

## **Supplementary Material**

**“Three components of glucose dynamics – value, variability, and autocorrelation – are independently associated with coronary plaque vulnerability”**

### **Table of Contents**

<b>Supplementary Texts .....</b>	<b>2</b>
<b>Supplementary Figures .....</b>	<b>5</b>
<b>Supplementary References.....</b>	<b>22</b>

## Supplementary Texts

### CGM-derived indices are effective in predicting coronary plaque vulnerability

To characterize CGM-derived indices in estimating the risk of CAD, we examined Spearman's correlation coefficients ( $r$ ) between CGM-derived indices and the ratio of necrotic core to total plaque volume (%NC) (Fig. S1). %NC, a widely used parameter of plaque vulnerability, was assessed by VH-IVUS. We performed this analysis using a previously described cohort consisting of 8 individuals with normal glucose tolerance (NGT), 16 with IGT, and 29 with T2DM (Otowa-Suematsu et al., 2018). For comparison, we also investigated FBG, HbA1c, OGTT-derived indices, body mass index (BMI), triglycerides (TGs), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), systolic blood pressure (SBP), and diastolic BP (DBP).

We constructed a correlation network connecting relationships with  $Q < 0.05$  (Fig. S1A). The  $Q$  values were calculated by Spearman's correlation test followed by multiple testing adjustment using the Benjamini-Hochberg method. The correlation network showed that AC\_Var was statistically significantly correlated with %NC ( $r = 0.35$ ; 95% confidence interval [CI], 0.09–0.57) (Fig. S1A). By contrast, AC\_Var displayed relatively weak correlations with other indices, including CGM\_Mean ( $r = -0.02$ ; 95% CI, -0.30–0.24) and CGM\_Std ( $r = 0.15$ ; 95% CI, -0.14–0.43) (Fig. S1A, B).

Twelve CGM-derived indices, namely, ADRR, MAGE, JINDEX, CGM\_Std (Std of glucose levels measured by CGM), CGM\_Mean (mean glucose levels measured by CGM), GRADE, MVALUE, AC\_Var, LI, HBGI, CONGA, and MODD, exhibited significant correlations with %NC (Fig. S1C). By contrast, with the exception of the insulinogenic index (I.I.), OGTT-derived indices, as well as other indices including FBG and HbA1c, displayed relatively weak correlations with %NC (Fig. S1C, blue, magenta, and green). Of note, this study enrolled individuals with well-controlled serum cholesterol and BP levels (Fig. S1D).

The weak correlations between cholesterol and BP-related indices and %NC (Fig. S1C) do not mean that cholesterol and BP are not associated with %NC. Consequently, our subsequent analysis focused primarily on indices unrelated to cholesterol and BP. Several relationships showed some linearity, including the association between ADRR and %NC (Fig. S2).

To predict %NC and further assess the efficacy of CGM-derived indices in estimating %NC, we conducted multiple linear regression analyses including not only CGM\_Mean, CGM\_Std, and AC\_Var but also other indices. Given that our study only enrolled individuals with well-controlled serum cholesterol levels and BP, as indicated in Figure S1D, we included only BMI, FBG, HbA1c, OGTT-derived indices, and CGM-derived indices as the input variables.

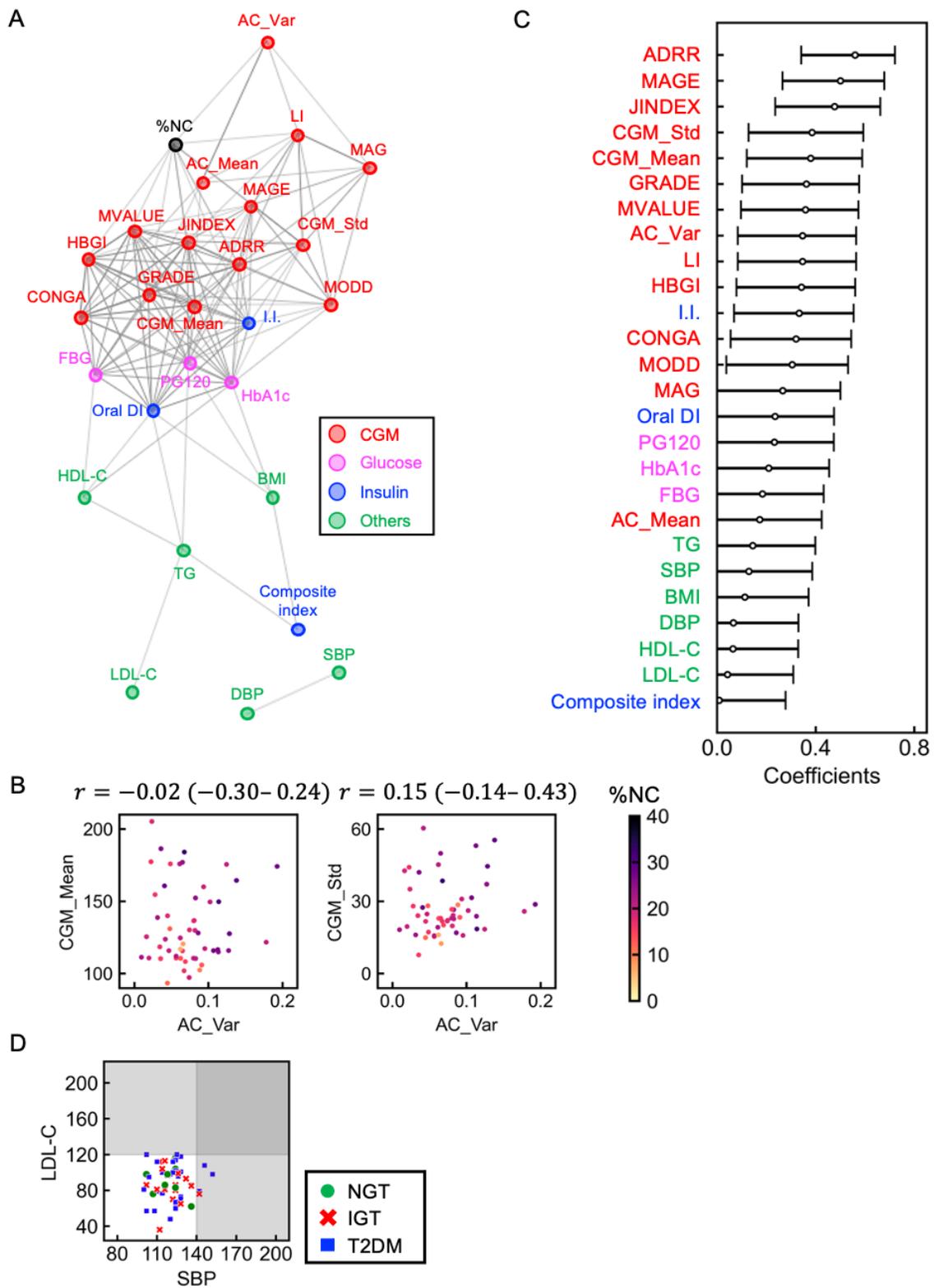
We evaluated three distinct models: the first model included all variables as input variables (referred to as Full); the second model excluded CGM-derived indices (referred to as CGM-); and the third model included only CGM-derived indices as input variables (referred to as CGM-only). To assess the multicollinearity among the input variables, we examined the VIF for each variable. We removed variables with the highest VIF one by one until all variables had VIF values less than 10. This process resulted in 14 variables for the Full model (Fig. S3A), 7 variables for the CGM- model (Fig. S3B), and 8 variables for the CGM-only model (Fig. S3C). AC\_Mean and AC\_Var were included within the 14 variables (Fig. S3A), and 8 variables (Fig. S3C), suggesting that AC\_Mean and AC\_Var exhibited relatively low multicollinearity with other indices.

