Glycemic Variability Measures

Evan Olawsky

11/15/2019

This document details the calculations made in the R script `GV measures.R`. That script is based upon the Easy GV spreadsheet constructed by Dr. Nathan R. Hill of Oxford University. In many cases, Easy GV does not give results consistent with the formulas presented in the original manuscripts. In these cases, `GV measures.R` gives an option to calculate either the Easy GV version or the original manuscript option.

Throughout this document, let X_t be a glucose reading at time t. Let n be the total number of glucose readings. Time is assumed to be measured in minutes since the first recording in the data set. Glucose readings can be measured in either mg/dL or mmol/L. Note that 1 mmol of glucose is equal to 18 mg of glucose.

All functions require a vector `x` of glucose readings. This vector should be numeric and should not include any blank entries. Some functions additionally require a vector `times` of times. This vector should also be numeric and should not include any blank entries. Currently, the function `read.CGM` can take a Dexcom output file as input and return a data.frame which includes properly formatted `x` and `times` vectors.

The wrapper function `GV` returns all of the following metrics simultaneously.

Continuous overall net glycemic action (CONGA) (McDonell et al. 2005)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `n`, the number of hours between "partner" observations. Null value is 1.
- `s`, the number of minutes of slack used when searching for partners. Null value is 1.
- `method`, either "manuscript" or "easy". Null value is "manuscript".

For a glucose measurement X_t at time t, let D_t be the difference between X_t and the mean of all glucose measurements made n hours prior to X_t, plus or minus s minutes. Let T be the set of times with a D_t value and let k be the number of such observations. Finally, let $\overline{D} = \sum D_t/k$. Then, the original manuscript version is

$$CONGA_M(n) = \sqrt{\frac{\sum_T (D_t - \overline{D})^2}{k-1}}$$

Furthermore, let $\overline{D}^* = \sum|D_t|/k$. Then the Easy GV version is

$$CONGA_{GV}(n) = \sqrt{\frac{\sum_T (X_t - \overline{D}^*)^2}{k-1}}$$

Lability Index (LI) (Ryan et al. 2004)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `k`, length of time (in minutes) used to find partners. Null value is 60.
- `s`, the number of minutes of slack used when searching for partners. Null value is 1.

For a glucose measurement X_t at time t, let D_t be the difference between X_t and the mean of all glucose measurements made k minutes prior to X_t, plus or minus s minutes. Let T be the set of times with a D_t value and let k be the number of such observations. Then

$$LI = \frac{1}{k}\sum_{t\in T}(D_t)^2$$

J-index (Wojcicki 1995)

Parameters include

- `x`, a vector of glucose readings
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".

Let $\overline{X}$ be the mean of all glucose values, and let $SD(X)$ be the standard deviation of all glucose values. If the units are mg/dL,

$$J = \frac{1}{1000}(\overline{X} + SD(X))^2$$

and if the units are mmol/L,

$$J = \frac{18^2}{1000}(\overline{X} + SD(X))^2$$

Low / High Blood Glucose Index (LBGI, HBGI) (Kovatchev et al. 2003) and (Gaynanova, Urbanek, and Punjabi 2018)

Parameters include

- `x`, a vector of glucose readings
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".
- `method`, one of "manuscript", "easy", or "corrected". Null value is "manuscript". "Corrected" refers to the Gaynanova paper's recommendation.

If the units are mg/dL, let $f(x) = 1.509(\ln(x)^{1.084} - 5.381)$. If the units are mmol/L, let $f(x) = 1.509(\ln(18x)^{1.084} - 5.381)$. Let $rl(x) = cf(x)^2$ when $f(x) < 0$ and $rl(x) = 0$ otherwise. Let $rh(x) = cf(x)^2$ when $f(x) > 0$ and $rh(x) = 0$ otherwise. In the original manuscript, Kovatchev et al use $c = 10$. Gaynanova et al recommend $c = 22.77$. Both measures can be obtained from this code, by setting `method` to `corrected` or `manuscript`, respectively. Then, the original manuscript version is

$$LBGI_M = \frac{\sum rl(x_t)}{n}$$

$$HBGI_M = \frac{\sum rh(x_t)}{n}$$

where n is the total number of glucose readings.

