# The measurement and consequences of daily glucose variability

**Document status and date:**
Published: 01/01/2021

**DOI:**
10.26481/dis.20211020yf

**Document Version:**
Publisher's PDF, also known as Version of record

# The Measurement and Consequences of Daily Glucose Variability

*A meaningful addition to the mean?*

# The Measurement and Consequences of Daily Glucose Variability

## *A meaningful addition to the mean?*

**Promotores**

Prof.dr. C.D.A. Stehouwer
Prof.dr. M.C.G.J. Brouwers
Prof.dr. N.C. Schaper

**Beoordelingscommissie**

Prof.dr. M.H. Prins (voorzitter)
Prof.dr.ir. E.E. Blaak
Prof.dr. C.J.J. Tack, Radboud Universiteit
Prof.dr. J.H. de Vries, Universiteit van Amsterdam
Dr. T.T. van Sloten

# Table of contents

# Chapter 1

General introduction and outline of the thesis

## Diabetes as a global health concern

Diabetes mellitus is an increasingly prevalent worldwide health concern[1]. Current global estimates indicate that diabetes affects 463 million individuals, a figure that is expected to rise to 700 million by 2045[2]. Diabetes is, by definition, characterized by hyperglycemia (i.e., elevated blood glucose concentrations)[3], and comprises several distinct chronic and metabolic diseases, the most common of which are type 1 (5-10%) and type 2 diabetes (90-95%)[4]. Type 1 diabetes is caused by autoimmune destruction of the insulin-producing beta cells of the endocrine pancreas, which results in absolute insulin deficiency[4]. Its onset is relatively rapid and commonly occurs at a young age[4]. By contrast, type 2 diabetes is predominantly a consequence of the pancreatic inability to overcome peripheral insulin resistance (i.e., relative insulin deficiency), which is caused by the detrimental effects of obesity and an unhealthy lifestyle in conjunction with a genetic predisposition[1,4]. Type 2 diabetes typically affects middle-aged and elderly individuals, and develops insidiously over a period of several years[1,4]. The intermediate phase during which the glucose metabolism status (GMS) transitions from normal glucose metabolism (NGM) to type 2 diabetes is commonly called prediabetes[4].

Individuals with diabetes have a twofold higher risk of all-cause mortality, mostly as a consequence of cardiovascular disease (CVD)[5], compared to their diabetes-free peers[6-8]. In addition, diabetes is associated with a multitude of other morbidities, which include the classic microvascular complications (i.e., nephropathy, neuropathy, and retinopathy)[9,10], as well as heart failure[11], cancer[12], non-alcoholic fatty liver disease[13], depression[14], and dementia[14]. Diabetes and its co-morbidities not only have debilitating effects at an individual level[15], but also negatively impact society through their large economic burden[16,17]. Hence, treatment of diabetes, which commonly entails a combination of lifestyle interventions and medication, is primarily focused on prevention of the development or progression of the aforementioned complications[18]. In addition to blood pressure control, lipid profile management, and smoking cessation[9,19], one of the main ways to achieve this is by improving glycemic control[3].

## Glycemic control: attacking the average

Glycated hemoglobin (HbA$_{1c}$) is currently the preferred method to quantify glycemic control[3]. It indirectly reflects average blood glucose levels over a period of two to three months[4]. In healthy individuals, HbA$_{1c}$ values are normally below 40 mmol/mol[4]. The Diabetes Control and Complications Trial (DCCT) was the first to convincingly show—in individuals with type 1 diabetes—that targeted lowering of HbA$_{1c}$ from on average ~74 to ~53 mmol/mol substantially improved microvascular complication risk[20]. Since then, several other landmark studies further reported on the benefits of HbA$_{1c}$-based glycemic control with regard to microvascular complication risk in individuals with type 2 diabetes[21,22], and CVD and mortality risk in individuals with type 1 diabetes[23,24]. Accordingly, HbA$_{1c}$ has become the gold-standard method for assessment of overall

glycemic control and complication risk, as well as a primary outcome measure in the development of new glucose-lowering medication and diabetes management technologies[25].

Despite the validity of its central role in diabetes care, HbA$_{1c}$ has specific shortcomings[26]. Importantly, it fails to provide sufficient information on daily glucose dynamics[27]. Examples of individuals with similar HbA$_{1c}$ values but very different glucose fluctuation profiles illustrate this notion. In **Figure 1.1**, the glucose profile of participant B is evidently characterized by greater variability. Such an example underscores the point that potentially harmful hyperglycemic peaks and hypoglycemic nadirs can go unnoticed when HbA$_{1c}$ is the sole focus of glycemic therapy. The danger of this was made clear by the Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial[28]. ACCORD was discontinued after a 3.5-year follow-up, because of higher mortality rates in the intensive-therapy group, which was presumed to be caused by more frequently occurring hypoglycemia[29]. Hence, a case can be made against attacking the average without paying heed to what transpires above and below the mean. In addition, a post-hoc analysis of the DCCT indicated that even in case of equal HbA$_{1c}$ values the risk of diabetic retinopathy was lower in the intensive treatment group[30]. At the time, the authors proposed that intensive treatment, which entailed multiple daily insulin injections, could have improved risk via its effects on daily glucose fluctuations[30]. This observation, although later contended[31], taken together with the results of several experimental studies, led to the hypothesis that variability in glucose concentrations may be a modifiable risk factor for diabetic complications, irrespective of HbA$_{1c}$[32,33].