We predicted %NC using these 14 (Fig. S3A, D), 7 (Fig. S3B, E), and 8 (Fig. S3C, F) variables. The Akaike Information Criterion (AIC) of the CGM-only model (Fig. S3F) was found to be lower than that of both the CGM- model (Fig. S3E) and the Full model (Fig. S3D), indicating the effectiveness of CGM-derived indices in predicting %NC. Inclusion of SBP, DBP, TG, LDL-C, and HDL-C as the input variables in the Full model and the CGM- model did not

change the result that the AIC of these models (Fig. S3G-J) was higher than that of the CGM-only model (Fig. S3F).

To verify the reproducibility of the relationship among the input variables, we compared VIF values calculated from this dataset (VIF<sub>t</sub>) with those from previously reported datasets (VIF<sub>p1</sub> (Sugimoto et al., 2023) and VIF<sub>p2</sub> (Hall et al., 2018)). These analyses revealed significant correlations between VIF<sub>t</sub> and VIF<sub>p1</sub> (Fig. S3K), as well as between VIF<sub>t</sub> and VIF<sub>p2</sub> (Fig. S3L). Furthermore, the VIF values for AC\_Mean and AC\_Var consistently tended to be low, underscoring the reproducibility of our observation that AC\_Mean and AC\_Var show relatively low multicollinearity with other CGM-derived indices.

## Supplementary Figures



**Figure S1. Relationship among clinical parameters.**

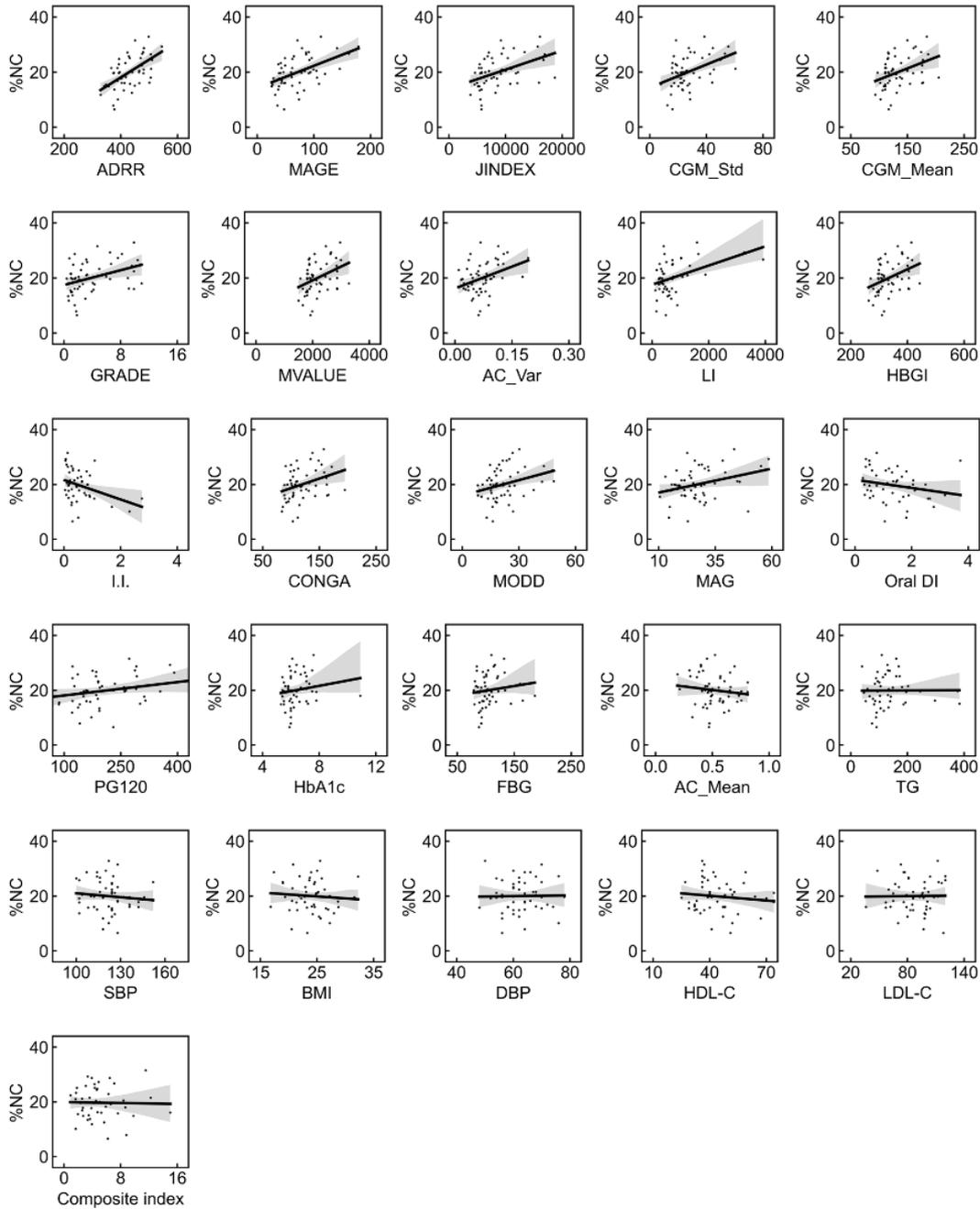
(A) A spring layout of the correlation network involving %NC (black), 14 CGM-derived

