbott Freestyle. The complaints about the "BUILT IN METER" refer to the Abbott CGM receiver, which besides displaying CGM values is also a glucose meter.

Discussion

The adverse event database presents a unique opportunity to explore a dataset based on real users that would normally not be available since manufacturer interactions with users are generally not public. The number of adverse event text descriptions for diabetes devices in 2019 was almost one million records (974 866)! The analysis in this study was selective rather than comprehensive.

The purpose in looking at adverse event data was not to imply that these devices are unsuitable for use but rather to

Table 1. Percentages for Various 670G Parameters.

	Yes	No	Unknown
Auto mode on?	38	49	14
Hospitalized?	19	81	
	Hyper	Hypo	Normal
Glucose result	63	34	2
	Malfunction	Injury	Death
Event code	9	92	0
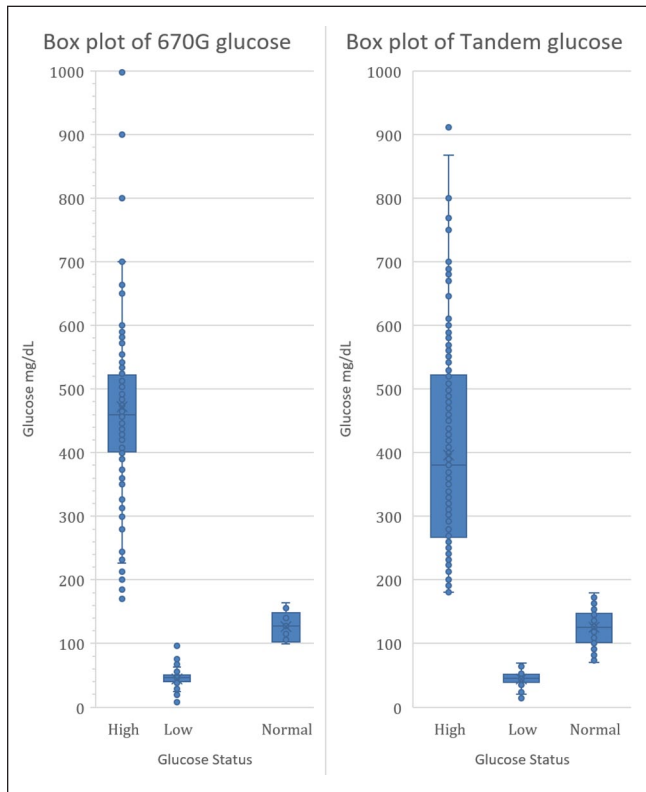
Figure 1. Box plots of 670G and Tandem glucose values.

**Table 2.** Percent Hyper and Hypo Events as a Function of Auto Mode Status.

Auto mode	Hyper	Hypo	Difference	Difference lower 95% CI	Difference upper 95% CI
Off	73	27	45.2	44.7	45.8
On	59	41	17.5	16.7	18.4

see if the adverse event data provided insights into device malfunctions. User error can never be discounted as a malfunction cause and according to the medical device reporting regulation,[4] user error must be reported as a *device malfunction*.

Table 3. Errors for the Tandem Pump System.

Problem	N	Percent
Malfunction alarm	3809	28.5
(CGM) Error 42	3575	26.8
Low blood glucose	1499	11.2
Occlusion alarm occurred	1424	10.7
Insulin gauge was inaccurate	1008	7.6
Cartridge change error	599	4.5
Altitude alarm	518	3.9
Pump touch screen was unresponsive	359	2.7
Wake button was unresponsive	238	1.8
Cartridge leaked	205	1.5
High blood glucose	116	0.9

CGM, continuous glucose monitor.
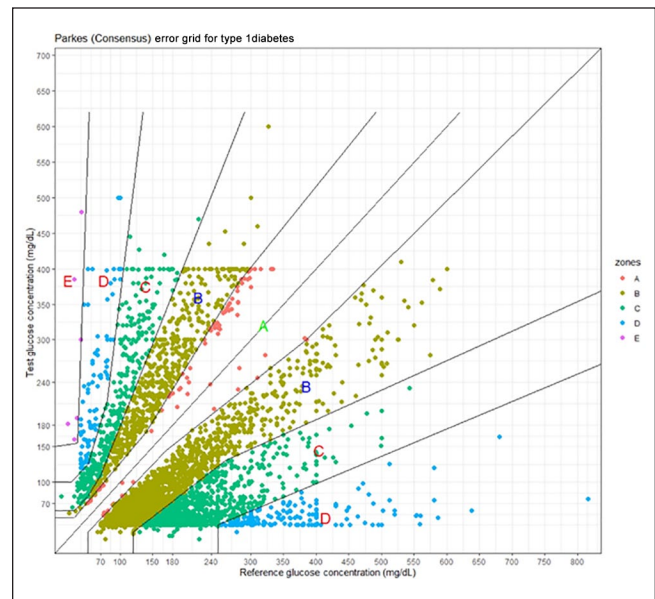
Figure 2. Parkes error grid for Dexcom continuous glucose monitors. Y = Dexcom; X = blood glucose meter.