## Daily glucose variability: somewhere, beyond the mean

Glucose or glycemic variability (GV) refers to fluctuations in glucose concentrations over a defined period of time[34]. A clear distinction needs to be made between short-term and long-term GV, as they are measured differently and represent different underlying etiologic concepts[34,35]. Short-term or daily are the adjectives used when GV has been continuously measured over a period of days to weeks[34,36]. It reflects the actual variability of daily or day-to-day glucose patterns. By contrast, long-term or visit-to-visit GV refers to the variance in classic glycemic indices (e.g., HbA$_{1c}$, fasting plasma glucose) that have been periodically measured over weeks, months, or years[34,35]. While it may reflect daily glucose fluctuations to some extent, long-term GV is viewed to largely represent difficult to measure factors that affect glycemic control (e.g., therapy adherence, multimorbidity, infections)[35]. Henceforth, this thesis will focus on the measurement and consequences of daily GV.

**A**

**Participant A**



**B**

**Participant B**



**C**

**Glycemic indices**

|  | HbA$_{1c}$ mmol/mol | FPG mmol/L | MSG mmol/L | SD mmol/L | CV % | TIR % | IGP mmol/L |
|---|---|---|---|---|---|---|---|
| Participant A | 44 | 6.7 | 8.1 | 1.4 | 16.7 | 90.1 | 7.6 |
| Participant B | 44 | 6.6 | 8.0 | 2.5 | 31.0 | 76.0 | 12.7 |
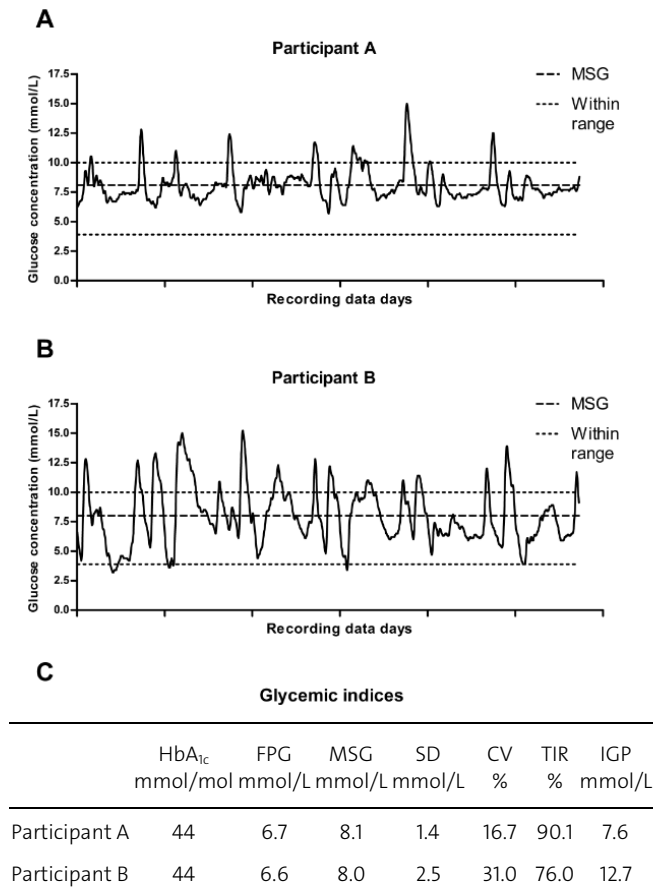
**Figure 1.1**    **Six-day continuous glucose monitoring profiles of two study participants diagnosed with type 2 diabetes who had similar HbA$_{1c}$, fasting plasma glucose and mean sensor glucose concentrations, but substantially different glucose variability profiles.** Higher SD, CV, and IGP values, as well as lower TIR all reflect greater glucose variability. HbA$_{1c}$, FPG, and IGP were measured in venous blood; MSG, SD, CV, and TIR were measured with continuous glucose monitoring. MSG, mean sensor glucose; FPG, fasting plasma glucose; SD, standard deviation; CV, coefficient of variation; TIR, time in range; IGP, oral glucose tolerance test-derived incremental glucose peak.

Importantly, it should be noted that daily glucose fluctuations are part of normal physiology[34,37,38]. Despite highly sophisticated pancreatic regulation, which physiologically maintains blood glucose concentrations within a 4 to 6 mmol/L range[39], it is impossible to accomplish absolutely stable concentrations[27]. The question then becomes at what point GV should be regarded abnormal. Thus far, such GV reference values have been sparsely investigated[40,41], especially in a setting without dietary

restrictions[42]. In this thesis, we aim to establish GV reference values based on a large sample of individuals with NGM (*Chapter 2*). A logical follow-up question would be whether elevated daily GV is also harmful. While several reviews and opinion pieces posit there to be sufficient biological rationale and *in vivo* evidence supporting this[33,34,43,44], other articles argue against it[25,45,46]. The absence of scientific consensus is partly due to the immaturity of the field (i.e., lack of large-scale studies), which stems from the fact that until the start of this century accurate measurement of GV has been challenging[47]. Before I go into the consequences of daily GV, I will first address its measurement.