indices (red), 3 blood glucose level-related indices (magenta), 3 insulin sensitivity or secretion-related indices (blue), and 6 other clinical indices (green) obtained from a single blood test or physical measurement. Connections denote relationships with  $Q < 0.05$ . The width of the edges is proportional to the corresponding correlation coefficient.

(B) Scatter plots for AC\_Var versus CGM\_Mean (the left), and AC\_Var versus CGM\_Std (the right). Each point corresponds to the values for a single subject. Subjects were colored based on the value of %NC.  $r$  is Spearman's correlation coefficient, and the value in parentheses is 95% CI.

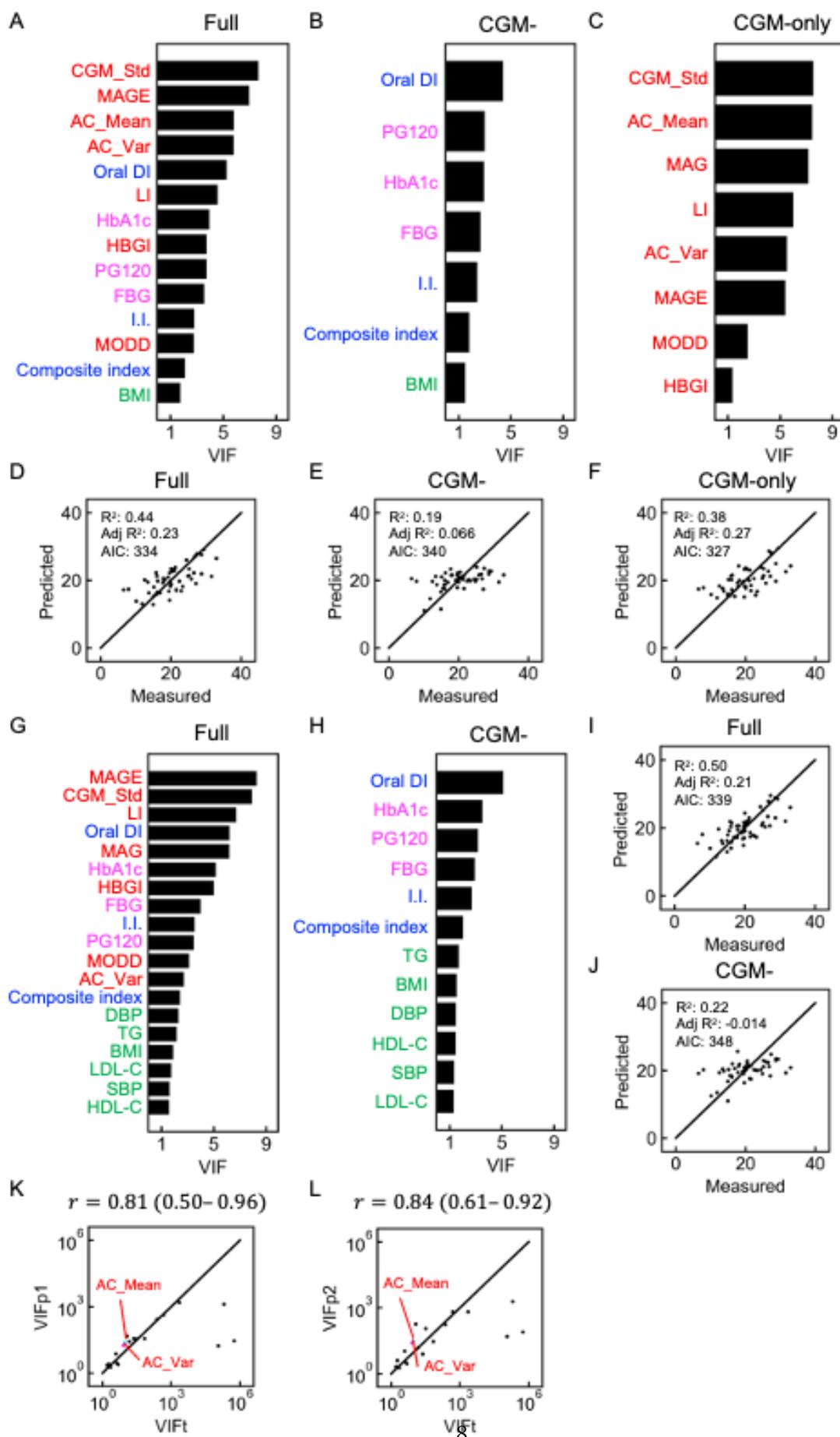
(C) The absolute values of Spearman's correlation coefficients between clinical parameters and %NC. Bars represent the 95% CIs.

(D) Scatter plot for SBP and LDL-C. Gray shaded areas indicate the range of values for high SBP ( $>140$  mmHg) or high LDL-C ( $>120$  mg/dL). Each point corresponds to the values for a single participant. Green circles, red cross marks, and blue squares indicate normal glucose tolerance, IGT, and T2DM, respectively.



**Figure S2. Relationship between the clinical indices and coronary plaque vulnerability (%NC).**

Scatter plots and fitted linear regression lines for each clinical index versus %NC. Each point corresponds to the values for a single subject. Gray shaded area indicates the 95% CI.