Table 4. Number and Percent of Dexcom Continuous Glucose Monitor Glucose Values in a Parkes Error Grid.

Zone	N	Percent
A	86	1.6
B	3328	61.2
C	1741	32.0
D	273	5.0
E	7	0.1

Medtronic 670G

Hybrid closed-loop systems offer users the ability to have their basal insulin rates automatically regulated and adjusted every five minutes. By enabling "auto mode" on their

Table 5. Complaints for the Abbott Freestyle.

Complaint	N	Percent
Allergic skin reaction	2325	54.8
Hypoglycemia	531	12.5
Infection	461	10.9
Hyperglycemia	360	8.5
Replace sensor	196	4.6
Erratic + Built-in meter	136	3.2
Erratic + Freestyle Libre Reader	128	3.0
Scan again in 10 minutes	71	1.7
Sensor error	31	0.7

devices, patients with diabetes can automatically have the correct amount of insulin administered as needed.[5] Despite the convenience of this feature it was recently reported that about 30% of a sample of Medtronic 670G pump users turned off auto mode.[6-8]

The source of the glucose values in FOI_TEXT was not provided. Possible sources are the Guardian 3 sensor, the Contour glucose meter, or a hospital lab glucose assay. The Guardian 3 sensor has a range of 40-400 mg/dL. Of the 491 glucose values provided, 69 were either exactly at 40 or 400 mg/dL (69/491 = 14%). Thus, we suspect that these values came from the Guardian 3 sensor. This means that some glucose data are censored with some values likely higher than 400 mg/dL and lower than 40 mg/dL. The Contour range is 20-600 mg/dL. Thus, values at 600 mg/dL may also be censored. Values above 600 mg/dL or lower than 20 mg/dL are likely from a hospital lab glucose assay.

Our dataset is a biased sample with respect to whether auto mode was on or off because our sample came from adverse event records rather than from a random sample of Medtronic 670G pump users. Moreover, auto mode will not be active for most pump errors.[9]

The percentage of auto mode being off vs on could be influenced by an alarm condition that will turn off auto mode. Thus, auto mode being off due to an alarm is confounded with a user choosing not to use auto mode. However, when auto mode is on, there are no likely alarm states that affect pump performance and the pump software is controlling insulin flow. The fact that there are more hypo events when auto mode is on rather than off is somewhat alarming since one would like to minimize the more dangerous hypo events at the expense of more hyper events as is the case for when the auto mode is off. The database reports were not very detailed, so we were unable to draw any conclusions regarding the origin of this finding.

We have no way of knowing what caused the adverse events. Possibilities include user error, inaccurate sensor readings, or pump malfunctions (including the algorithm to deliver insulin). Pumps were not often returned by the users for analysis by Medtronic, leading to many inconclusive reports regarding the pump's role in the adverse events.

Tandem T:slim X2 Basal-IQ

The Tandem t:slim X2 Basal-IQ partially implemented a hybrid closed-loop system by using an algorithm to prevent hypos.[10] Of the 579 glucose values in our dataset that were not in the 70-180 mg/dL range, 89.7% were hyperglycemic. This is higher than the hyperglycemic results found for the Medtronic 670G in Table 2 and perhaps not surprising since the Basal-IQ pump version was designed to prevent hypos.

The most frequent complaint error in Table 3 (malfunction alarm) is a generic term for a system pump error. The second most frequent error (CGM error 42) relates to the Dexcom sensor. Thus, the Tandem pump is really a system (CGM plus pump) as is true for the Medtronic pump. Many users stated that the altitude alarm occurred in spite of the altitude being within limits.

Dexcom CGMs

The subset of data containing the terms "INACCURACY" and "METER READING" were all Dexcom brands with 87% G6, 11% G5, and 2% G4. We took advantage of the use of these terms by Dexcom personnel to facilitate our SQL queries. Other brands use different terms when results between CGM and BGM do not agree; hence, there is no implication that Dexcom brands are more inaccurate than other brands.

The range of the Dexcom CGMs is 40-400 mg/dL. This explains the straight lines at Y = 400 mg/dL and Y = 40 mg/dL in Figure 2. Note that in Figure 2, CGM values below 40 and above 400 were nevertheless in the dataset.

