medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

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

Authors

Hikaru Sugimoto,[1] Ken-ichi Hironaka,[2] Tomoko Yamada,[3] Natsu Otowa-Suematsu,[3] Yushi Hirota,[3] Hiromasa Otake,[4] Ken-Ichi Hirata,[3] Kazuhiko Sakaguchi,[3] Wataru Ogawa,[3]* and Shinya Kuroda[1,2,5]*

Affiliations

[1]Department of Biochemistry and Molecular Biology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[2]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[3]Division of Diabetes and Endocrinology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kusunoki-cho 7-5-1, Chuo-ku, Kobe, Hyogo 650-0017, Japan

[4]Division of Cardiovascular Medicine, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kusunoki-cho 7-5-1, Chuo-ku, Kobe, Hyogo 650-0017, Japan

[5]Lead contact

*Corresponding authors: Wataru Ogawa, ogawa@med.kobe-u.ac.jp; Shinya Kuroda, skuroda@bs.s.u-tokyo.ac.jp

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

SUMMARY

Impaired glucose homeostasis leads to many complications, with coronary artery disease (CAD) being a major contributor to healthcare costs. However, current CAD screening methods lack efficacy. Here, we predicted CAD using easy-to-measure indices, including continuous glucose monitoring (CGM)-derived indices. We found that CGM-derived indices, particularly ADRR and AC_Var, exhibited stronger predictive capabilities for CAD compared to commonly used diabetes diagnostic indices such as fasting blood glucose (FBG), hemoglobin A1C (HbA1c), and plasma glucose level at 120 min during oral glucose tolerance tests (PG120). Factor analysis identified three distinct components underlying glucose dynamics – value, variability, and autocorrelation – each independently associated with CAD. Remarkably, ADRR was influenced by the first two components, and AC_Var was influenced by the third component. FBG, HbA1c, and PG120 were influenced only by the value component, making them insufficient for CAD prediction. CGM-derived indices reflecting the three components can outperform traditional diabetes diagnostic methods in CAD prediction. (150/150 words)

Keywords: Continuous glucose monitoring; oral glucose tolerance test; virtual histology-intravascular ultrasound; coronary artery disease

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

INTRODUCTION

Diabetes mellitus (DM) leads to a variety of complications, primarily caused by vascular damage.[1] Among the complications, coronary artery disease (CAD) accounts for a large fraction of the morbidity, mortality, and healthcare costs in patients with type 2 DM (T2DM).[2] Various prognostic models[1,3] and diagnostic markers[2] have been developed to predict CAD; however, screening of CAD can be ineffective, costly, or laborious.[2,4] More effective approaches for identifying individuals at high risk for complications using readily obtained clinical variables are warranted.

Blood glucose levels are among the readily obtained predictors of the complications.[5] The disrupted conditions of glucose dynamics seen in impaired glucose tolerance (IGT) and T2DM are partly characterized by high concentrations of blood glucose levels.[6] High concentrations of blood glucose levels have been defined as having high hemoglobin A1c (HbA1c) levels, fasting blood glucose (FBG) levels, and plasma glucose concentration at 120 min during the oral glucose tolerance test (OGTT) (PG120).[6] These indices, especially HbA1c, are associated with complications of T2DM.[7]

Beyond the absolute value of glucose concentration, glucose variability also contributes to the prognosis of the complications[5,6,8–11] and all-cause mortality.[12] Continuous glucose monitoring (CGM) can estimate short-term glycemic variability,[13] and is reportedly able to predict T2DM complications.[14] Standard deviation (Std) of glucose levels (CGM_Std), mean amplitude of glycemic excursion (MAGE), mean of daily difference (MODD), and continuous overlapping net glycemic action (CONGA) are established indices of glycemic variability, of which CGM_Std and MAGE are more highly correlated with coronary plaque properties.[15] Among glucose level-related indices, including HbA1c and FBG, MAGE is an independent determinant of coronary plaque instability.[16]

Other CGM-derived indices such as average daily risk ratio (ADRR), lability index (LI), J-index, mean absolute glucose (MAG), and glycemic risk assessment in diabetes

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

equation (GRADE) have also been developed.[17] We recently showed that AC_Mean and AC_Var, which are calculated from the autocorrelation function of glucose levels measured by CGM, can detect decreased abilities in glucose regulation that cannot be captured by FBG, HbA1c, or the other conventional CGM-derived indices.[18] The characteristics of glucose dynamics can also be estimated from insulin concentrations. The disposition index (DI), which is the product of insulin sensitivity and insulin secretion, reflects and predicts glycemic disability beyond FBG.[19] Several other glucose-related indices have also been reported; however, how these indices can be combined to deduce T2DM complications including CAD has yet to be established.

The objectives of this study were to determine (1) which clinical parameters are particularly useful in predicting CAD, (2) what factors underlie these indices, and (3) how these factors are associated with CAD. We investigated the characteristics of 14 CGM-derived indices: 12 relatively well-known CGM-derived indices[17] and 2 indices (AC_Mean and AC_Var) as well as OGTT-derived indices, and investigated the relationship between these parameters and coronary plaque vulnerability assessed by virtual histology-intravascular ultrasound (VH-IVAS), a strong predictor of coronary events. We showed that CGM-derived indices are particularly useful in the prediction. In addition, we showed that three components, namely, value, variability, and autocorrelation, underly blood glucose level-related indices, and that the three are independently associated with coronary plaque vulnerability.

RESULTS

CGM_Mean, CGM_Std, and AC_Var independently contribute to the prediction of 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

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

necrotic core to total plaque volume (%NC) (Figs. 1A, 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.[15] 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).

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. 1A). 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. 1A, blue, magenta, and green), suggesting that CGM-derived indices are effective in predicting %NC. Of note, this study enrolled individuals with well-controlled serum cholesterol and BP levels (Fig. S2). The weak correlations between cholesterol and BP-related indices and %NC (Fig. 1A) 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.

For a more comprehensive assessment of the association among clinical parameters, we constructed a correlation network connecting relationships with $Q < 0.05$ (Fig. 1B). 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. 1B). 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. 1B, C).