**Figure S3. Multiple linear regression analyses for predicting %NC.**

(A-C) VIF for each variable remaining after iteratively removing the variable with the highest VIF until all VIF values were less than 10. The input variables in (A) included the following 21 variables: BMI, FBG, HbA1c, PG120, I.I., composite index, oral DI, CGM\_Mean, CGM\_Std, CONGA, LI, JINDEX, HBGI, GRADE, MODD, MAGE, ADRR, MVALUE, MAG, AC\_Mean, and AC\_Var. Input variables in (B) included only those indices that excluded CGM-derived indices from the above 21 variables. The input variables of (C) included only the CGM-derived indices from the above 21 indices.

(D-F) Scatter plots for predicted %NC versus measured %NC. The black diagonal lines represent  $y=x$ . The model in (D) used the 14 variables shown in Figure 2A. The model in (E) used the 7 variables shown in Figure 2B. The model in (F) used the 8 variables shown in Figure 2C. Each point corresponds to the values for a single subject.

(G, H) VIF of each variable remaining after removing the variables with the highest VIF one by one until the VIF of all variables are less than 10. The input variables of (G) include the following 26 variables: BMI, SBP, DBP, TGs, LDL-C, HDL-C, FBG, HbA1c, PG120, I.I., composite index, oral DI, CGM\_Mean, CGM\_Std, CONGA, LI, JINDEX, HBGI, GRADE, MODD, MAGE, ADRR, MVALUE, MAG, AC\_Mean, and AC\_Var. Only indices that excluded CGM-derived indices from the above 26 indices were included as the input variables in (H).

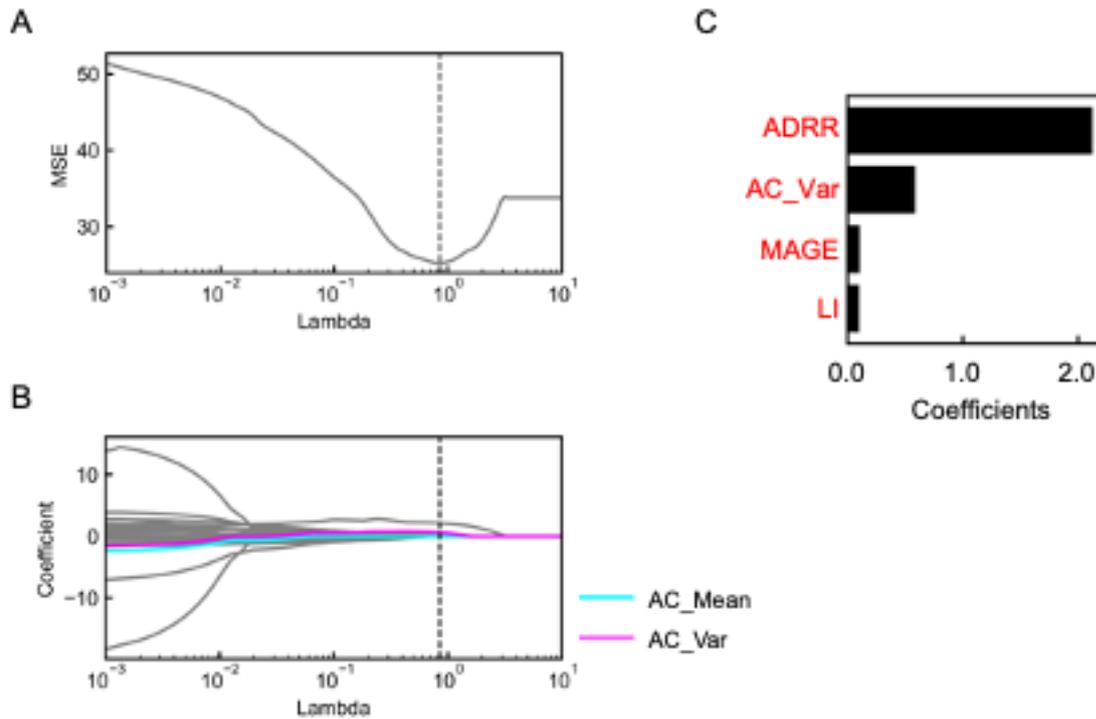
(I) Scatter plots for predicted %NC versus measured %NC. The model included the 19 variables shown in Figure S3G as the input variables. Each point corresponds to the values for a single subject.

(J) Scatter plots for predicted %NC versus measured %NC. The model included the 12 variables shown in Figure S3H as the input variables.

(K) Scatter plot of the VIF of the indices measured in this study (VIFt) versus that of indices

measured in a previous study (VIFp1) (Sugimoto et al., 2023).  $r$  is Spearman's correlation coefficient, and the value in parentheses is the 95% CI.

(L) Scatter plot of the VIF of the indices measured in this study (VIFt) versus that of the indices measured in a previous study (Hall et al., 2018) (VIFp2).