The paucity of data in the A zone is understandable since the complaint for these adverse events was inaccuracy.

The event type was classified as malfunction 98.4% of the time with only 1.6% of events classified as injury despite the fact that 5.1% of the results were in the D and E zones and 37.1% in C or higher zones. Most of the injury events were associated with outcomes such as emergency medical treatment or hospitalization and not the location of the point in the Parkes error grid.

In the Parkes error grid[11] values in both the A and B zones should have no impact on clinical outcomes. Yet, clearly users do not feel this way since they complained about inaccuracy for B zone results, which occurred 61.2% of the time. As an example, a value of 200 (CGM) vs 326 (BGM) is in the middle of the lower B zone and is an error of –39%. Regardless of the clinical implications, –39% is a large error to a user.

Abbott Freestyle

The most frequent complaint in Table 5 (54.8%) for the Abbott Freestyle was an allergic skin reaction. In the entire dataset of Dexcom G6 entries (eg, not the inaccuracy subset), the terms skin and allergic were found only 0.2%. Another striking difference was that Abbott classified events as injury 94.8% of the time whereas in the entire Dexcom G6 dataset ($N = 125\ 753$), only 1.1% events were classified as injury.

Conclusion

One must not lose sight of the fact that these results must be interpreted with two limitations:

1. The number of events is misleading without rates (the denominator of usage).
2. There is no way to know if all events that should be reported have been reported and have been properly classified.

One reaction to these limitations is to ignore the adverse event database. Alternatively, one can try to find information in the data while respecting the limitations.

For example, the Medtronic 670G system had a higher percentage of hypo events when the auto mode was on than in manual mode. Tandem abnormal glucose events were largely (89.7%) hyperglycemic. Users complained about Dexcom CGM inaccuracy for B zone results, which occurred 61.2% of the time. The most frequent Freestyle adverse event was an allergic skin reaction.

We need to try to eliminate the above limitations:

1. The FDA needs to enforce the medical device reporting regulation to ensure consistent reporting among manufacturers.
2. We need usage information to obtain rates.

A detailed compilation of adverse events regarding a product could benefit all stakeholders: manufacturers, clinicians, educators, and regulators. The compilation would be extremely valuable for ensuring betterment of the product and the safety of the users.

Abbreviations

FDA, Food and Drug Administration; MAUDE, Manufacturer and User Facility Device Experience; CGM, continuous glucose monitor; BGM, blood glucose meter; hyper, hyperglycemic; hypo, hyperglycemic; vba, visual basic for applications; SQL, structured query language.

ORCID iD

Jan S. Krouwer [iD] https://orcid.org/0000-0003-2300-076X

References

1. Krouwer JS. Reducing glucose meter adverse events by using reliability growth with the FDA MAUDE database. *J Diabetes Sci Technol*. 2019;(5):959-962.
2. U.S. Food & Drug Administration. Manufacturer and user facility device experience database - (MAUDE). https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/manufacturer-and-user-facility-device-experience-database-maude. Accessed June 23, 2020.
3. ega: error grid analysis. https://rdrr.io/cran/ega/. Accessed June 23, 2020.
4. U.S. Food & Drug Administration. CFR - Code of federal regulations title 21. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=803&showFR=1. Accessed May 24, 2020.
5. Medtronic. 670 insulin pump system. https://www.medtronicdiabetes.com/products/minimed-670g-insulin-pump-system. Accessed June 23, 2020.
6. Messer LH, Berget C, Vigers T, et al. Real world hybrid closed-loop discontinuation: predictors and perceptions of youth discontinuing the 670G system in the first 6 months. *Pediatr Diabetes*. 2020;21(2):319-327.
7. Lal RA, Basina M, Maahs D, Hood K, Buckingham B, Wilson DM. One year clinical experience of the first commercial hybrid closed-loop system. *Diabetes Care*. 2019;42(12):2190-2196.
8. Stone MP, Agrawal P, Chen X, et al. Retrospective analysis of 3-month real-world glucose data after the MiniMed 670G system commercial launch. *Diabetes Technol Ther*. 2018;20(10):689-692.
9. Medtronic. MINIMED® 670G system user guide. https://www.medtronicdiabetes.com/sites/default/files/library/download-library/user-guides/MiniMed%20670G%20System%20User%20Guide.pdf. Accessed June 23, 2020.
10. Tandem Basal-IQ Technology. https://www.tandemdiabetes.com/providers/products/basal-iq. Accessed June 23, 2020.
11. Parkes JL, Slatin SL, Pardo S, Ginsberg BH. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care*. 2000;23(8):1143-1148.