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

We previously reported that AC_Var, calculated from the autocorrelation function of glucose levels, can capture glucose handling capacities that cannot be captured by conventional CGM-derived indices including CGM_Mean and CGM_Std.[18] Based on the study, we hypothesized that AC_Var can identify individuals with high %NC independently from CGM_Mean and CGM_Std. To test this hypothesis, we conducted multiple regression analysis with CGM_Mean, CGM_Std, and AC_Var as input variables and %NC as the objective variable (Fig. 1D). The variance inflation factor (VIF), an index of multicollinearity, for CGM_Mean, CGM_Std, and AC_Var was 1.1, 1.1, and 1.0, respectively, suggesting that CGM_Mean, CGM_Std, and AC_Var had low multicollinearity with each other. The R^2 of the model that predicted %NC from the three indices was 0.36. CGM_Mean, CGM_Std, and AC_Var had statistically significant independent positive correlations with %NC (Fig. 1D), suggesting that CGM_Mean, CGM_Std, and AC_Var are independently associated with %NC. Notably, the R^2 of the model that predicted %NC from FBG, HbA1c, and PG120, which have been used as diagnostic markers of diabetes, was only 0.05 (Fig. 1D). Multivariate analyses including other indices also indicated that AC_Mean and AC_Var had relatively low multicollinearity with other CGM-derived indices (Supplementary text). Collectively, we conclude that CGM-derived indices are useful for predicting %NC compared to indices used for diabetes diagnosis, and that CGM_Mean, CGM_Std, and AC_Var independently contribute to the prediction of %NC.

CGM-derived indices, particularly ADRR, AC_Var, MAGE, and LI are effective in predicting coronary plaque vulnerability

To avoid overfitting and investigate which input variables are particularly useful in estimating %NC, we performed Least Absolute Shrinkage and Selection Operator (LASSO) regression (Fig. 2).[20] LASSO uses L1 regularization to produce models with fewer parameters and has been widely applied to feature selection in predictive modeling.[21] We

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

included BMI, FBG, HbA1c, OGTT-derived indices, and CGM-derived indices as the input variables. The leave-one-out cross-validation identified the optimal regularization coefficient, lambda, as 0.849 (Fig. 2A). At the lambda, the coefficients of ADRR, AC_Var, MAGE, and LI were estimated to be non-zero coefficients (Fig. 2B, C), suggesting that CGM-derived indices, particularly ADRR, AC_Var, MAGE, and LI, contribute to the prediction of %NC. Even with the inclusion of SBP, DBP, TG, LDL-C, and HDL-C as additional input variables, the results remained consistent, with the coefficients of ADRR, AC_Var, MAGE, and LI still estimated as non-zero coefficients (Fig. S4).

To further validate the results of the LASSO, we also performed Partial Least Squares (PLS) regression and investigated the VIP scores[22] (Fig. 2D). The VIP scores of ADRR, AC_Var, MAGE, and LI, which were estimated to be non-zero coefficients by LASSO, were higher than 1, indicating that these four variables especially contribute to the prediction of %NC.

Three components of dysglycemia – high value, high variability, and low autocorrelation – are associated with coronary plaque vulnerability

To identify factors underlying the clinical parameters and investigate how these factors are associated with %NC, we conducted exploratory factor analysis. Factor analysis reduces interrelated indices into a smaller set of underlying common factors, and has been employed to examine the interdependencies among various clinical parameters[23–25] and DM complications.[26]

To determine the optimal number of underlying factors, we investigated Bayesian information criterion (BIC) and minimum average partial (MAP) methods, indicating that five or six factors were appropriate, respectively. We first set the number of underlying factors as 5. Figure 3A indicates that FBG, HbA1c, PG120, I.I., oral DI, CGM_Mean, CONGA, HBGI, MVALUE, GRADE, JINDEX, and ADRR clustered as factor 1. Given that

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

most of these indices are related to the value of blood glucose concentration, factor 1 was labeled "value." CGM_Std, MAGE, LI, MAG, MODD, JINDEX, and ADRR clustered as factor 2. Given that these indices are related to glucose variability, factor 2 was labeled "variability." Given that the definition of JINDEX is based on the sum of CGM_Mean and CGM_Std, and that of ADRR is based on both high and low values of glucose, the result that JINDEX and ADRR clustered in both factors 1 and 2 is plausible. Given that autocorrelation-derived indices, AC_Mean and AC_Var, clustered as factor 3, factor 3 was labeled "autocorrelation." BMI, PG120, composite index, and oral DI clustered as factor 4. Factor 4 did not include any CGM-derived indices. Given that this combination of indices indicates a decrease in oral DI and associated increase in blood glucose due to decreased insulin sensitivity, factor 4 was labeled "sensitivity (without CGM)". PG120, I.I., oral DI, and MAG clustered as factor 5. Factor 5 did not have positive loadings of any CGM-derived indices. Given that this combination of the indices indicates a decrease in oral DI and associated increase in blood glucose due to decreased insulin secretion (I.I.), factor 5 was labeled "secretion (without CGM)." The cumulative percentages of the total variance of the factors were 39%, 60%, 70%, 75%, and 80%, respectively.

The validity of the factor analysis was assessed according to previous studies.[23,26] To evaluate the applicability of the factor analysis, the Kaiser-Meyer-Olkin (KMO) and Bartlett's spherical test were performed. The KMO test indicated that the value of the measure of sampling adequacy for this data was 0.64, and Bartlett's spherical test indicated that the variables were statistically significantly intercorrelated ($P < 0.01$), suggesting that this dataset was applicable for the factor analysis. To evaluate internal consistency, Cronbach's α (Fig. 3B) and item–total correlations were calculated for each factor. Cronbach's α was 0.97 for factor 1, 0.93 for factor 2, 0.90 for factor 3, 0.72 for factor 4, and 0.66 for factor 5; these values were larger than 0.65 (Fig. 3B), suggesting that the internal consistency was satisfactory. While Cronbach's α of factor 5 was relatively low, exclusion of

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

MAG increased the Cronbach's α to 0.84, indicating that the association between factor 5 and decrease in oral DI and associated increase in blood glucose due to decreased insulin secretion could be considered reliable. Item–total correlations ranged from 0.63 to 0.97 for factor 1, 0.72 to 0.94 for factor 2, 0.82 for factor 3, 0.54 to 0.76 for factor 4, and 0.37 to 0.86 for factor 5. With the exception of MAG, item–total correlations ranged from 0.84 to 0.91 for factor 5. The correlation coefficient of MAG was 0.37, which can be considered a modest correlation,[23] and the item–total correlations were generally reasonably strong in demonstrating reliability.