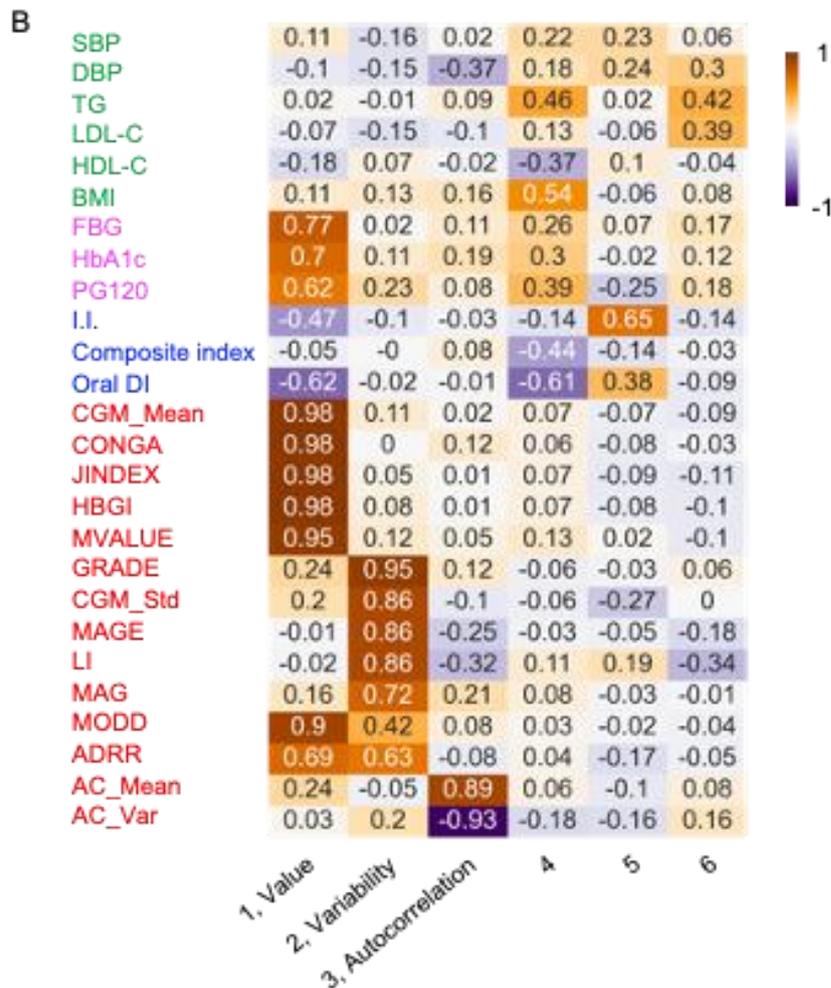
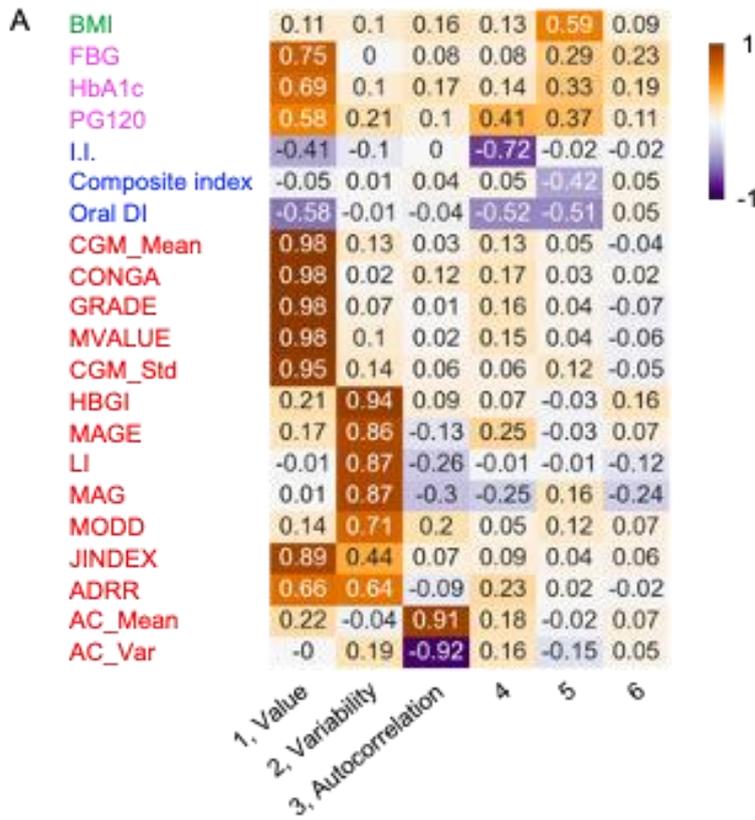


**Figure S4. LASSO regression analysis for predicting %NC including SBP, DBP, TGs, LDL-C, and HDL-C.**

(A) Relationship between regularization coefficients (lambda) and the MSE based on the leave-one-out cross-validation in predicting %NC. Dotted vertical line indicates the optimal lambda that provides the least MSE. The optimal lambda was 0.849.

(B) LASSO regularization paths along the lambda in predicting %NC. Cyan, magenta, and gray lines indicate the estimated coefficients of AC\_Mean, AC\_Var, and the other input variables, respectively. Dotted vertical line indicates the optimal lambda.

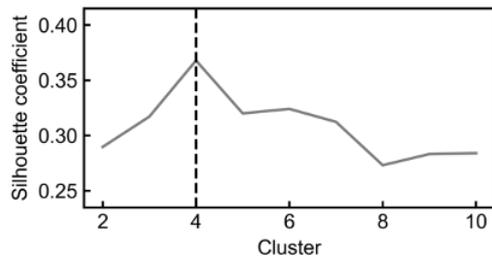
(C) Estimated coefficients with the optimal lambda. Only variables with non-zero coefficients are shown. Input variables include the following 26 variables: BMI, SBP, DBP, TGs, LDL-C, HDL-C, FBG, HbA1c, PG120, I.I., composite index, oral DI, CGM\_Mean, CGM\_Std, CONGA, LI, JINDEX, HBGI, GRADE, MODD, MAGE, ADRR, MVALUE, MAG, AC\_Mean, and AC\_Var.



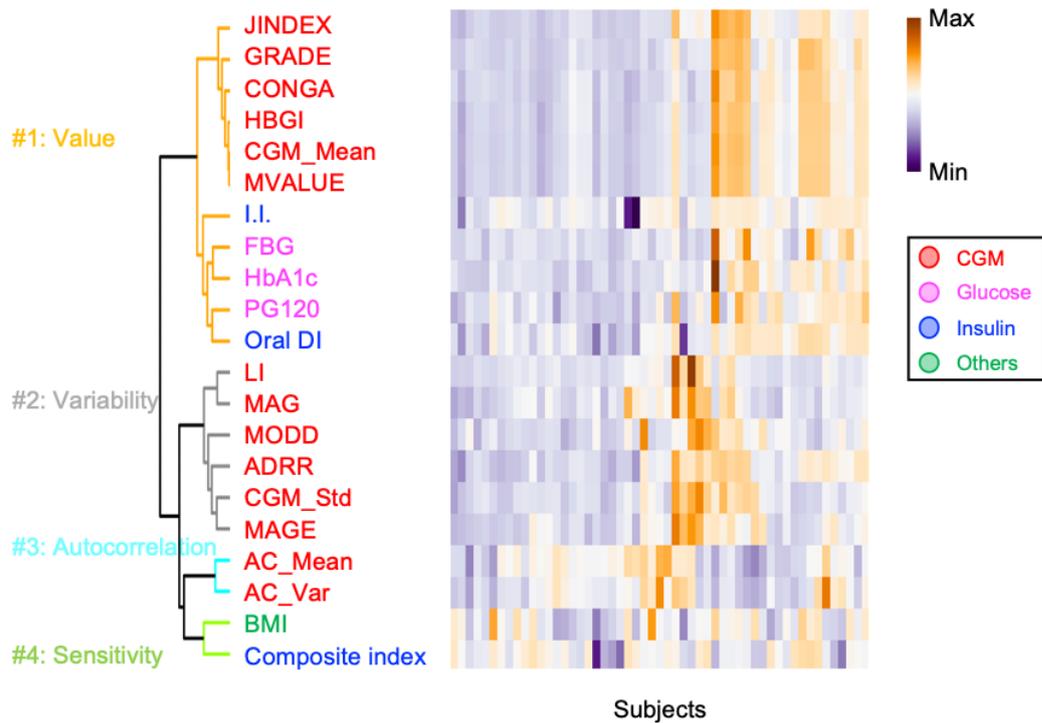
**Figure S5. Factor analyses of the clinical parameters.**