The Easy GV version is

$$LBGI_{GV} = \frac{\sum rl(x_t)}{\sum I(rl(x_t) > 0)}$$

$$HBGI_{GV} = \frac{\sum rh(x_t)}{\sum I(rh(x_t) > 0)}$$

Glycemic Risk Assessment Diabetes Equation (GRADE) (Hill et al. 2007)

Parameters include

- `x`, a vector of glucose readings
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".
- `method`, either "manuscript" or "easy". Null value is "manuscript".
- `c1`, the glucose value below which readings are considered hypoglycemic
- `c2`, the glucose value above which readings are considered hyperglycemic

If the units are mg/dL, let

$$g(x) = \min(425[\log(\log(x/18)) + C]^2, 50)$$

If the units are mmol/L, let

$$g(x) = \min(425[\log(\log(x)) + C]^2, 50)$$

Where the logarithm is base ten in both cases. Let $C = 1.6$ for the manuscript calculation and let $C = 1.5554147$ for the Easy GV calculation.

For the manuscript calculation, $GRADE_M$ is the mean of the $g(x_t)$. For the Easy GV calculation, $GRADE_{GV}$ is the median of the $g(x_t)$.

We also calculate the contributions of hypoglycemia, euglycemia, and hyperglycemia to the GRADE score.

$$\text{Hypo percentage} = \frac{\sum_{x_t<C_1} g\,(x_t)}{\sum_{\text{all } x_t} g\,(x_t)}$$

$$\text{Eu percentage} = \frac{\sum_{C_1<x_t<C_2} g\,(x_t)}{\sum_{\text{all } x_t} g\,(x_t)}$$

$$\text{Hyper percentage} = \frac{\sum_{x_t>C_2} g\,(x_t)}{\sum_{\text{all } x_t} g\,(x_t)}$$

If the units are mg/dL, the default values for C_1 and C_2 are 70.2 and 140.4$ If the units are mmol/L, the defaults are 3.9 and 7.8.

Mean of Daily Differences (MODD) (Molnar, Taylor, and Ho 1972)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `s`, the number of minutes of slack used when searching for partners. Null value is 1.
- `method`, either “manuscript” or “easy”. Null value is “manuscript”.

For a glucose measurement X_t at time t, let D_t be the difference between X_t and the mean of all glucose measurements made 24 hours prior to X_t, plus or minus s minutes. Let T be the set of times with a D_t value and let k be the number of such observations.

Then, the original manuscript version is

$$MODD_M = \frac{1}{k}\sum_{T} |\,D_t|$$

Let $T^- = T \backslash max(t \in T)$. Then, the Easy GV version is

$$MODD_{GV} = \frac{1}{K-1}\sum_{T^-} |\,D_t|$$

Mean Amplitude of Glycemic Excursions (MAGE) (Service et al. 1970)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times

Note that the original manuscript for MAGE is not very precise and does not lead to an obvious calculation of MAGE. While Easy GV does not appear to calculate MAGE in the same way as the original manuscript, the Easy GV version of MAGE is the only one we present here.

Let $D_t = X_t - X_{t-1}$. Then let E be the set of all D_t whose absolute value exceeds the standard deviation of all glucose readings from the day that D_t occurred. Then let E^+ be the set that contains the positive D_t values in E, with size $\#E^+$. Let E^- be the set that contains the negative D_t values in E, with size $\#E^+$. We then report separate positive and negative MAGE values and the averaged MAGE value:

$$MAGE_+ = \frac{1}{\#E^+}\sum_{E^+} D_t$$

$$MAGE\_ = \frac{1}{\#E^-}\sum_{E^-} D_t$$

$$MAGE = (MAGE_+ + MAGE\_)/2$$

Average Daily Risk Range (ADRR) (Kovatchev et al. 2006)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".
- `method`, either "manuscript" or "easy". Null value is "manuscript".

If the units are mg/dL, let

$$f(x) = [1.509(\ln(x)^{1.084} - 5.381)]$$

If the units are mmol/L, let

$$f(x) = [1.509(\ln(18x)^{1.084} - 5.381)]$$

Let $rl(x) = 10f(x)^2$ when $f(x) < 0$ and $rl(x) = 0$ otherwise. Let $rh(x) = 10f(x)^2$ when $f(x) > 0$ and $rh(x) = 0$ otherwise. Denote $(x_1^d, \dots, x_{n_d}^d)$ as the n_d glucose values on day d. Then let

$$LR^d = \max(rl(x_1^d), \ldots, rl(x_{n_d}^d))$$

and

$$HR^d = max(rh(x_1^d), \ldots, rh(x_{n_d}^d))$$

for day d. Let D be the total number of days where glucose levels were measured. Then, the original manuscript version is

$$ADRR_M = \frac{1}{D}\sum_{d=1}^{D}(LR^d + HR^d)$$

The Easy GV version gives high and low measures separately.