We also investigated a 6-factor solution (Fig. S5A). Factors 1, 2, and 3 could be interpreted as value, variability, and autocorrelation, respectively, similar to the 5-factor solution. Given that factor 6 had no factor loadings ≥ 0.3, we applied the 5-factor solution in the subsequent analysis. Furthermore, the inclusion of SBP, DBP, TG, LDL-C, and HDL-C into the input variables did not change the presence of the three components (value, variability, and autocorrelation) in glucose dynamics (Fig. S5B). Since we only included only individuals with well-controlled serum cholesterol and BP levels in this study, we applied the 5-factor solution without these indices (Fig. 3A) to the following analysis.

To further examine the stability of the results of the factor analysis, we also conducted hierarchical clustering analysis (Fig. S6). The optimal number of clusters was determined based on silhouette analysis. A large positive silhouette coefficient indicates that each cluster is compact and distinct from the others. The analysis indicated that the four clusters were appropriate (Fig. S6A). Clusters 1, 2, and 3 can be interpreted as value, variability, and autocorrelation, respectively (Fig. S6B). This confirmed the consistency of the clustering analysis with the factor analysis.

To investigate the reproducibility of the factor analysis, we conducted factor analyses using the previously reported datasets (Fig. S7). Factors that can be interpreted as value, variability, and autocorrelation were observed in both the Japanese[18] (Fig. S7A) and

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

American[27] (Fig. S7B) datasets, suggesting the reproducibility of glucose dynamics with these three components.

To investigate the association between these underlying factors and %NC, we investigated the correlation between the factor scores and %NC (Fig. 3C). The factor value and variability showed significant positive correlations with %NC, whereas autocorrelation showed a significant negative correlation. Factors 4 and 5, which were less related to the CGM-derived indices, showed weaker correlations with %NC. Collectively, we conclude that glucose dynamics has three components – value, variability, and autocorrelation – and that these three components are associated with %NC.

Overview of the three components of dysglycemia: high value, high variability, and low autocorrelation

We showed the existence of three components of glucose dynamics: value, variability, and autocorrelation. We also showed that %NC tended to increase with higher value, higher variability, and lower autocorrelation. To investigate whether each component can be varied separately, and to overview the characteristics of glucose dynamics with different values of the components, we simulated glucose fluctuations using a previously reported mathematical model[28] (Figs. 4, S8).

We could generate glucose fluctuations with almost the same standard deviation (Std) and AC_Var but with a different mean (Fig. 4A). Similarly, we could also simulate glucose fluctuations with almost the same mean and AC_Var but different Std, and with almost the same mean and Std but different AC_Var. These three components could be changed separately by changing the parameters within the range of values for NGT individuals (Fig. S8B). Individuals with higher AC_Var tended to have higher %NC (Fig. 1); however, comparing the glucose dynamics with higher and lower AC_Var, the maximum value of blood glucose was lower in individuals with higher AC_Var (Fig. 4A).

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

We also investigated the relationship between the three components and shapes of the glucose response curve after OGTT. Patterns of the glucose response curve after OGTT were heterogeneous, and four distinct patterns, denoted class 1–4, were previously identified (Fig. 4B).[29] The classes with a high mean did not necessarily have high Std or AC_Var (Fig. 4C). The classes with high Std did not necessarily have high mean and AC_Var (Fig. 4C), consistent with the result that mean, Std, and AC_Var had low multicollinearity with each other. Compared to class 2, only class 1 was lower in mean and Std, and was higher in AC_Var. Compared to class 2, only class 3 was higher in mean, Std, and AC_Var. Compared to class 2, only class 4 was higher in mean and Std, and lower in AC_Var. Collectively, the three components could characterize the previously reported four distinct patterns during OGTT.

Class 3 was characterized by normal FBG and PG120 values, but is reportedly associated with increased risk of diabetes and higher all-cause mortality rate, suggesting that subgroups at high risk may not be identified by investigating only FBG and PG120.[29] Std and AC_Var were high in class 3 (Fig. 4C), suggesting that high Std and high AC_Var indicate glycemic disability independent of PG120.

DISCUSSION

Here, we showed the existence of three distinct components in glucose dynamics, namely value, variability, and autocorrelation. We also showed that the higher the value, the higher the variability; and the lower the autocorrelation, the more vulnerable the coronary plaque. We previously reported that AC_Var, an index reflecting autocorrelation, can detect decreased abilities in glucose regulation independently of other CGM-derived indices including CGM_Mean and CGM_Std, which reflect value and variability components, respectively.[18] Diabetes diagnosis has been based on elevated FBG, PG120, and HbA1c levels. However, these indices primarily reflect only the value component of glucose

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

dynamics, and consequently the predictive performance of prediction model for %NC using FBG, PG120, and HbA1c was relatively modest compared to that of the model using all three components of glucose dynamics. This result is partially consistent with a previous notion that glucose dynamics include two components: amplitude and timing.[30] Collectively, for the effective prediction of glycemic control capacity and diabetic complications, a comprehensive examination of all three components (value, variability, and autocorrelation) is imperative.

We also showed that CGM-derived indices, especially ADRR and AC_Var, contribute to the prediction of %NC by using LASSO (Fig. 2). Given that the definition equation for ADRR is affected by both high and low concentrations of blood glucose,[17] it is likely affected by both glucose concentration values and glycemic variability. Factor analysis (Fig. 3) also showed that ADRR clustered as both factor 1 (value) and factor 2 (variability). Since three factors, value, variability, and autocorrelation, contribute independently to the prediction of the complication, it would be useful to examine ADRR, which is influenced by both value and variability, and AC_Var, which is influenced by autocorrelation, in predicting %NC with a minimal number of variables. Therefore, the result of the LASSO showing that ADRR and AC_Var are particularly effective in predicting %NC is consistent with the results of the factor analysis that the three components contribute to the prediction.

This study also provided evidence that autocorrelation can vary independently from the value and variability components by using simulated data. As shown in Figure 4, these three components could be varied independently by simply changing the parameters related to glucose regulation within the range of NGT individuals. In addition, simulated glucose dynamics indicated that even subjects with low autocorrelation did not necessarily have high maximum and minimum blood glucose levels. This study also indicated that these three components qualitatively corresponded to the four distinct glucose patterns observed after glucose administration, which were identified in a previous study.[29] Glycemic variability is

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

involved in T2DM complications by oxidative stress and endothelial dysfunction;[6,31] however, the reasons why the three components, especially autocorrelation, independently contribute to the prediction remain unknown. The underlying biological mechanisms and effects of the three components on living systems need to be investigated in future studies.