Factor analyses after orthogonal rotation. The values were based on the factor loadings. The columns represent each factor. The rows represent input indices. The analyses with the 21 variables (A), and the 26 variables (B).

A



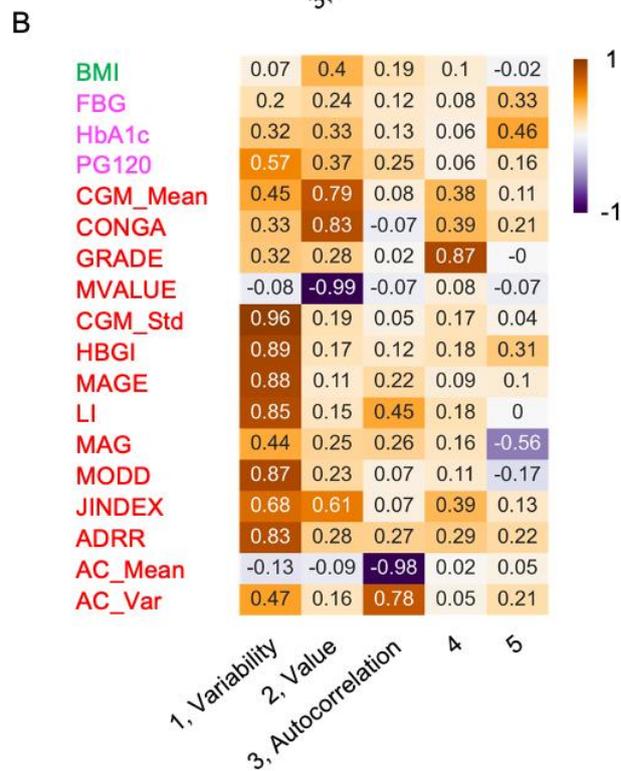
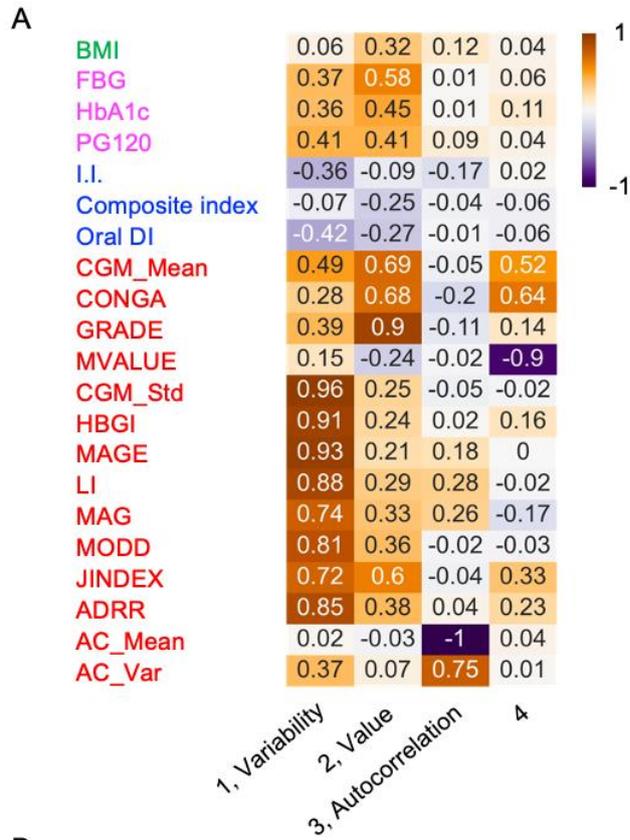
B



**Figure S6. Hierarchical clustering analysis of metabolic syndrome-related indices.**

(A) Relationship between the number of cluster and the silhouette coefficient. Dotted vertical line indicates the optimal cluster number which provides the best silhouette coefficient.

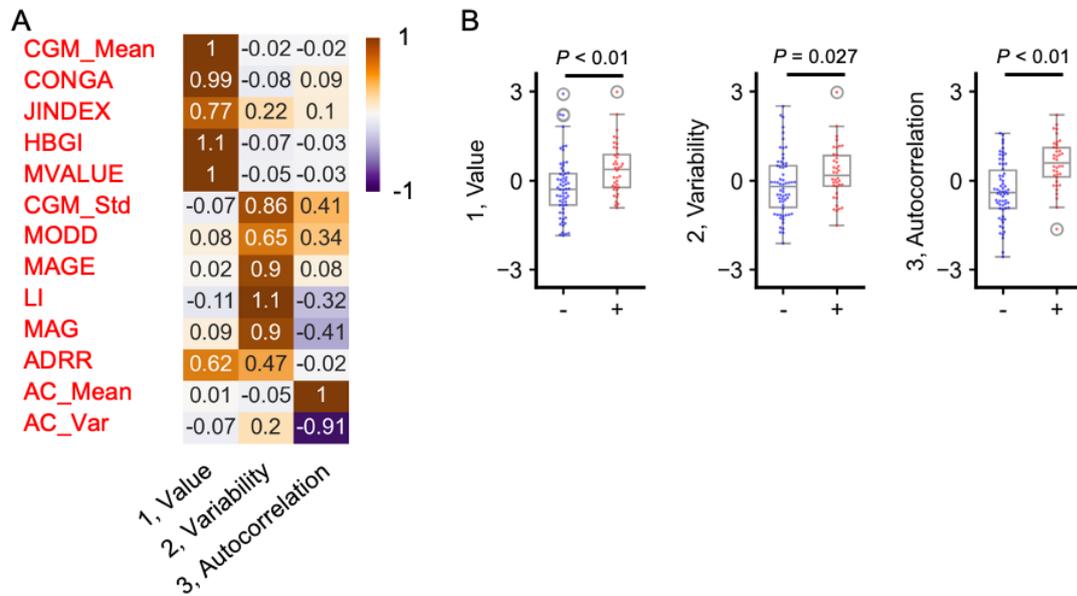
(B) Hierarchical clustering analysis of the standardized metabolic syndrome-related indices using Euclidean distance as a metric with the Ward method. The columns represent the standardized value of each index. The rows represent individual subjects. The indices are grouped and sorted according to their degree of relatedness.



