Journal of Diabetes Science and Technology
Volume 6, Issue 3, May 2012

COMMENTARY

The Challenges of Measuring Glycemic Variability

David Rodbard, M.D.

Abstract

This commentary reviews several of the challenges encountered when attempting to quantify glycemic variability and correlate it with risk of diabetes complications. These challenges include (1) immaturity of the field, including problems of data accuracy, precision, reliability, cost, and availability; (2) larger *relative error* in the estimates of glycemic variability than in the estimates of the mean glucose; (3) high correlation between glycemic variability and mean glucose level; (4) multiplicity of measures; (5) correlation of the multiple measures; (6) duplication or reinvention of methods; (7) confusion of measures of glycemic variability with measures of quality of glycemic control; (8) the problem of multiple comparisons when assessing relationships among multiple measures of variability and multiple clinical end points; and (9) differing needs for routine clinical practice and clinical research applications.

J Diabetes Sci Technol 2012;6(3):712-715

Let us consider some of the reasons why measurement of glycemic variability is more challenging than measurement of a mean glucose or hemoglobin A1c and identify some of the pitfalls.

1. **Measurement of glycemic variability is challenging because it is relatively new.** With all due respect to the seminal work of Service and colleagues[1–3] who introduced the mean amplitude of glucose excursion (MAGE), mean of daily differences (MODD), and several other measures dating back to the early 1970s, the large-scale study of glycemic variability began in earnest only with the advent of commercially available continuous glucose monitoring (CGM) and the scientific interest and controversy in the hypothesis that long-term complications of diabetes could be linked to glycemic variability.[4–6] The field is new and immature. The accuracy, precision, stability, reliability, and availability of CGM are improving[7,8] but not yet comparable to that of self-monitoring of blood glucose (SMBG) and laboratory measurements. Methods for smoothing and preprocessing of signals from the glucose sensor may affect results obtained for measures of glycemic variability.[9]

2. **There is intrinsically greater uncertainty in the measurement of variability than in the measurement of the mean.** When data obey a Gaussian distribution, the expected standard error of the mean

Author Affiliation: Biomedical Informatics Consultants LLC, Potomac, Maryland

Abbreviations: (CGM) continuous glucose monitoring, (CV) coefficient of variation, (GRADE) glycemic risk assessment diabetes equation, (MAD) mean absolute difference, (MAG) mean absolute glucose rate of change, (MAGE) mean amplitude of glucose excursion, (MODD) mean of daily differences, (SD) standard deviation, (SEM) standard error of the mean, (SMBG) self-monitoring of blood glucose

Keywords: continuous glucose monitoring, glycemic variability, hypoglycemia, quality of glycemic control, self-monitoring of blood glucose

Corresponding Author: David Rodbard, M.D., Biomedical Informatics Consultants LLC, 10113 Bentcross Dr., Potomac MD 20854-4721; email address drodbard@comcast.net

(SEM) is $\sigma/\sqrt{N}$, where σ is the true (population) standard deviation (SD) and N is the number of independent measurements on which the mean is based. In contrast, the expected standard error of the SD[10] is given by $se(SD) = \sigma/\sqrt{(2N)}$, which is 0.71 times the SEM. For most patients with diabetes, the SD is only 25% to 50% of the mean glucose (typically 35%). Thus the relative error in the estimate of the SD, $100 \times \{se(SD)\}/SD$, and most of the other measures of glycemic variability, would be 0.71/0.35 or approximately two-fold larger than the relative error for the estimate of the mean. For numerical example, if the mean of 25 *independent* glucose measurements were 150 and the SD was 50, then the SEM would be $50/\sqrt{25} = 50/5 = 10$, i.e., the mean ± SEM is 150 ± 10, or there is a relative error of 10/150 = 0.067, i.e., a 6.7% *relative error* in the estimate of the mean. The standard error of the SD would be expected to be $50/\sqrt{(2*25)} = 7.1$ mg/dl, corresponding to a *relative error* in the estimate of the SD of 7.1/50 = 0.142 or 14.2%. Thus the mean can often be estimated fairly well (±6.7%), but the SD is relatively poorly determined because of random sampling error (±14.2%). This is under a best-case scenario when the data are Gaussian and when there are a sufficient number of *independent* observations.[10] Measurement of glycemic variability will be subject to even larger measurement errors in the presence of outliers, when distributions are non-Gaussian, and when glucose measurements are not independent (e.g., due to the autocorrelation of CGM measurements).