The current study had several limitations. LASSO and factor analysis indicated that CGM-related features were particularly important in predicting %NC. However, these results do not mean that other clinical parameters do not associate with T2DM complications, because we only included subjects with well-controlled serum cholesterol and BP levels in this study. Moreover, a previous study indicated that components of interday variability and hypoglycemia exist under CGM-derived indices,[25] but we did not observe these components. This may be because the number of T2DM subjects was small and the measurement period was short (2 days) in this study. Factor analyses of data with a longer measurement period that include more patients with T1DM and T2DM could change the results. In addition, although we used three different data sets with a total of 174 subjects, the sample size still tended to be small to examine the relationship among the variables examined in this study. Larger and prospective studies are required for a more accurate assessment of the variables in the prediction of the abnormality.

In conclusion, glucose dynamics has three components: value, variability, and autocorrelation. These three components are associated with coronary plaque vulnerability. CGM-derived indices reflecting these three components can be valuable predictive tools for T2DM complications, compared to conventional diabetes diagnostic markers reflecting only the value component. To facilitate this CGM-derived prediction, we created a web application that performs a multiple regression model with these three components as input variables (https://cgm-basedregression.streamlit.app/).

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

METHODS

Subjects and measurements

A previously reported dataset[15] was used in this study. The retrospective observational study was approved by the ethics committee of Kobe University Graduate School of Medicine (UMIN000018326; Kobe, Japan). Briefly, a 75-g OGTT, CGM, and percutaneous coronary intervention (PCI) were carried out on 53 participants. Of note, with a type I error of 0.05, a power of 0.8, and an expected correlation coefficient of 0.4, a sample size of 47 was required to detect a significant difference from zero in the correlation coefficient. The iPro2 CGM system (Medtronic, Northridge, CA, USA) was used, and data from the second and third days of CGM were analyzed. During PCI, VH-IVUS was carried out to assess the plaque components.

Participants aged 20–80 years with LDL-C levels < 120 mg/dL under statin administration or <100 mg/dL under other treatments for dyslipidemia, including lifestyle intervention, were included in this study. Participants with acute coronary syndrome, unsuitable anatomy for virtual VH-IVUS, poor imaging by VH-IVUS, hemodialysis, inflammatory disease, shock, low cardiac output, or concurrent malignant disease were excluded from this study.

Calculation of clinical indices

CGM-derived indices:

Fourteen CGM-derived indices were evaluated: 12 relatively well-known CGM-derived indices[17] and 2 indices (AC_Mean and AC_Var), which were indicated to capture glucose handling capacities independently of the 12 indices.[18] AC_Mean and AC_Var are the mean and variance of the autocorrelation coefficients at lags 1–30, respectively. CGM_Mean and CGM_Std indicate the mean value and Std of glucose levels measured by CGM, respectively. CONGA, LI, JINDEX, HBGI, GRADE, MODD, MAGE, ADRR, MVALUE, and MAG

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

were calculated using EasyGV software.[17] The calculating formulae of these indices are shown in Table S1. They were calculated from the CGM data measured every 5 min.

OGTT-derived indices:

Three OGTT-derived indices were calculated as previously described.[15] I.I. indicates insulin secretion, which can be calculated from the ratio of the increment of immunoreactive insulin (IRI) to that of plasma glucose at 30 min after onset of the OGTT. Composite index indicates insulin sensitivity, which can be calculated from fasting plasma glucose, fasting IRI, mean blood glucose levels, and mean serum IRI concentrations during the OGTT. Oral DI was calculated from the product of composite index and the ratio of the area under the insulin concentration curve from 0 to 120 min to that for plasma glucose from 0 to 120 min, without using the data measured at 90 min, in the OGTT.

VH-IVUS-derived index:

VH-IVUS was carried out using the Eagle Eye Platinum 3.5-Fr 20-MHz catheter (Volcano, Rancho Cordova, CA, USA), as previously described.[15] The %NC, a widely used parameter of plaque vulnerability, was evaluated.

Prediction models and statistical analyses

In this study, we conducted multiple linear regression, LASSO regression, and PLS regression. The input variables in these models included 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. In conducting these models, z-score normalization on each input variable was performed.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

The predictive performance of multiple linear regression was evaluated by the coefficient of determination (R^2), the adjusted coefficient of determination (Adj R^2), and AIC. The multicollinearity of the input variables was estimated by VIF. LASSO regression is a kind of linear regression with L1 regularization.[20] The optimal regularization coefficient, lambda, was based on leave-one-out cross-validation and mean-squared error. The importance of the input variables in predicting %NC was evaluated by the VIP scores[22] that were generated from PLS regression. These models were conducted using scikit-learn, a python-based tool kit.

Relationships among indices were also evaluated using Spearman's correlation coefficients (r), and the correlation coefficients were reported with 95% CIs through bootstrap resampling. The number of resamples performed to form the distribution was set at 10000. Benjamini–Hochberg's multiple comparison test was also performed with a significance threshold of $Q < 0.05$.

Factor analysis and hierarchical clustering analysis

The intercorrelations of the clinical parameters and their associations with %NC were assessed using exploratory factor analyses and hierarchical clustering analyses. We followed the previously reported approach[23,24,26,32] with some modifications in conducting our exploratory factor analyses. BIC and MAP methods were used to determine the number of underlying factors. Variables with factor loadings of ≥0.30 were used in interpretation. To improve the interpretation, orthogonal (varimax) rotation was used. To evaluate the applicability of the factor analysis, KMO and Bartlett's spherical test were performed. To evaluate internal consistency of each factor, Cronbach's α and item–total correlations were calculated. The association of the factor scores with %NC was assessed using Spearman's correlation.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Hierarchical clustering analysis was conducted using a method that combines a Euclidean distance measure and Ward linkage. It was adopted after Z score normalization. I.I., composite index, oral DI, and AC_Mean were inverted negatively so that the value of indices increased in subjects with abnormalities. The quality of the hierarchical clustering analysis was evaluated based on silhouette analysis.[33]

Mathematical model used for simulating the characteristics of glucose dynamics

In simulating the characteristics of glucose dynamics, we used a simple and stable model,[28] which can be written as follows:

$$\frac{dG}{dt} = -k_{\mathrm{glu}}G - k_{\mathrm{sen}}IG + k_{\mathrm{pro}} + f$$

$$\frac{dI}{dt} = \frac{k_{\mathrm{sec}}}{k_{\mathrm{tim}}}\int_{t-k_{\mathrm{tim}}}^{t} G\,ds - k_{\mathrm{cle}}I$$

where the variables G and I denote blood glucose and insulin concentrations, respectively. We simulated 240 min profiles of G, and calculated the mean, Std, and AC_Var of G. The parameters were changed within the range participants could take.[28] Five mg/dL/min glucose was applied for 10 min at 30 min as the external input of glucose f (Fig. S8A).