**Figure S7. Factor analyses of the clinical parameters using the previous datasets.**

Factor analyses after orthogonal rotation. The values were based on the factor loadings. The

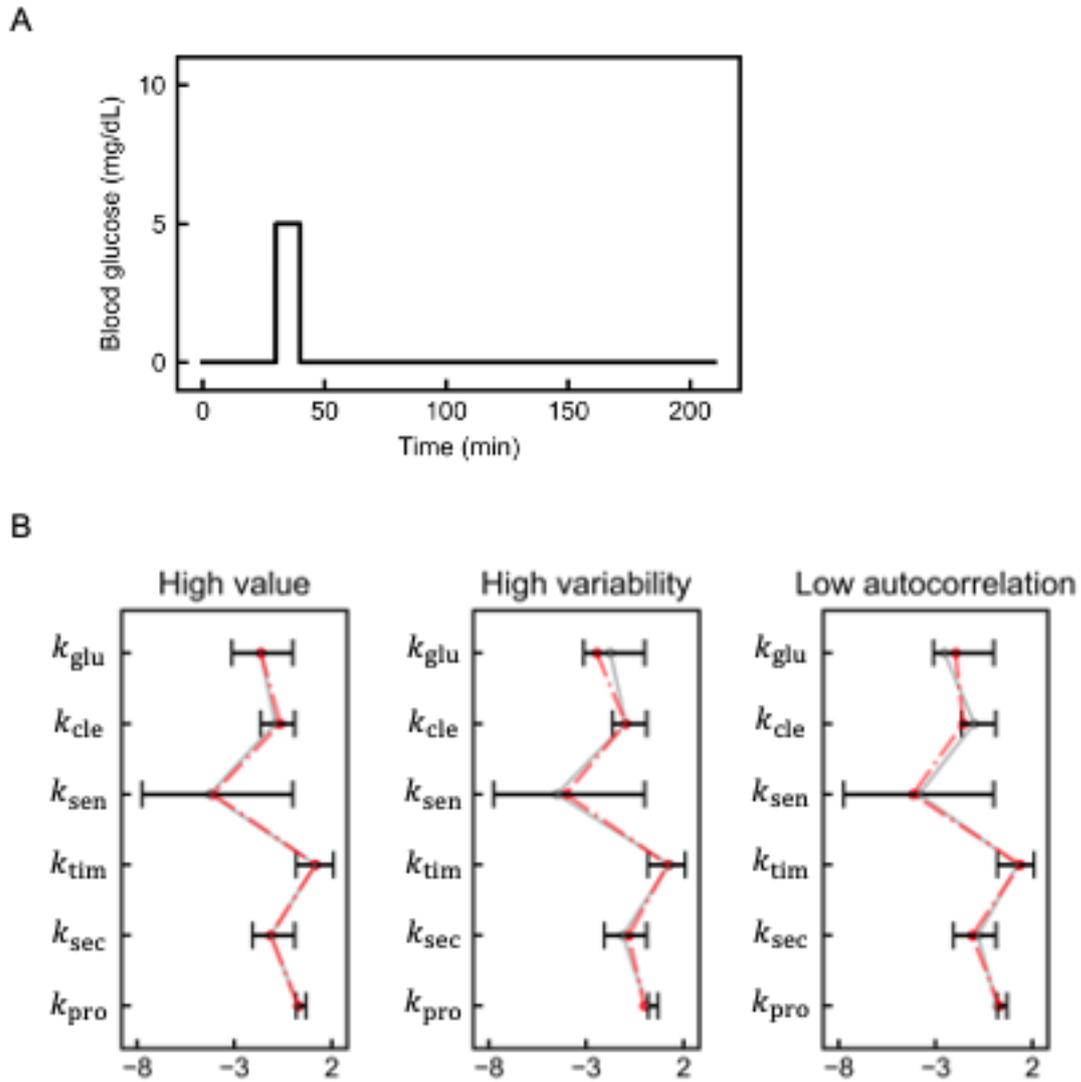
columns represent each factor. The rows represent input variables. The analyses used the Japanese data from a previous study (Sugimoto et al., 2023) (A), and the American data from a previous study (Hall et al., 2018) (B).



**Figure S8. Factor analysis of continuous glucose monitoring (CGM)-derived indices using a previously reported dataset.**

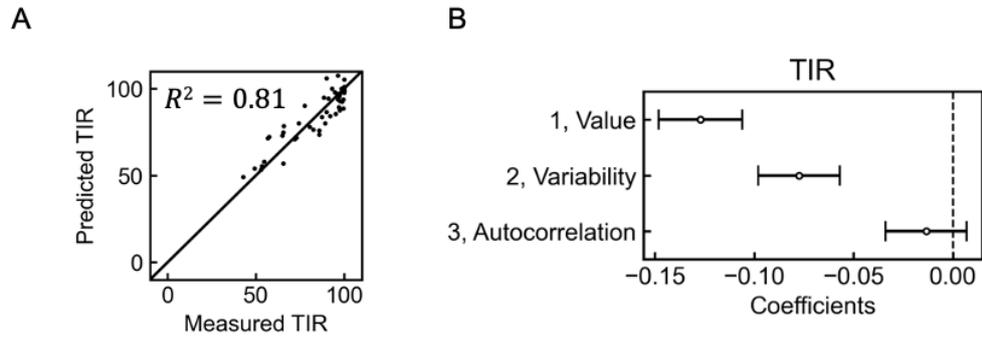
(A) Factor analysis of the CGM-derived indices. The heat map shows the factor loadings, with columns representing each factor and rows representing the input variables. The analyses used data from 100 Chinese individuals from a previous study (Zhao et al., 2023).