$$ADRR_L = \frac{1}{D}\sum_{d=1}^{D}(LR^d)$$

$$ADRR_H = \frac{1}{D}\sum_{d=1}^{D}(HR^d)$$

M-value (Schlichtkrull, Munck, and Jersild 1965)

Parameters include

- `x`, a vector of glucose readings
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".
- `index`, a value to be considered a 'standard' blood glucose value, in mg/dL. Null value is 120.
- `method`, either "manuscript" or "easy". Null value is "manuscript".

After conversion of all glucose values to mg/dL, let $M^* = (10\log\frac{x}{\text{index}})^3$ and let $W = (max(x_i) - min(x_i))/20$. The log used in that equation is base 10.

Then, the original manuscript version is

$$M_M = \frac{1}{N}\sum_{i=1}^{N}|M^*| + W$$

The Easy GV version is

$$M_{GV} = \frac{1}{N}\sum_{i=1}^{N}|M^*|$$

Mean Absolute Glucose (MAG) (Hermanides et al. 2010)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times

$$MAG = \frac{\sum_{i=1}^{N-1} | x_{i+1} - x_i |}{(\max(t) - \min(t))/60}$$

where N is the total number of glucose values.

Coefficient of variation (CV)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `overall`, a logical, equal to `TRUE` you want the CV for the entire dataset, or equal to `FALSE` if you would prefer many CV values over a moving window
- `interval`, size (in hours) of the moving window to be used if `overall` is false. Null value is 1.

$$CV = SD(X)/\overline{X},$$

where X is a vector of glucose readings, potentially restricted to a particular time window.

Standard deviation (SD)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `overall`, a logical, equal to `TRUE` you want the SD for the entire dataset, or equal to `FALSE` if you would prefer many SD values over a moving window
- `interval`, size (in hours) of the moving window to be used if `overall` is false. Null value is 1.

$$SD = \sqrt{\frac{1}{N-1}\sum_{t \in T} (x_t - \overline{x})^2},$$

where T is a set of times (potentially restricted to a particular window) and N is the size of T.

Area Under the Curve (AUC)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `thresh`, a threshold above (or below) which you wish to calculate the AUC. Default is 100.
- `above`, a logical indicating whether you wish to calculate area above the threshold value (`TRUE`) or below it (`FALSE`). Default is `TRUE`.

If `above == T`,

$$AUC_+ = \sum_{i=1}^{N-1} I\,(x_i \geq v) I(x_{i+1} \geq v)(min(x_i - v, x_{i+1} - v)(t_{i+1} - t_i) + |x_{i+1} - x_i|(t_{i+1} - t_i)/2)$$

If `above == F`,

$$AUC_- = -\sum_{i=1}^{N-1} I\,(x_i \leq v) I(x_{i+1} \leq v)(min(x_i - v, x_{i+1} - v)(t_{i+1} - t_i) + |x_{i+1} - x_i|(t_{i+1} - t_i)/2)$$

where v is the threshold value and N is the length of the glucose vector.

For each excursion beyond this threshold value, this calculation does not include the triangular area from the threshold to the first glucose value beyond the threshold, nor from the last glucose beyond the threshold back to the threshold. Hence a single glucose value beyond the threshold is not captured by the calculation.

Time spent in range (TIR) (Battelino and others 2019)

Parameters include

- `x`, a vector of glucose readings
- `low`, the lower bound of the range
- `high`, the upper bound of the range

This function gives the percentage of glucose readings that fall in a given range (l, u).

$$TIR = \sum_{i=1}^{N} I\,(l \leq x_i \leq u)/N$$

Battelino et al suggest five ranges: below 54 mg/dL, 55-70, 71-180, 181-250, above 250.

Glucose Management Indicator (GMI) (Bergenstal and others 2018)

Parameters include

- `x`, a vector of glucose readings
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".