Characterization of glucose patterns during the OGTT

We investigated the characteristics of previously reported glucose patterns during the OGTT.[29] In the study, 5861 subjects without diabetes underwent the OGTT with measurements of glucose levels at three time points (0, 30, and 120 min), and four distinct glucose patterns associated with long-term outcomes including diabetes onset, CVD, and all-cause mortality rate were identified. For the calculation of mean, Std, and AC_Var of glucose levels, each time point was linearly imputed. Here, AC_Var was calculated from the autocorrelation function at lags 1–20, as we had glucose data available for only 2 h after the OGTT.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

RESOURCE AVAILABILITY

Lead contact

For any additional information and requests regarding resources and reagents, please contact the lead contact, Shinya Kuroda (skuroda@bs.s.u-tokyo.ac.jp).

Data and code availability

The CGM data that support the findings of this study are available from the GitHub repository (https://github.com/HikaruSugimoto/CGM_regression_app). The code that calculates AC_Mean and AC_Var and that performs regression analysis with CGM-derived indices as input variables are available from the repository (https://github.com/HikaruSugimoto/CGM_regression_app) and the web application (https://cgm-basedregression.streamlit.app/).

Materials availability

The study did not generate any new material.

Acknowledgments

Personal Thanks. We thank our laboratory members for critically reading this manuscript

Author Contributions. H.S. analyzed the data. H.S., K.H., T.Y., N.O.S., Y.H., H.O., K.H., K.S., W.O., and S.K. wrote the manuscript. W.O. and S.K. supervised the study.

Conflict of Interest. The authors have no conflicts of interest to declare.

Funding and Assistance. This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (JP21H04759), CREST, the Japan Science and Technology Agency (JST) (JPMJCR2123), and The Uehara Memorial Foundation.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Figures

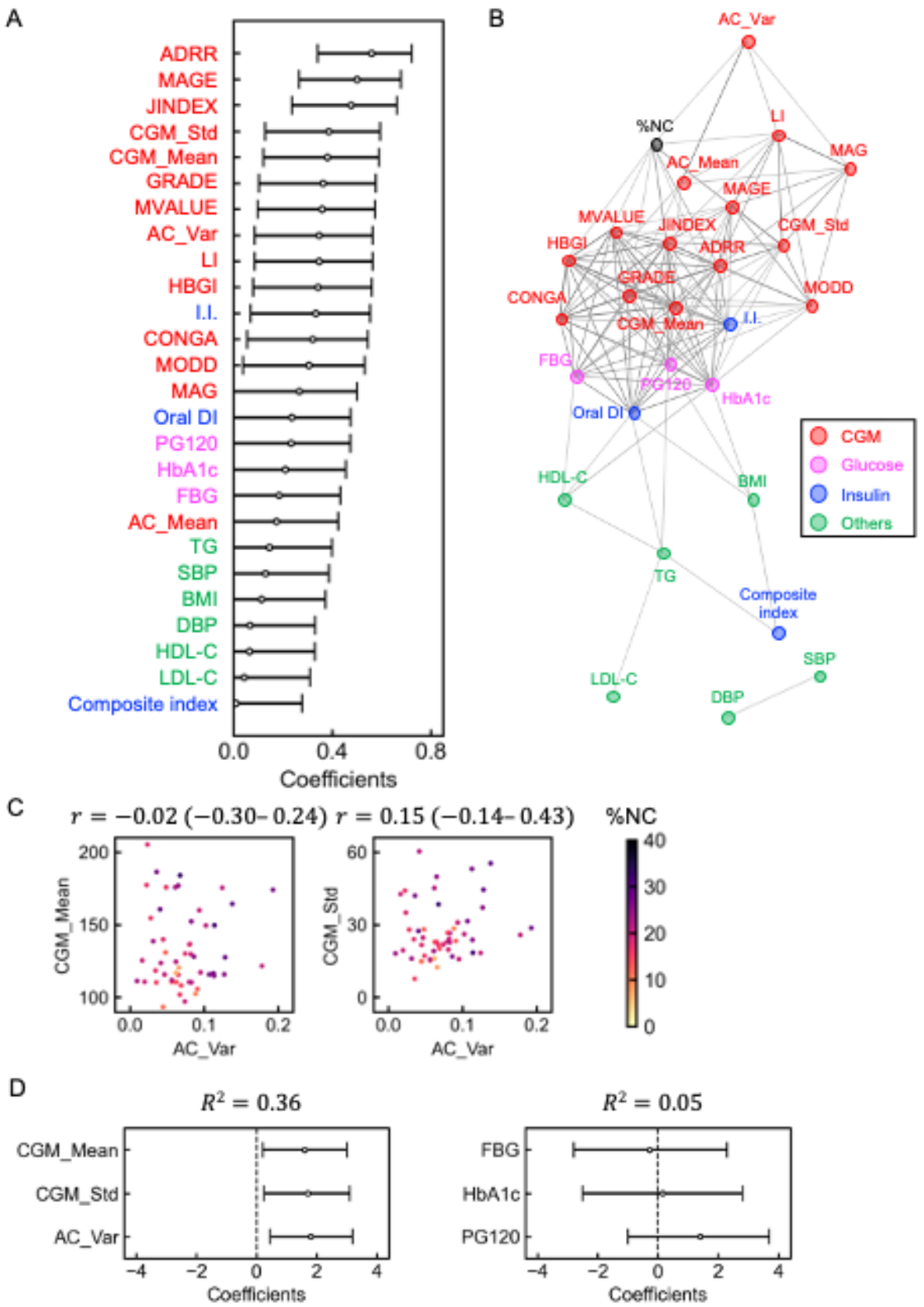

Figure 1. Relationship among clinical parameters.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

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

(B) 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.