## The measurement of daily glucose variability

Continuous glucose monitoring (CGM) is the gold-standard method to measure daily GV, as it is best able to assess its many aspects, which include amplitude, frequency, and duration of glucose fluctuations[26,48]. A continuous glucose monitor is a wearable body sensor that is normally worn abdominally or on the back of the upper arm and automatically measures glucose concentrations—most commonly in the interstitial fluid—at 5- or 15-minute intervals[49,50]. In order to calibrate the sensor data, most CGM devices require capillary self-measurement of blood glucose (SMBG) several times a day[49]. A differentiation can be made between real-time or personal CGM, which enables the user to inspect their glucose value at all times, or retrospective or professional CGM, which only allows for glucose profile evaluation after the device has been removed[49].

CGM became commercially available in 2000[51]. Since then, it is increasingly being used in a clinical setting, mostly by individuals with type 1 diabetes, who either use CGM solitarily or incorporated in a closed-loop insulin delivery system[49,52]. Such a system combines CGM, an insulin infusion device, and a control algorithm in order to continuously regulate blood glucose levels[53,54]. While CGM devices are steadily being improved upon with regard to, among others, their accuracy, size, and user-friendliness[51], certain inherent drawbacks persist. These include ~10-minute sensor delay, a consequence of interstitial glucose measurement, and brief periods of malfunction[51,53,55], both of which can be particularly problematic in case of closed-loop dosing systems. In this thesis, we investigate whether machine learning-based glucose prediction can overcome these obstacles (*Chapter 3*).

In addition to its clinical use, CGM has been increasingly utilized in large epidemiological studies. It, however, remains challenging to do so, mainly because of the costliness and relative invasiveness of CGM itself and the often required SMBG to calibrate device[47,50]. From both a scientific (i.e., data reliability) and logistic viewpoint (i.e., device costs, participant strain), it is useful to establish the minimum number of consecutive CGM days required to reliably measure CGM-derived indices. While this has been established for individuals with type 1 diabetes (i.e., 12-15 days)[56-58], it is

currently unclear for individuals with other GMS categories. In this thesis, we assess how many days are minimally needed to reliably measure the main CGM-derived indices in individuals with NGM, prediabetes, or type 2 diabetes (*Chapter 2*).

The detailed glucose profile data gathered with CGM can be used to calculate indices of average glycemia (e.g., mean sensor glucose; MSG) as well as a wide array of GV indices[27,36], such as standard deviation (SD), coefficient of variation (CV; SD/MSG * 100%), interquartile range, mean amplitude of glycemic excursion, and mean of daily differences. Based on expert opinion, SD and CV are presently the recommended indices for reporting GV, as they are easiest to calculate and interpret[26,48]. Nevertheless, researchers frequently propose more intricate techniques for CGM-based quantification of GV[59-61]. The spectral clustering-derived 'glucotype', for example, has recently been promoted as a more comprehensive GV measure than the currently used indices[59]. The glucotype algorithm processes the CGM glucose profiles with use of spectral clustering, and then categorizes individuals based on their time spent within each of the three glucotypes, which are asserted to represent increasing levels of GV (i.e., low, moderate, or severe variability)[59]. At present, the glucotypes' validity and generalizability to other study populations are unclear. In this thesis, we investigate whether the glucotype has added value for GV assessment (*Chapter 4*).

CGM can additionally be used to calculate time in range (TIR), a relatively novel glycemic index that is defined as the percentage of time spent between glucose concentrations of 3.9 and 10.0 mmol/L[26]. TIR is strongly related to $HbA_{1c}$, and is therefore regarded as a measure of overall glycemic control[62]. However, TIR and GV are also inversely linked[63], as GV entails hyperglycemic peaks (i.e., higher time above range) as well as hypoglycemic nadirs (i.e., higher time below range)[64]. As such, TIR is a CGM-derived index that combines aspects of both mean glycemia and GV (**Figure 1.1**).

If CGM is unavailable, GV can be quantified via alternative measures. First, glucose profiles based on frequently performed SMBG can be used to calculate most of the aforementioned indices of GV[27,46]. However, as measurement error is strongly related to the sampling frequency, SMBG should—depending on the GV index used—ideally be performed at least 6-24 times per day[65]. Second, certain studies have used 1,5-anhydroglucitol (1,5-AG), a polyol that competes with glucose for renal reabsorption, as an inverse surrogate marker of GV[66-68]. However, its validity has been criticized, especially due to its limited added value at elevated glucose concentrations[37,46]. Last, while certain oral glucose tolerance test (OGTT)-derived glucose values could be used a proxies of GV[69], it is currently unclear which are most suitable. In this thesis, we assess which OGTT-derived indices most strongly associate with GV and, thus, could be used as valid GV proxies (*Chapter 2*). We hypothesize that the incremental glucose peak (IGP), which is calculated by subtracting fasting plasma glucose from the highest OGTT glucose value minus (**Figure 1.2**), best reflects SD, as both indices represent a glucose excursion relative to the mean.
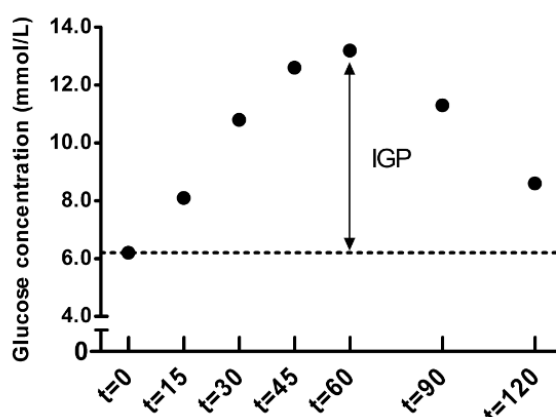
Figure 1.2    The calculation of the incremental glucose peak with use of the seven-point oral glucose tolerance test profile of a study participant with prediabetes. IGP, incremental glucose peak. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points.