3. **The magnitude of glycemic variability is highly correlated with the level of the mean.** The strong correlation of SD with the mean glucose[11–14] makes it extremely difficult to evaluate whether clinical outcomes are related to the mean, to glycemic variability, or to both. One can use multiple regression to try to separate the effects of variability and mean level.[12] However, in view of the larger measurement error in the estimate of glycemic variability than in the mean (discussed earlier), one would expect higher correlations of biological and clinical effects with the mean glucose than with the SD of glucose, even if both factors were equally important in terms of their biological or pathophysiological roles. Unlike the case with *in vitro* studies,[15] it is difficult to systematically and independently vary the mean and the glycemic variability, although this has been achieved in one study (e.g., HEART2D).[16]

4. **Uncertainty and ambiguity regarding the choice of the "right" measure of variability.** We are now blessed with a multiplicity of measures of glycemic variability. These include SD, coefficient of variation (CV), interquartile range, MAGE, MODD, mean absolute glucose (MAG) change per hour, and several subtypes of the SD,[11–27] including total SD (SD_T), SD *within* days (SD_w), *between* daily means (SD_{dm}), *between* days (SD_b), *between* days after correction for variability in *daily means* ($SD_{b\ //\ dm}$), variability by time of day for the mean glucose profile from several days ($SD_{hh:mm}$), measures of the stability of the glycemic profile or similarity of glucose patterns from day to day (root mean square error),[17–19] a "distance travelled"[20] that is closely related to the MAG rate of change,[21] glucose fluctuation as defined by Mori and associates,[22] and combinations of several of these measures.[11,19,23] Workers in the field need some time to evaluate and compare the performance of the various measures that have been proposed. This will enable us to examine the correlations among the several measures of variability[13,14] and see if a consensus begins to emerge regarding which is the "best" or "most sensitive" or "more reliable" or "most reproducible" and "easiest to understand" of the various measures. Further, we need to examine which combinations of parameters, with what kind of weighting, will give the best correlation with clinically important events and long-term complications.[11,14,19,23]

5. **High degree of correlation of different measures of glycemic variability.** Not only do we have a multiplicity of measures, but these measures are highly correlated among themselves. This makes it difficult to determine which is the "right" measure, the "best" measure, or the measure with the greatest sensitivity to detect an effect such as a change in oxidative stress, coronary artery calcification,[14] cardiovascular disease, or measures of psychological factors such as depression or anxiety. If the measures are highly correlated, perhaps it matters less which measure one uses.

   Each of the various parameters have their own problems with measurement error. Baghurst and colleagues have shown that the random sampling errors in the SD are significantly and consistently smaller than for MAGE or for CONGA1 (continuous overlapping net glycemic index, i.e. standard deviation of successive differences

of glucose values separated by 1 h). Baghurst also introduced the concepts of MAGE considering only upstrokes (MAGE+), MAGE considering only downstrokes (MAGE_), and the average of the two (MAGEavge). Mathematicians and statisticians have long known that SD is the most "efficient" statistic *if* the data are subject to a normal distribution, and this is often true even when the data depart from a Gaussian distribution.

6. **Rediscovery, reinvention, and renaming of various measures.** With so much excitement and activity in the field, some workers are describing "new" measures that have already been described under another name. For example, the "glycemic fluctuation"[22] is mathematically interchangeable with a conventional statistical method, the mean absolute difference (MAD) of glucose from the mean, multiplied by a constant (24 h) to convert the MAD to an area under the curve. For data obeying a Gaussian distribution, the MAD is directly proportional to the SD, according to the following formula:[28] MAD = Ö(2/p) × SD. This relationship applies to a Gaussian distribution; other relationships are available for other types of distributions.[28] MAD = √(2/π) × SD. Thus glycemic fluctuation[22] is expected to be proportional to the SD with a slope of 24 × √(2/3.1416) = 19.15.

"Distance travelled" as proposed by Marling and colleagues [20] for analysis of CGM data would be mathematically equivalent to the MAG introduced by Hermanides and associates[21] in the context of SMBG or laboratory glucose values *if glucose data were equally spaced on the time axis*. Authors should examine the empirical correlation of "new" measures with previously proposed measures and also compare the theoretical or mathematical properties of the methods to detect relationships, similarities, or identities.