(C) 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.

(D) Multiple regression analysis between %NC and CGM_Mean, CGM_Std, and AC_Var (the left). That between %NC and FBG, HbA1c, and PG120 (the right). Bars represent the 95% CIs of the coefficients.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

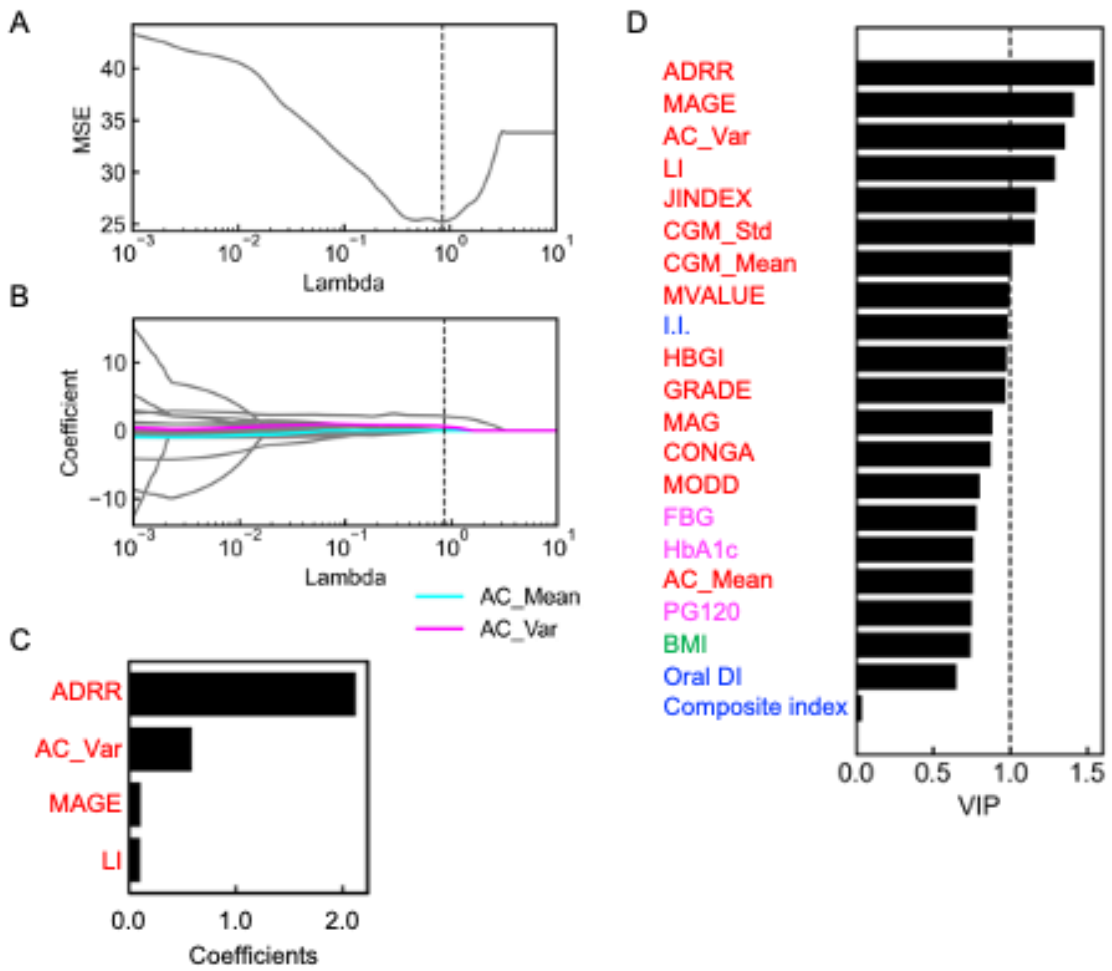

Figure 2. LASSO and PLS regression analyses for predicting %NC.

(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, which 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 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.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

(D) VIP generated from the PLS regression predicting %NC. Variables with a VIP ≥ 1 (the dotted line) were considered to significantly contribute to the prediction.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