## The consequences of daily glucose variability

First, it is important to reiterate the distinction between daily and long-term GV, which are—in the context of GV consequences—occasionally used interchangeably[34,44]. Notwithstanding its demonstrated association with several adverse outcomes (e.g., CVD and mortality), long-term GV is viewed to largely reflect stability of glycemic control rather than daily glucose fluctuations[35]. A wide array of potential consequences of daily GV have thus far been studied, which range from the molecular processes (e.g., oxidative stress[70]), to preclinical or functional measures (e.g., carotid-intima media thickness[71]), and frank complications (e.g., CVD[72]). As stated previously, there is currently no scientific consensus on whether daily GV is harmful independent of its contribution to mean glycemia[25,33,34,43-46]. Below, we provide a concise overview of the current studies on the adverse consequences of daily GV.

The biological pathways through which daily GV could exert harmful effects have been extensively explored[32]. The hyperglycemia-induced generation of oxidative stress is known to be a central process in the pathophysiology of diabetic complications[73]. The concept that GV further aggravates this process is supported by several *in vitro* studies that observed higher levels of oxidative stress after glucose oscillations than after to stable hyperglycemic concentrations[32,70]. Compared to stable hyperglycemia, GV was further found to increase cellular apoptosis, protein kinase C activation, and collagen synthesis[32]. However, these results have not been convincingly translated to *in vivo* studies. While GV was associated with oxidative stress in individuals with type 2 diabetes[74,75], such findings were not replicated in individuals with type 1 diabetes[76,77] or

individuals without diabetes[78]. Further, hypoglycemic nadirs could also contribute to the development of complications[79,80].

Currently, clinical studies have mostly focused on CVD (e.g.,[71,72,81]), microvascular disease (e.g.,[82-84]), and cognitive impairment (e.g.,[61,68,85]) as adverse consequences of daily GV. Although the overall results have been conflicting, those derived from the largest study populations tend to oppose rather than support an independent pathophysiological role in the development of complications[46,81]. However, quite a few of these studies were limited by specific methodological concerns. First, most CGM-based studies had a small sample size (n<300, e.g.,[61,71,85]), which negatively impacts statistical power and, thus, increases the risk of type I and type II errors[86,87]. Second, associations with outcome measures were frequently inadequately adjusted for confounders (e.g.,[61,68]), the most important of which is mean glycemia[47,88], which could have led to overestimated associations. Third, several studies have suboptimally measured GV (e.g., via SMBG[72,82-84] or 1,5-AG[67,68]), which could have led to underestimated associations due to random measurement error[89]. In this thesis, we aim to sufficiently address these points when investigating the associations of daily GV with several arterial measures (*Chapters 5 and 6*), endothelial function (*Chapter 5*), retinal nerve fiber layer thickness (*Chapter 7*), and cognitive performance (*Chapter 8*).

## Studies used in this thesis

To address the aforementioned knowledge gaps, we used data from The Maastricht Study, the PRE-D Trial, and OhioT1DM Dataset.

The Maastricht Study is an ongoing prospective population-based cohort study that is characterized by an extensive phenotyping approach and primarily focuses on investigating the causes, consequences, and comorbidities of type 2 diabetes[90]. Individuals aged between 40 and 75 years and living in the southern part of the Netherlands were eligible for participation. Mass media campaigns and targeted mailings were utilized for participant recruitment, and individuals with type 2 diabetes were oversampled by design. In order to facilitate the study of daily GV, all participants of The Maastricht Study included between 19 September 2016 and 13 September 2018 were invited to undergo professional CGM (iPro2 and Enlite Glucose Sensor; Medtronic) as a part of their standard measurements. In the same period, a group of recently included individuals were asked to return for a separate CGM research visit. This thesis includes cross-sectional data of the first 8,005 participants included in The Maastricht Study. Of these participants, 896 have undergone CGM.

The PRE-D Trial is a Danish, investigator-initiated, randomized, controlled, parallel, open-label, superiority trial, designed to study the effects of four different interventions (i.e., 10mg of dapagliflozin once daily; 850mg of metformin twice daily; 30-minute interval training sessions 5 days per week; or lifestyle advice) in individuals with prediabetes[91]. The primary endpoint of the PRE-D Trial is a reduction in CGM-

measured GV. Secondary endpoints include concomitant changes in glucose metabolism measures, body weight, cardiorespiratory fitness, blood pressure, plasma lipids, physical activity, and dietary intake. This thesis includes baseline data of the 116 participants included, all of whom underwent CGM (iPro2 and Enlite Glucose Sensor; Medtronic).

The OhioT1DM Dataset is publically available for scientific purposes and contains data of 6 individuals with type 1 diabetes who were all using insulin pump therapy in combination with CGM (Enlite Glucose Sensor; Medtronic)[92]. The participants provided interstitial glucose values every five minutes for an eight-week period. The OhioT1DM Dataset is commonly used for the evaluation of prediction model performance (e.g.,[93,94]).