(B) Box plots comparing factors 1 (value), 2 (variability), and 3 (autocorrelation) between individuals without (-) and with (+) diabetic macrovascular complications. The boxes represent the interquartile range, with the median shown as a horizontal line. Mann–Whitney U tests were used to assess differences between groups, with  $P$  values  $< 0.05$  considered statistically significant.



**Figure S9. Infused glucose and parameters used in simulating glucose dynamics.**

(A) The amount of the external input of glucose  $f$  (see Methods). (B) The parameters used in the simulation. The values of these parameters shown in red and gray correspond to the color of the simulated glucose dynamics shown in Figure 4A. Bars indicate the range of values for NGT individuals (De Gaetano and Arino, 2000).



**Figure S10. Multiple regression analysis between Time in range (TIR) and factor scores in Figure 3.**

(A) Scatter plots for predicted TIR versus measured TIR. Each point corresponds to the values for a single subject.

(B) Bars represent the 95% CIs of the coefficients of the regression model.

**Table S1 Calculating formulae of the continuous glucose monitoring-derived indices.**

Name	Formulae	Variables
CGM_Mean	$\frac{1}{N} \sum_{n=1}^N G_n$	$G$ =glucose measured $N$ =total number of readings
CGM_Std	$\sqrt{\frac{1}{N-1} \sum_{n=1}^N (G_n - \bar{G})^2}$	$G$ =glucose measured $N$ =total number of readings
CONGA	$\sqrt{\frac{1}{k-1} \sum_{t=t_1}^{t_k} (D_t - \bar{D})^2}$ $\bar{D} = \frac{1}{k} \sum_{t=t_1}^{t_k} D_t, D_t = G_t - G_{t-60}$	$G$ =glucose measured $k$ =number of observations with an observation 60 min ago $t$ =time
LI	$\sum_{n=1}^{N-1} \frac{(G_n - G_{n+1})^2}{t_{n+1} - t_n}$	$G$ =glucose measured $N$ =total number of readings $t$ =time
JINDEX	$0.324 \times (\text{CGM\_Mean} + \text{CGM\_Std})^2$	
HBGI	$\frac{1}{N} \sum_{i=1}^N \text{rh}(x_i)$	$x$ = nonlinear transformation of glucose measured $N$ =total number of readings $\text{rh}$ =risk value associated with a high glucose
GRADE	$\text{median}(425 \times \{\log [\log (G_n)] + 0.16\}^2)$	$G$ =glucose measured
MODD	$\frac{1}{k} \sum_{t=t_1}^{t_k}  G_t - G_{t-1440} $	$G$ =glucose measured $k$ =number of observations with

		<p>an observation 24 h ago</p> <p><math>t</math> =time</p>
MAGE	$\sum \frac{\lambda}{x} \text{ if } \lambda > v$	<p><math>\lambda</math> =blood glucose changes from peak to nadir</p> <p><math>x</math> =number of observations</p> <p><math>v</math> =1 Std of mean glucose for 24 h period</p>
ADRR	$\frac{1}{N} \sum_{n=1}^N [\text{LR} + \text{HR}]$	<p><math>N</math> =total number of readings</p> <p>LR =risk value attributed to low glucose</p> <p>HR =risk value attributed to high glucose</p>
M-value	$\frac{1}{N} \sum_{t=t_1}^{t_k} \left  10 \log \frac{18G_t}{\text{IGV}} \right ^3$	<p><math>G</math> =glucose measured</p> <p><math>k</math> = number of observations</p> <p>IGV=ideal glucose value</p> <p><math>t</math> =time</p>
MAG	$\frac{1}{T} \sum_{n=1}^{N-1} (G_n - G_{n+1})$	<p><math>G</math> =glucose measured</p> <p><math>N</math> =total number of readings</p> <p><math>T</math> =total time</p>
AC_Mean	$\frac{1}{30} \sum_{l=1}^{30} \text{AC}_l$	<p>AC =autocorrelation of glucose</p> <p><math>l</math> =lag</p>
AC_Var	$\frac{1}{29} \sum_{l=1}^{30} (\text{AC}_l - \text{AC\_Mean})^2$	<p>AC =autocorrelation of glucose</p> <p><math>l</math> =lag</p>

## Supplementary References

- De Gaetano A, Arino O. 2000. Mathematical modelling of the intravenous glucose tolerance test. *J Math Biol* **40**:136–168.
- Hall H, Perelman D, Breschi A, Limcaoco P, Kellogg R, McLaughlin T, Snyder M. 2018. Glucotypes reveal new patterns of glucose dysregulation. *PLoS Biol* **16**:e2005143.
- Otowa-Suematsu N, Sakaguchi K, Komada H, Nakamura T, Sou A, Hirota Y, Kuroda M, Shinke T, Hirata K-I, Ogawa W. 2018. Comparison of the relationship between multiple parameters of glycemic variability and coronary plaque vulnerability assessed by virtual histology-intravascular ultrasound. *J Diabetes Investig* **9**:610–615.
- Sugimoto H, Hironaka K-I, Nakamura T, Yamada T, Miura H, Otowa-Suematsu N, Fujii M, Hirota Y, Sakaguchi K, Ogawa W, Others. 2023. Improved Detection of Decreased Glucose Handling Capacities via Novel Continuous Glucose Monitoring-Derived Indices: AC\_Mean and AC\_Var. *medRxiv* 2023–2009.
- Zhao Q, Zhu J, Shen X, Lin C, Zhang Y, Liang Y, Cao B, Li J, Liu X, Rao W, Wang C. 2023. Chinese diabetes datasets for data-driven machine learning. *Sci Data* **10**:35.