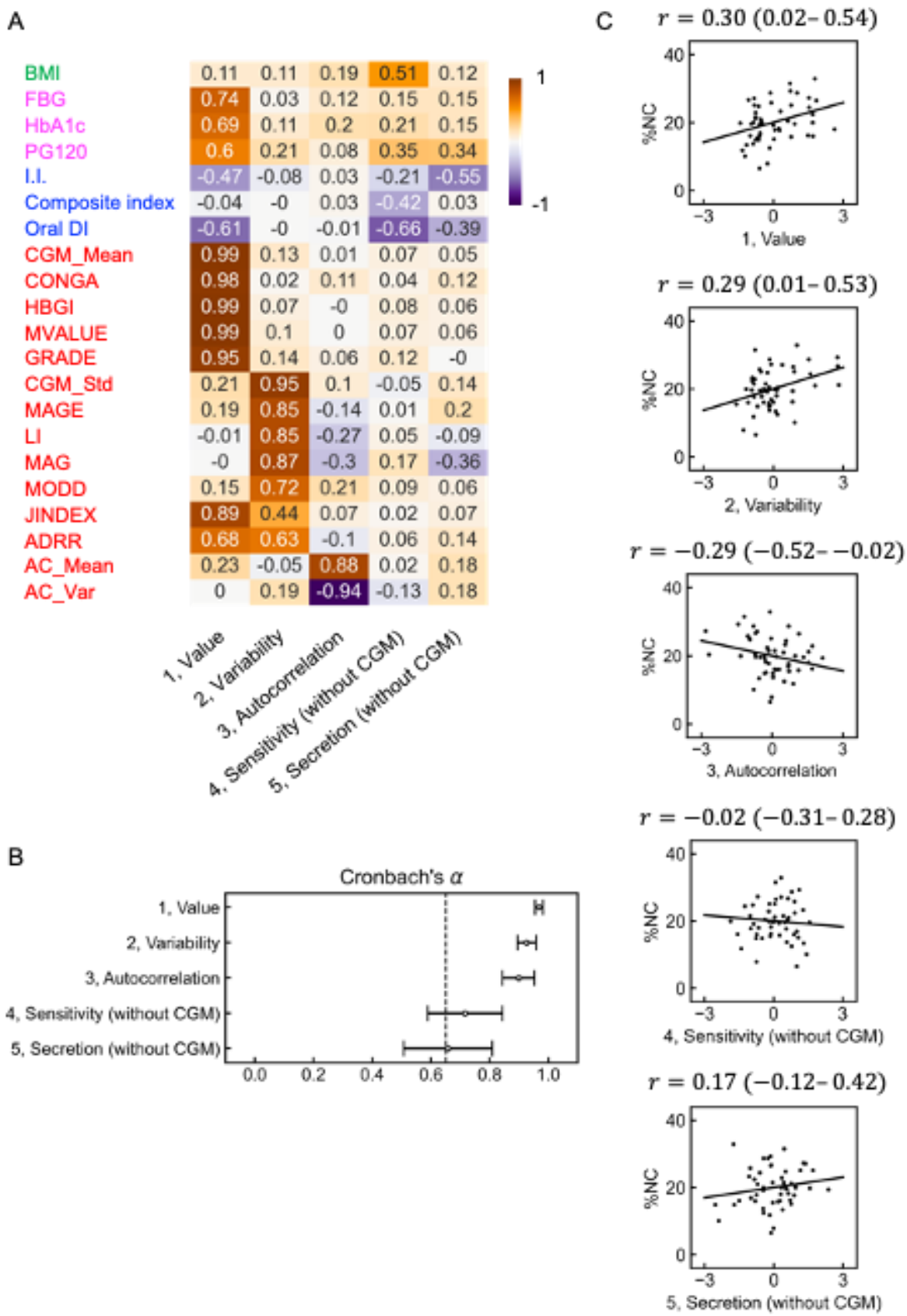

Figure 3. Factor analysis of the clinical parameters.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

(A) Factor analysis after orthogonal rotation. The values and colors were based on the factor loadings. The columns represent each factor. The rows represent input indices.

(B) Cronbach's α for each factor. Bars represent the 95% CI.

(C) Scatter plots and fitted linear regression lines for factor scores versus %NC. Each point corresponds to the values for a single subject. *r* is Spearman's correlation coefficient, and the value in parentheses is the 95% CI.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

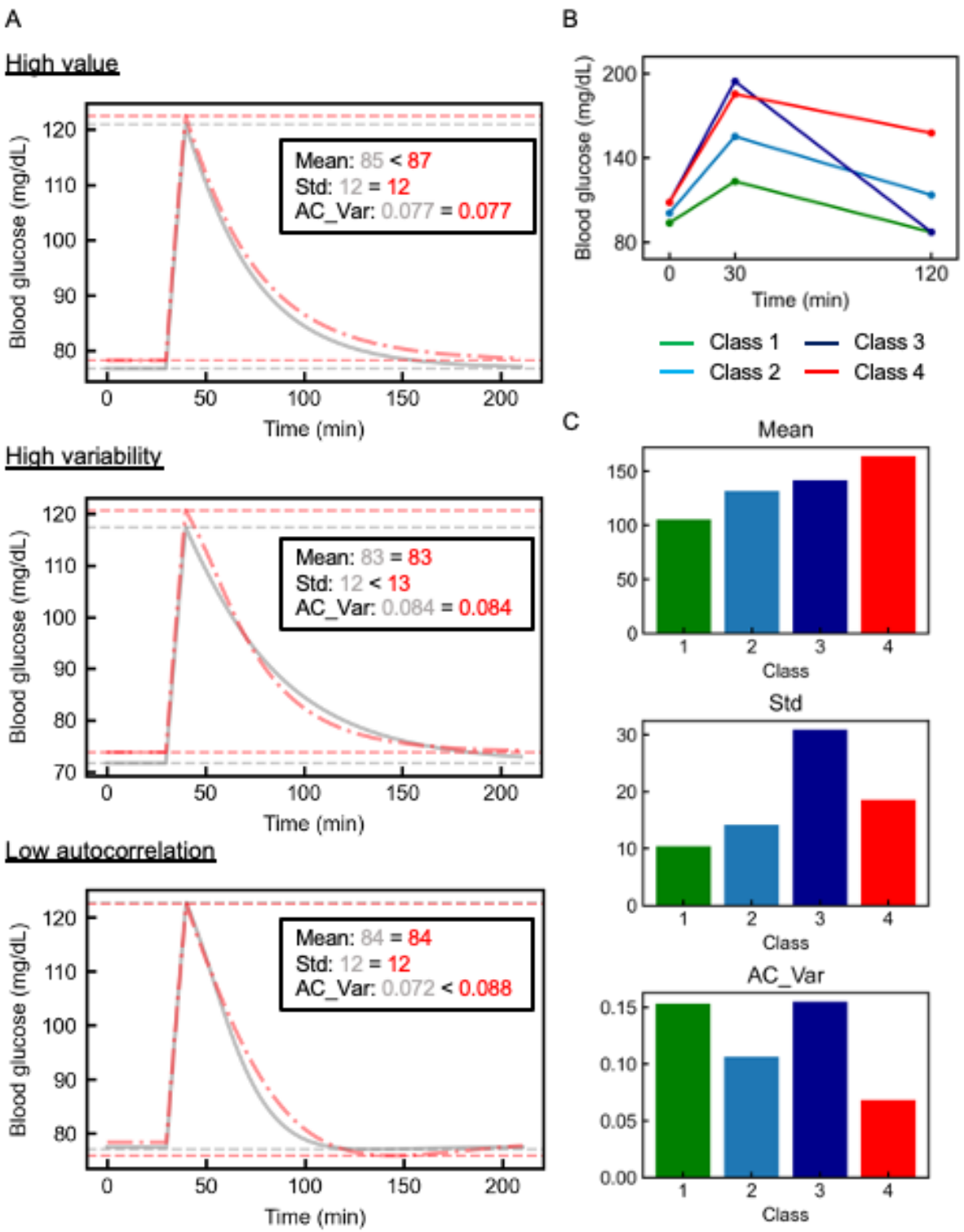

Figure 4. Overview of the three components of glucose dynamics.

(A) 240 min simulated glucose concentration. The colors of the line are based on the mean value (Mean), Std, and AC_Var of the simulated blood glucose. Red and gray dotted horizontal lines indicate the minimum or maximum values of blood glucose, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

(B) Previously reported patterns of blood glucose during the OGTT.[29] Green, class 1; light blue, class 2; dark blue, class 3; red, class 4.