## Outline and aims of this thesis

This thesis has two main aims. The first aim was to assess methodological aspects of CGM and daily GV measurement. The second aim was to investigate the adverse consequences of daily GV. **Figure 1.3** provides an overview of the topics investigated in this thesis.

In *Chapters 2 through 4* (Part I), we assessed methodological aspects of CGM and the measurement of daily GV. In *Chapter 2*, we investigated the minimum number of days required to reliably measure CGM-derived indices, assessed the reference values of these indices, and studied their correlation with established glycemic indices. In *Chapter 3*, we investigated to whether machine learning-based glucose prediction can be used to overcome certain inherent drawbacks of CGM (i.e., ~10-minute sensor delay and periods of malfunction). In *Chapter 4*, we investigated whether the 'glucotype' by Hall et al.[59] has added value for GV assessment and discussed the results in the context of precision medicine.

In *Chapters 5 through 8* (Part II), we investigated potential adverse consequences of daily GV, either measured via CGM or based on an OGTT-derived proxy (i.e., IGP). In *Chapter 5*, we studied the associations of the OGTT-based IGP with measures of arterial stiffness, structure, and function, and measures of endothelial function. In *Chapter 6*, we studied the associations of CGM-measured GV and TIR with measures of arterial stiffness, structure, and function. In *Chapter 7*, we studied the associations of GMS, measures of glycemia, and indices of daily GV with retinal nerve fiber layer thickness, which is an index of neurodegeneration and precursor of neuropathy and retinopathy. In *Chapter 8*, we studied the associations of CGM-measured GV and TIR, as well as the OGTT-based IGP with overall cognitive performance.

In *Chapter 9*, I provide a summary of the main findings of this thesis and discuss them in the context of current literature. I further review the methodological challenges of measuring GV in a large epidemiological setting, and address directions

for future research. Last, in *Chapter 10*, the scientific and societal impact of this thesis is examined.
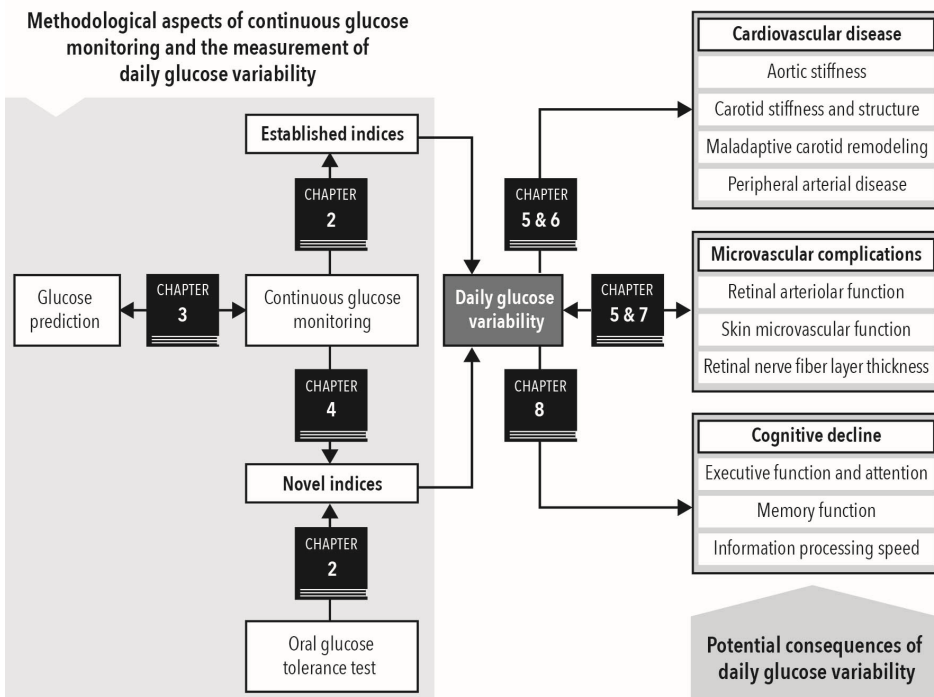