Let $\overline{x}$ be the mean of all glucose readings taken. If the units are mg/dL, then

$$GMI = 3.31 + 0.02392\overline{x}$$

If the units are mmol/L,

$$12.71 + 4.70587\overline{x}$$

Number of episodes per day

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `thresh`, a threshold, where glucoses below the threshold are considered as part of an episode
- `len`, the minimum length of an episode
- `gap`, the typical gap between CGM measurements, in minutes

This function counts the number of "episodes" where glucose values remain below a certain threshold `thresh` for a period of at least `len` minutes. Then the number of episodes is divided by the amount of days that the sensor was active. This amount is calculated by taking the total time (the time between the first and last measurements), subtracting any gaps in time that are longer than `gap`+2 minutes and then adding back `gap` minutes for each of the gaps subtracted away.

Glycemic Variability Percentage (GVP) (Peyser et al. 2018)

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times

Let $\Delta x_i = x_i - x_{i-1}$ and $\Delta t_i = t_i - t_{i-1}$ for $i = 2, \ldots, n$. Then let $L = \sum_{i=2}^{n} \sqrt{\Delta x_i^2 + \Delta t_i^2}$ and $L_0 = \sum_{i=2}^{n} \Delta\, t_i$. Then $GVP = (L/L_0 - 1) \times 100$.

Distance Travelled (Marling et al. 2011)

Parameters include

- x, a vector of glucose readings

Let $\Delta x_i = x_i - x_{i-1}$ for $i = 2, \ldots, n$. Then the distance travelled is equal to $\sum_{i=2}^{n} |\Delta x_i|$.

Other functions

read.CGM

Parameters include

- `file`, the name of the file (in CSV format) to be read-in
- `timezero`, set to `"first"` if the first glucose reading should be considered time zero and set to `"midnight"` if midnight of the day of the first reading should be considered time zero. Default is `"first"`.
- `na.rm`, a logical that is `TRUE` if you wish to exclude all readings that are missing glucose values or time stamps and `FALSE` if not. Default is `TRUE`.
- `skip`, the number of lines in the data file to skip before beginning to read in data
- `calib.col`, the number or name of the column containing information regarding calibration status of each glucose entry
- `calib.tag`, the character value used to denote calibration rows in `calib.col`
- `mult.sensors`, a logical that is `TRUE` if you wish to split the data set into parts corresponding to different CGM sensors and `FALSE` if not. Default is `FALSE`.
- `sensor.times`, a vector of times (in the same format as the time data) that correspond to the beginning of a new CGM sensor. These times are used to split the data between multiple sensors if `mult.sensors` is `TRUE`. If `sensor.times` is `NA`, the data is split automatically at every gap of `sensor.gap` or more minutes.
- `sensor.gaps`, a number specifying the minimum gap (in minutes) for which we should split the data into two pieces. Default is 120.
- `time.col`, the number or name of the column containing time data
- `gluc.col`, the number or name of the column containing glucose data
- `time.sep`, character that separates date from time in your time data
- `time.format`, specify date and time formats according to the specification used in the `chron` package. Default is `c(dates = "m/d/y", times = "h:m:s")`.
- `high.ind`, character value that identifies high glucose values in the data. Default is "High".
- `high.value`, numeric value by which to replace glucose values equal to `high.ind`. Default is 400.
- `low.ind`, character value that identifies low glucose values in the data. Default is "Low".

- `low.value`, numeric value by which to replace glucose values equal to `low.ind`. Default is 40.

This function takes in data from a CGM and converts it into a data frame with one column of glucose readings and one column of times (in minutes). These two columns can then be used with any of the glucose variability functions.

plot.CGM

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".

This function returns a plot of blood glucose over time.

plot.diff

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `n`, the number of hours between "partner" observations. Null value is 1.
- `s`, the number of minutes of slack used when searching for partners. Null value is 1.
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".

This function returns a plot of the `n`-hour changes in glucose values over time.

plot.symm

Parameters include

- `x`, a vector of glucose readings
- `times`, a vector of corresponding times
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".

This function returns a plot of the "symmetrized" glucose values used in calculating BGI and ADRR.

GV

Parameters include

- `x`, a vector of glucose readings

- `times`, a vector of corresponding times
- `unit`, either "mg" if the units are mg/dL or "mmol" if the units are mmol/L. Null value is "mg".
- `m.index`, a value to be considered a 'standard' blood glucose value, in mg/dL. Null value is 120.
- `k`, length of time (in minutes) used to find partners. Null value is 60.
- `s`, the number of minutes of slack used when searching for partners. Null value is 1.
- `conga.n`, the number of hours between "partner" observations. Null value is 1.
- `interval`, size (in hours) of the moving window to be used if `overall` is false. Null value is 1.
- `thresh`, a threshold above (or below) which you wish to calculate percentages. Default is 100

This is a wrapper function that outputs a table with all 14 metrics, calculated for both manuscript and Easy GV methods, if applicable.