(C) Mean, Std, and AC_Var of the glucose during the OGTT. Colors are based on the class shown in Figure 4B.

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

References

1. Fiarni, C., Sipayung, E. M. & Maemunah, S. Analysis and Prediction of Diabetes Complication Disease using Data Mining Algorithm. *Procedia Comput. Sci.* **161**, 449–457 (2019).
2. Bax, J. J. *et al.* Screening for coronary artery disease in patients with diabetes. *Diabetes Care* **30**, 2729–2736 (2007).
3. Ravaut, M. *et al.* Predicting adverse outcomes due to diabetes complications with machine learning using administrative health data. *NPJ Digit Med* **4**, 24 (2021).
4. Young, L. H. *et al.* Cardiac outcomes after screening for asymptomatic coronary artery disease in patients with type 2 diabetes: the DIAD study: a randomized controlled trial. *JAMA* **301**, 1547–1555 (2009).
5. Psoma, O., Makris, M., Tselepis, A. & Tsimihodimos, V. Short-term Glycemic Variability and Its Association With Macrovascular and Microvascular Complications in Patients With Diabetes. *J. Diabetes Sci. Technol.* 19322968221146810 (2022).
6. Monnier, L., Colette, C. & Owens, D. R. Glycemic variability: the third component of the dysglycemia in diabetes. Is it important? How to measure it? *J. Diabetes Sci. Technol.* **2**, 1094–1100 (2008).
7. Selvin, E. *et al.* Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. *N. Engl. J. Med.* **362**, 800–811 (2010).
8. Gorst, C. *et al.* Long-term Glycemic Variability and Risk of Adverse Outcomes: A Systematic Review and Meta-analysis. *Diabetes Care* **38**, 2354–2369 (2015).
9. Zhou, J. J., Schwenke, D. C., Bahn, G., Reaven, P. & VADT Investigators. Glycemic Variation and Cardiovascular Risk in the Veterans Affairs Diabetes Trial. *Diabetes Care* **41**, 2187–2194 (2018).
10. Gerbaud, E. *et al.* Glycemic Variability Is a Powerful Independent Predictive Factor of Midterm Major Adverse Cardiac Events in Patients With Diabetes With Acute Coronary Syndrome. *Diabetes Care* **42**, 674–681 (2019).
11. Su, G. *et al.* Association of glycemic variability and the presence and severity of coronary artery disease in patients with type 2 diabetes. *Cardiovasc. Diabetol.* **10**, 19 (2011).
12. Cai, J. *et al.* Impact of the complexity of glucose time series on all-cause mortality in patients with type 2 diabetes. *J. Clin. Endocrinol. Metab.* **108**, 1093–1100 (2023).
13. Service, F. J. Glucose variability. *Diabetes* **62**, 1398–1404 (2013).
14. Tang, X. *et al.* Glycemic variability evaluated by continuous glucose monitoring system is associated with the 10-y cardiovascular risk of diabetic patients with well-controlled HbA1c. *Clin. Chim. Acta* **461**, 146–150 (2016).

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

15. Otowa-Suematsu, N. *et al.* 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 (2018).
16. Okada, K. *et al.* Association between blood glucose variability and coronary plaque instability in patients with acute coronary syndromes. *Cardiovasc. Diabetol.* **14**, 111 (2015).
17. Hill, N. R. *et al.* Normal reference range for mean tissue glucose and glycemic variability derived from continuous glucose monitoring for subjects without diabetes in different ethnic groups. *Diabetes Technol. Ther.* **13**, 921–928 (2011).
18. Sugimoto, H. *et al.* Improved Detection of Decreased Glucose Handling Capacities via Novel Continuous Glucose Monitoring-Derived Indices: AC_Mean and AC_Var. *medRxiv* 2023−2009 (2023).
19. Utzschneider, K. M. *et al.* Oral disposition index predicts the development of future diabetes above and beyond fasting and 2-h glucose levels. *Diabetes Care* **32**, 335–341 (2009).
20. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.* **58**, 267–288 (1996).
21. Wei, H. *et al.* Environmental chemical exposure dynamics and machine learning-based prediction of diabetes mellitus. *Sci. Total Environ.* **806**, 150674 (2022).
22. Wold, S., Sjöström, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometrics Intellig. Lab. Syst.* **58**, 109–130 (2001).
23. Cappelleri, J. C., Gerber, R. A., Kourides, I. A. & Gelfand, R. A. Development and factor analysis of a questionnaire to measure patient satisfaction with injected and inhaled insulin for type 1 diabetes. *Diabetes Care* **23**, 1799–1803 (2000).
24. Oh, J.-Y., Hong, Y. S., Sung, Y.-A. & Barrett-Connor, E. Prevalence and factor analysis of metabolic syndrome in an urban Korean population. *Diabetes Care* **27**, 2027–2032 (2004).
25. Augstein, P. *et al.* Q-Score: development of a new metric for continuous glucose monitoring that enables stratification of antihyperglycaemic therapies. *BMC Endocr. Disord.* **15**, 22 (2015).
26. Guo, W., Zhou, Q., Jia, Y. & Xu, J. Cluster and Factor Analysis of Elements in Serum and Urine of Diabetic Patients with Peripheral Neuropathy and Healthy People. *Biol. Trace Elem. Res.* **194**, 48–57 (2020).
27. Hall, H. *et al.* Glucotypes reveal new patterns of glucose dysregulation. *PLoS Biol.* **16**, e2005143 (2018).
28. De Gaetano, A. & Arino, O. Mathematical modelling of the intravenous glucose tolerance test. *J. Math. Biol.* **40**, 136–168 (2000).

medRxiv preprint doi: https://doi.org/10.1101/2023.11.21.23298816; this version posted November 22, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

29. Hulman, A. *et al.* Glucose patterns during an oral glucose tolerance test and associations with future diabetes, cardiovascular disease and all-cause mortality rate. *Diabetologia* **61**, 101–107 (2018).
30. Cobelli, C. & Facchinetti, A. Yet Another Glucose Variability Index: Time for a Paradigm Change? *Diabetes Technol. Ther.* **20**, 1–3 (2018).
31. Ceriello, A. *et al.* Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in normal and type 2 diabetic patients. *Diabetes* **57**, 1349–1354 (2008).
32. Lakka, H.-M. *et al.* The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA* **288**, 2709–2716 (2002).
33. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).