**Figure 1.3**    Schematic overview of the thesis aims.

# References

1.  Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol. 2018;14(2):88-98.
2.  Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. Diabetes Res Clin Pract. 2019;157:107843.
3.  American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S66-S76.
4.  American Diabetes A. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S14-S31.
5.  Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical Update: Cardiovascular Disease in Diabetes Mellitus: Atherosclerotic Cardiovascular Disease and Heart Failure in Type 2 Diabetes Mellitus - Mechanisms, Management, and Clinical Considerations. Circulation. 2016;133(24):2459-502.
6.  Preis SR, Hwang SJ, Coady S, Pencina MJ, D'Agostino RB, Sr., Savage PJ, et al. Trends in all-cause and cardiovascular disease mortality among women and men with and without diabetes mellitus in the Framingham Heart Study, 1950 to 2005. Circulation. 2009;119(13):1728-35.
7.  Taylor KS, Heneghan CJ, Farmer AJ, Fuller AM, Adler AI, Aronson JK, et al. All-cause and cardiovascular mortality in middle-aged people with type 2 diabetes compared with people without diabetes in a large U.K. primary care database. Diabetes Care. 2013;36(8):2366-71.
8.  Bragg F, Holmes MV, Iona A, Guo Y, Du H, Chen Y, et al. Association Between Diabetes and Cause-Specific Mortality in Rural and Urban Areas of China. JAMA. 2017;317(3):280-9.
9.  American Diabetes A. 11. Microvascular Complications and Foot Care: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S135-S51.
10. Stehouwer CDA. Microvascular Dysfunction and Hyperglycemia: A Vicious Cycle With Widespread Consequences. Diabetes. 2018;67(9):1729-41.
11. Dunlay SM, Givertz MM, Aguilar D, Allen LA, Chan M, Desai AS, et al. Type 2 Diabetes Mellitus and Heart Failure: A Scientific Statement From the American Heart Association and the Heart Failure Society of America: This statement does not represent an update of the 2017 ACC/AHA/HFSA heart failure guideline update. Circulation. 2019;140(7):e294-e324.
12. Tsilidis KK, Kasimis JC, Lopez DS, Ntzani EE, Ioannidis JP. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. BMJ. 2015;350:g7607.
13. Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. Nat Rev Gastroenterol Hepatol. 2017;14(1):32-42.
14. van Sloten TT, Sedaghat S, Carnethon MR, Launer LJ, Stehouwer CDA. Cerebral microvascular complications of type 2 diabetes: stroke, cognitive dysfunction, and depression. Lancet Diabetes Endocrinol. 2020;8(4):325-36.
15. Jing X, Chen J, Dong Y, Han D, Zhao H, Wang X, et al. Related factors of quality of life of type 2 diabetes patients: a systematic review and meta-analysis. Health Qual Life Outcomes. 2018;16(1):189.
16. Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Barnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. Lancet Diabetes Endocrinol. 2017;5(6):423-30.
17. Bommer C, Sagalova V, Heesemann E, Manne-Goehler J, Atun R, Barnighausen T, et al. Global Economic Burden of Diabetes in Adults: Projections From 2015 to 2030. Diabetes Care. 2018;41(5):963-70.
18. American Diabetes A. 4. Comprehensive Medical Evaluation and Assessment of Comorbidities: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S37-S47.
19. American Diabetes A. 10. Cardiovascular Disease and Risk Management: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S111-S34.

20. Diabetes C, Complications Trial Research G, Nathan DM, Genuth S, Lachin J, Cleary P, et al. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993;329(14):977-86.

21. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet. 1998;352(9131):837-53.

22. Group AC, Patel A, MacMahon S, Chalmers J, Neal B, Billot L, et al. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. N Engl J Med. 2008;358(24):2560-72.

23. Nathan DM, Cleary PA, Backlund JY, Genuth SM, Lachin JM, Orchard TJ, et al. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. N Engl J Med. 2005;353(25):2643-53.

24. Writing Group for the DERG, Orchard TJ, Nathan DM, Zinman B, Cleary P, Brillon D, et al. Association between 7 years of intensive treatment of type 1 diabetes and long-term mortality. JAMA. 2015;313(1):45-53.

25. Bergenstal RM. Glycemic Variability and Diabetes Complications: Does It Matter? Simply Put, There Are Better Glycemic Markers! Diabetes Care. 2015;38(8):1615-21.

26. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care. 2019;42(8):1593-603.

27. Kovatchev B, Cobelli C. Glucose Variability: Timing, Risk Analysis, and Relationship to Hypoglycemia in Diabetes. Diabetes Care. 2016;39(4):502-10.

28. Action to Control Cardiovascular Risk in Diabetes Study G, Gerstein HC, Miller ME, Byington RP, Goff DC, Jr., Bigger JT, et al. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med. 2008;358(24):2545-59.

29. Dluhy RG, McMahon GT. Intensive glycemic control in the ACCORD and ADVANCE trials. N Engl J Med. 2008;358(24):2630-3.

30. The relationship of glycemic exposure (HbA1c) to the risk of development and progression of retinopathy in the diabetes control and complications trial. Diabetes. 1995;44(8):968-83.

31. Lachin JM, Genuth S, Nathan DM, Zinman B, Rutledge BN, Group DER. Effect of glycemic exposure on the risk of microvascular complications in the diabetes control and complications trial-- revisited. Diabetes. 2008;57(4):995-1001.

32. Hirsch IB, Brownlee M. Should minimal blood glucose variability become the gold standard of glycemic control? J Diabetes Complications. 2005;19(3):178-81.

33. Brownlee M, Hirsch IB. Glycemic variability: a hemoglobin A1c-independent risk factor for diabetic complications. JAMA. 2006;295(14):1707-8.

34. Ceriello A, Monnier L, Owens D. Glycaemic variability in diabetes: clinical and therapeutic implications. Lancet Diabetes Endocrinol. 2019;7(3):221-30.

35. Gorst C, Kwok CS, Aslam S, Buchan I, Kontopantelis E, Myint PK, et al. Long-term Glycemic Variability and Risk of Adverse Outcomes: A Systematic Review and Meta-analysis. Diabetes Care. 2015;38(12):2354-69.

36. Rodbard D. Glucose Variability: A Review of Clinical Applications and Research Developments. Diabetes Technol Ther. 2018;20(S2):S25-S215.

37. Service FJ. Glucose variability. Diabetes. 2013;62(5):1398-404.

38. DeVries JH. Glucose variability: where it is important and how to measure it. Diabetes. 2013;62(5):1405-8.

39. Roder PV, Wu B, Liu Y, Han W. Pancreatic regulation of glucose homeostasis. Exp Mol Med. 2016;48:e219.

40. Zhou J, Li H, Ran X, Yang W, Li Q, Peng Y, et al. Establishment of normal reference ranges for glycemic variability in Chinese subjects using continuous glucose monitoring. Med Sci Monit. 2011;17(1):CR9-13.