7. **Need to distinguish between measures of "glycemic variability" and measures of "glycemic control."** A few measures of quality of glycemic control (hemoglobin A1c, mean glucose, fructosamine, glycated albumin) do not consider glycemic variability. However, most measures of glycemic control are sensitive to glycemic variability. Nevertheless, measures such as Schlichtkrull's M value, the J index, the high blood glucose index/low blood glucose index) family of indices [BGRI (blood glucose risk index), ADRR (average daily risk range), HBGI (high blood glucose index), LBGI (low blood glucose index),[26] the index of glycemic control family of methods [IGC (index of glycemic control), hypoglycemia index, hyperglycemia index],[17,18] and the glycemic risk assessment diabetes equation (GRADE) familyof methods (GRADE, $\%GRADE_{HYPOGLYCEMIA}$, $\%GRADE_{HYPERGLYCEMIA}$, and $\%GRADE_{EUGLYCEMIA}$)[27] are attempts to measure *quality* of glycemic control and not simply glycemic variability. A simple measure such as "percentage of glucose values falling within any specified range" (e.g., 80 to 180 mg/dl; or, conversely, the percentage of glucose values falling outside that range) is intended to be a measure of quality of glycemic control that is affected by glycemic variability but is not a measure of glycemic variability *per se*.[13]

8. **Need to correct for the problem of "multiple comparisons."** With so many criteria for glycemic variability, one is tempted to examine each of them for possible relationships with the intervention or the effect(s) being studied. However, this increases the probability that some statistical significance tests will be positive simply because of the number of tests being performed. If one were to have 10 measures of clinical outcomes and 10 measures of glycemic variability, one could calculate 100 significance tests and their corresponding P values. Several of these might be significant at the $P < .05$ level due to random chance (sampling error) alone. The problem becomes more complicated when the several measures of glycemic variability are highly correlated. Hence it is necessary to specify, *a priori*, the primary response variables to test a specified hypothesis. This may be based on an exploratory analysis of a small subset of the data or of an independent pilot study. If need be, one can use a "resampling" type of statistical calculation to ensure that the multiplicity of hypotheses being tested are not generating spurious indications of statistical significance due to chance events.[29,30]

9. **The possibility of differing needs for routine clinical practice and for clinical research.** For routine clinical practice, it is likely that the SD and the corresponding CV obtained using either SMBG or CGM will be sufficient to permit assessment of changes in glycemic variability with time or following therapeutic interventions, and to permit comparison with reference populations of patients with similar type, duration, and level of control of hemoglobin A1c or mean glucose.[11] For research applications, it will often be desirable to examine several additional

measures of glycemic control and variability. In many situations, use of CGM will be informative and cost-effective.

Conclusion

Measurement of glycemic variability is and will remain challenging for a number of reasons, including those discussed here. The scientific community is just now attaining the requisite level of experience to resolve the methodological issues, identify common pitfalls, and optimize methods for data collection, analysis, and display.

Acknowledgment:

This study was supported by Biomedical Informatics Consultants LLC, Potomac, MD.

References:

1. Service FJ, Molnar GD, Rosevear JW, Ackerman E, Gatewood LC, Taylor WF. Mean amplitude of glycemic excursions, a measure of diabetic instability. Diabetes. 1970;19(9):644–55.
2. Service FJ, Nelson RL. Characteristics of glycemic stability. Diabetes Care. 1980;3(1):58–62.
3. Service FJ, O'Brien PC, Rizza RA. Measurements of glucose control. Diabetes Care. 1987;10(2):225–37.
4. Cameron FJ, Baghurst PA, Rodbard D. Assessing glycemic variation: why, when and how? Pediatr Endocrinol Rev. 2010;7 Suppl 3:432–44.
5. Siegelaar SE, Holleman F, Hoekstra JB, DeVries JH. Glucose variability; does it matter? Endocr Rev. 2010;31(2):171–82.
6. Weber C, Schnell O. The assessment of glycemic variability and its impact on diabetes-related complications: an overview. Diabetes Technol Ther. 2009;11(10):623–33.
7. Keenan DB, Grosman B, Clark HW, Roy A, Weinzimer SA, Shah RV, Mastrototaro JJ. Continuous glucose monitoring considerations for the development of a closed-loop artificial pancreas system. J Diabetes Sci Technol. 2011;5(6):1327–36.
8. McGarraugh GV, Clarke WL, Kovatchev BP. Comparison of the clinical information provided by the FreeStyle Navigator continuous interstitial glucose monitor versus traditional blood glucose readings. Diabetes Technol Ther. 2010;12(5):365–71.
9. Bequette BW. Continuous glucose monitoring: real-time algorithms for calibration, filtering, and alarms. J Diabetes Sci Technol. 2010;4(2):404–18.
10. Ahn S, Fessler JA. Standard errors of mean, variance, and standard deviation estimators. http://www.eecs.umich.edu/~fessler/papers/files/tr/stderr.pdf.
11. Rodbard D. Clinical interpretation of indices of quality of glycemic control and glycemic variability. Postgrad Med. 2011;123(4):107–18.
12. Kilpatrick ES. Arguments for and against the role of glucose variability in the development of diabetes complications. J Diabetes Sci Technol. 2009;3(4):649–55.
13. Rodbard D, Bailey T, Jovanovic L, Zisser H, Kaplan R, Garg SK. Improved quality of glycemic control and reduced glycemic variability with use of continuous glucose monitoring. Diabetes Technol Ther. 2009;11(11):717–23.
14. Snell-Bergeon JK, Roman R, Rodbard D, Garg S, Maahs DM, Schauer IE, Bergman BC, Kinney GL, Rewers M. Glycaemic variability is associated with coronary artery calcium in men with type 1 diabetes: the coronary artery calcification in type 1 diabetes study. Diabet Med. 2010;27(12):1436–42.
15. Schisano B, Tripathi G, McGee K, McTernan PG, Ceriello A. Glucose oscillations, more than constant high glucose, induce p53 activation and a metabolic memory in human endothelial cells. Diabetologia. 2011;54(5):1219–26.
16. Siegelaar SE, Kerr L, Jacober SJ, Devries JH. A decrease in glucose variability does not reduce cardiovascular event rates in type 2 diabetic patients after acute myocardial infarction: a reanalysis of the HEART2D study. Diabetes Care. 2011;34(4):855–7.
17. Rodbard D. Interpretation of continuous glucose monitoring data: glycemic variability and quality of glycemic control. Diabetes Technol Ther. 2009;11 Suppl 1:S55–67.
18. Rodbard D. New and improved methods to characterize glycemic variability using continuous glucose monitoring. Diabetes Technol Ther. 2009;11(9):551–65.
19. Rodbard D. Glycemic variability: measurement and utility in clinical medicine and research--one viewpoint. Diabetes Technol Ther. 2011;13(11):1077–80.
20. Marling CR, Shubrook JH, Vernier SJ, Wiley MT, Schwartz FL. Characterizing blood glucose variability using new metrics with continuous glucose monitoring data. J Diabetes Sci Technol. 2011;5(4):871–8.
21. Hermanides J, Vriesendorp TM, Bosman RJ, Zandstra DF, Hoekstra JB, Devries JH. Glucose variability is associated with intensive care unit mortality. Crit Care Med. 2010;38(3):838–42.
22. Mori Y, Taniguchi Y, Sezaki K, Yokoyama J, Utsunomiya K. Liraglutide narrows the range of circadian glycemic variations in Japanese type 2 diabetes patients and nearly flattens these variations in drug-naive type 2 diabetes patients: a continuous glucose monitoring-based study. Diabetes Technol Ther. 2011;13(11):1139–44.
23. Thomas A, Schönauer M, Achermann F, Schnell O, Hanefeld M, Ziegelasch HJ, Mastrototaro J, Heinemann L. The "glucose pentagon": assessing glycemic control of patients with diabetes mellitus by a model integrating different parameters from glucose profiles. Diabetes Technol Ther. 2009;11(6):399–409.
24. Baghurst PA, Rodbard D, Cameron FJ. The minimum frequency of glucose measurements from which glycemic variation can be consistently assessed. J Diabetes Sci Technol. 2010;4(6):1382–5.
25. Baghurst PA. Calculating the mean amplitude of glycemic excursion from continuous glucose monitoring data: an automated algorithm. Diabetes Technol Ther. 2011;13(3):296–302.
26. Clarke W, Kovatchev B. Statistical tools to analyze continuous glucose monitor data. Diabetes Technol Ther. 2009;11 Suppl 1:S45–54.
27. Hill NR, Thompson B, Bruce J, Matthews DR, Hindmarsh P. Glycaemic risk assessment in children and young people with Type 1 diabetes mellitus. Diabet Med. 2009;26(7):740–3.
28. Wolfram MathWorld. Mean deviation. http://mathworld.wolfram.com/MeanDeviation.html. Accessed November 21, 2011.
29. Wikipedia. Resampling (statistics). http://en.wikipedia.org/wiki/Resampling_(statistics). Accessed November 23, 2011.
30. Penckofer S, Quinn L, Byrn M, Ferrans C, Miller M, Strange P. Does glycemic variability impact mood and quality of life? Diabetes Technol Ther. 2012;14(4):303–10.