References

Battelino, Tadej, and others. 2019. "Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations from the International Consensus on Time in Range." *Diabetes Care* 42: 1593–1603.

Bergenstal, Richard M., and others. 2018. "Glucose Management Indicator (Gmi): A New Term for Estimating A1c from Continuous Glucose Monitoring." *Diabetes Care* 41: 2275–80.

Gaynanova, Irina, Jacek Urbanek, and Naresh M. Punjabi. 2018. "Corrections of Equations on Glycemic Variability and Quality of Glycemic Control." *Diabetes Technology and Therapeutics* 20 (4): 317.

Hermanides, Jeroen, Titia M. Vriesendorp, Robert J. Bosman, Durk F. Zandstra, Joost B. Hoekstra, and J. Han DeVries. 2010. "Glucose Variability Is Associated with Intensive Care Unit Mortality." *Critical Care Medicine* 38 (3): 838–42.

Hill, N.R., P.C. Hindmarsh, R.J. Stevens, I.M. Stratton, J.C. Levy, and D.R. Matthews. 2007. "A Method for Assessing Quality of Control from Glucose Profiles." *Diabetic Medicine* 24: 753–58.

Kovatchev, Boris P., Daniel Cox, Anand Kumar, Linda Gonder-Frederick, and William L. Clarke. 2003. "Algorithmic Evaluation of Metabolic Control and Risk of Severe Hypoglycemia in Type 1 and Type 2 Diabetes Using Self-Monitoring Blood Glucose Data." *Diabetes Technology and Therapeutics* 5 (5): 817–28.

Kovatchev, Boris P., Erik Otto, Daniel Cox, Linda Gonder-Frederick, and William Clarke. 2006. "Evaluation of a New Measure of Blood Glucose Variability in Diabetes." *Diabetes Care* 29 (11): 2433–8.

Marling, Cynthia R., Jay H. Shubrook, Stanley J. Vernier, Matthew T. Wiley, and Frank L. Schwartz. 2011. “Characterizing Blood Glucose Variability Using New Metrics with Continuous Glucose Monitoring Data.” *Journal of Diabetes Science and Technology* 5 (4): 871–78.

McDonell, C.M., S.M. Donath, S.I. Vidmar, G.A. Werhter, and F.J. Cameron. 2005. “A Novel Approach to Continuous Glucose Analysis Utilizing Glycemic Variation.” *Diabetes Technology and Therapeutics* 7 (2): 253–63.

Molnar, G.D., W.F. Taylor, and M.M. Ho. 1972. “Day-to-Day Variation of Continuously Monitored Glycaemia: A Further Measure of Diabetic Instability.” *Diabetologia* 8: 342–48.

Peyser, Thomas A., Andrew K. Balo, Bruce A. Buckingham, Irl B. Hirsch, and Arturo Garcia. 2018. “Glycemic Variability Percentage: A Novel Method for Assessing Glycemic Variability from Continuous Glucose Monitor Data.” *Diabetes Technology and Therapeutics* 20 (1): 6–16.

Ryan, Edmond A., Tami Shandro, Kristy Green, Breay W. Path, Peter A. Senior, David Bigam, A.M. James Shapiro, and Marie-Christine Vantyghem. 2004. “Assessment of the Severity of Hypoglycemia and Glycemic Lability in Type 1 Diabetic Subjects Undergoing Islet Transplantation.” *Diabetes* 53: 955–62.

Schlichtkrull, J., O. Munck, and M. Jersild. 1965. “The M-Value, an Index of Blood-Sugar Control in Diabetics.” *Acta Medica Scandinavia* 177 (1): 95–102.

Service, F. John, George D. Molnar, John W. Rosevear, Eugene Ackerman, Lael C. Gatewood, and William F. Taylor. 1970. “Mean Amplitude of Glycemic Excursions, a Measure of Diabetic Instability.” *Diabetes* 19 (9): 644–55.

Wojcicki, J.M. 1995. “J-Index. A New Proposition of the Assessment of Current Glucose Control in Diabetic Patients.” *Hormone and Metabolic Research* 27: 41–42.