41. Zhou J, Li H, Ran X, Yang W, Li Q, Peng Y, et al. Reference values for continuous glucose monitoring in Chinese subjects. Diabetes Care. 2009;32(7):1188-93.

42.  Gude F, Diaz-Vidal P, Rua-Perez C, Alonso-Sampedro M, Fernandez-Merino C, Rey-Garcia J, et al. Glycemic Variability and Its Association With Demographics and Lifestyles in a General Adult Population. J Diabetes Sci Technol. 2017;11(4):780-90.
43.  Hirsch IB. Glycemic Variability and Diabetes Complications: Does It Matter? Of Course It Does! Diabetes Care. 2015;38(8):1610-4.
44.  Ceriello A. Glucose Variability and Diabetic Complications: Is It Time to Treat? Diabetes Care. 2020;43(6):1169-71.
45.  Cavalot F. Do data in the literature indicate that glycaemic variability is a clinical problem? Glycaemic variability and vascular complications of diabetes. Diabetes Obes Metab. 2013;15 Suppl 2:3-8.
46.  Siegelaar SE, Holleman F, Hoekstra JB, DeVries JH. Glucose variability; does it matter? Endocr Rev. 2010;31(2):171-82.
47.  Rodbard D. The challenges of measuring glycemic variability. J Diabetes Sci Technol. 2012;6(3): 712-5.
48.  Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care. 2017;40(12):1631-40.
49.  Klonoff DC, Ahn D, Drincic A. Continuous glucose monitoring: A review of the technology and clinical use. Diabetes Res Clin Pract. 2017;133:178-92.
50.  Wood A, O'Neal D, Furler J, Ekinci EI. Continuous glucose monitoring: a review of the evidence, opportunities for future use and ongoing challenges. Intern Med J. 2018;48(5):499-508.
51.  Rodbard D. Continuous Glucose Monitoring: A Review of Successes, Challenges, and Opportunities. Diabetes Technol Ther. 2016;18 Suppl 2:S3-S13.
52.  van den Boom L, Karges B, Auzanneau M, Rami-Merhar B, Lilienthal E, von Sengbusch S, et al. Temporal Trends and Contemporary Use of Insulin Pump Therapy and Glucose Monitoring Among Children, Adolescents, and Adults With Type 1 Diabetes Between 1995 and 2017. Diabetes Care. 2019;42(11):2050-6.
53.  Cobelli C, Renard E, Kovatchev B. Artificial pancreas: past, present, future. Diabetes. 2011;60(11):2672-82.
54.  Bruttomesso D. Toward Automated Insulin Delivery. N Engl J Med. 2019;381(18):1774-5.
55.  Blauw H, Keith-Hynes P, Koops R, DeVries JH. A Review of Safety and Design Requirements of the Artificial Pancreas. Ann Biomed Eng. 2016;44(11):3158-72.
56.  Xing D, Kollman C, Beck RW, Tamborlane WV, Laffel L, Buckingham BA, et al. Optimal sampling intervals to assess long-term glycemic control using continuous glucose monitoring. Diabetes Technol Ther. 2011;13(3):351-8.
57.  Neylon OM, Baghurst PA, Cameron FJ. The Minimum Duration of Sensor Data From Which Glycemic Variability Can Be Consistently Assessed. J Diabetes Sci Technol. 2014;8(2):273-6.
58.  Riddlesworth TD, Beck RW, Gal RL, Connor CG, Bergenstal RM, Lee S, et al. Optimal Sampling Duration for Continuous Glucose Monitoring to Determine Long-Term Glycemic Control. Diabetes Technol Ther. 2018;20(4):314-6.
59.  Hall H, Perelman D, Breschi A, Limcaoco P, Kellogg R, McLaughlin T, et al. Glucotypes reveal new patterns of glucose dysregulation. PLoS Biol. 2018;16(7):e2005143.
60.  Peyser TA, Balo AK, Buckingham BA, Hirsch IB, Garcia A. Glycemic Variability Percentage: A Novel Method for Assessing Glycemic Variability from Continuous Glucose Monitor Data. Diabetes Technol Ther. 2018;20(1):6-16.
61.  Cui X, Abduljalil A, Manor BD, Peng C-K, Novak V. Multi-scale glycemic variability: a link to gray matter atrophy and cognitive decline in type 2 diabetes. PloS one. 2014;9(1).
62.  Beck RW, Bergenstal RM, Cheng P, Kollman C, Carlson AL, Johnson ML, et al. The Relationships Between Time in Range, Hyperglycemia Metrics, and HbA1c. J Diabetes Sci Technol. 2019;13(4): 614-26.
63.  Rodbard D. Glucose Time In Range, Time Above Range, and Time Below Range Depend on Mean or Median Glucose or HbA1c, Glucose Coefficient of Variation, and Shape of the Glucose Distribution. Diabetes Technol Ther. 2020.

64. Rodbard D. Glycemic variability: measurement and utility in clinical medicine and research--one viewpoint. Diabetes Technol Ther. 2011;13(11):1077-80.

65. Baghurst PA, Rodbard D, Cameron FJ. The minimum frequency of glucose measurements from which glycemic variation can be consistently assessed. J Diabetes Sci Technol. 2010;4(6):1382-5.

66. Dungan KM. 1,5-anhydroglucitol (GlycoMark) as a marker of short-term glycemic control and glycemic excursions. Expert Rev Mol Diagn. 2008;8(1):9-19.

67. Selvin E, Rawlings AM, Grams M, Klein R, Steffes M, Coresh J. Association of 1,5-anhydroglucitol with diabetes and microvascular conditions. Clin Chem. 2014;60(11):1409-18.

68. Rawlings AM, Sharrett AR, Mosley TH, Ballew SH, Deal JA, Selvin E. Glucose Peaks and the Risk of Dementia and 20-Year Cognitive Decline. Diabetes Care. 2017;40(7):879-86.

69. Meier JJ, Baller B, Menge BA, Gallwitz B, Schmidt WE, Nauck MA. Excess glycaemic excursions after an oral glucose tolerance test compared with a mixed meal challenge and self-measured home glucose profiles: is the OGTT a valid predictor of postprandial hyperglycaemia and vice versa? Diabetes Obes Metab. 2009;11(3):213-22.

70. Piconi L, Quagliaro L, Assaloni R, Da Ros R, Maier A, Zuodar G, et al. Constant and intermittent high glucose enhances endothelial cell apoptosis through mitochondrial superoxide overproduction. Diabetes Metab Res Rev. 2006;22(3):198-203.

71. Magri CJ, Mintoff D, Camilleri L, Xuereb RG, Galea J, Fava S. Relationship of Hyperglycaemia, Hypoglycaemia, and Glucose Variability to Atherosclerotic Disease in Type 2 Diabetes. J Diabetes Res. 2018;2018:7464320.

72. Siegelaar SE, Kerr L, Jacober SJ, Devries JH. A decrease in glucose variability does not reduce cardiovascular event rates in type 2 diabetic patients after acute myocardial infarction: a reanalysis of the HEART2D study. Diabetes Care. 2011;34(4):855-7.

73. Brownlee M. The pathobiology of diabetic complications: a unifying mechanism. Diabetes. 2005;54(6):1615-25.

74. Monnier L, Mas E, Ginet C, Michel F, Villon L, Cristol JP, et al. Activation of oxidative stress by acute glucose fluctuations compared with sustained chronic hyperglycemia in patients with type 2 diabetes. JAMA. 2006;295(14):1681-7.

75. Ceriello A, Esposito K, Piconi L, Ihnat MA, Thorpe JE, Testa R, et al. Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in normal and type 2 diabetic patients. Diabetes. 2008;57(5):1349-54.

76. Wentholt IM, Kulik W, Michels RP, Hoekstra JB, DeVries JH. Glucose fluctuations and activation of oxidative stress in patients with type 1 diabetes. Diabetologia. 2008;51(1):183-90.

77. Monnier L, Colette C, Mas E, Michel F, Cristol JP, Boegner C, et al. Regulation of oxidative stress by glycaemic control: evidence for an independent inhibitory effect of insulin therapy. Diabetologia. 2010;53(3):562-71.

78. Wakil A, Smith KA, Atkin SL, Kilpatrick ES. Short-term glucose variability in healthy volunteers is not associated with raised oxidative stress markers. Diabetes Obes Metab. 2012;14(11):1047-9.

79. International Hypoglycaemia Study G. Hypoglycaemia, cardiovascular disease, and mortality in diabetes: epidemiology, pathogenesis, and management. Lancet Diabetes Endocrinol. 2019;7(5):385-96.

80. Gimenez M, Gilabert R, Monteagudo J, Alonso A, Casamitjana R, Pare C, et al. Repeated episodes of hypoglycemia as a potential aggravating factor for preclinical atherosclerosis in subjects with type 1 diabetes. Diabetes Care. 2011;34(1):198-203.

81. Lu J, Ma X, Shen Y, Wu Q, Wang R, Zhang L, et al. Time in Range Is Associated with Carotid Intima-Media Thickness in Type 2 Diabetes. Diabetes Technol Ther. 2020;22(2):72-8.

82. Siegelaar SE, Kilpatrick ES, Rigby AS, Atkin SL, Hoekstra JB, Devries JH. Glucose variability does not contribute to the development of peripheral and autonomic neuropathy in type 1 diabetes: data from the DCCT. Diabetologia. 2009;52(10):2229-32.

83. Kilpatrick ES, Rigby AS, Atkin SL. Effect of glucose variability on the long-term risk of microvascular complications in type 1 diabetes. Diabetes Care. 2009;32(10):1901-3.

84. Kilpatrick ES, Rigby AS, Atkin SL. The effect of glucose variability on the risk of microvascular complications in type 1 diabetes. Diabetes Care. 2006;29(7):1486-90.

85. Rizzo MR, Marfella R, Barbieri M, Boccardi V, Vestini F, Lettieri B, et al. Relationships between daily acute glucose fluctuations and cognitive performance among aged type 2 diabetic patients. Diabetes care. 2010;33(10):2169-74.
86. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14(5):365-76.
87. Ioannidis JP. Why most published research findings are false. PLoS Med. 2005;2(8):e124.
88. Rodbard D. Glycemic variability: challenges in interpretation. Diabetes Technol Ther. 2015;17(6):370-2.
89. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. BMJ. 2010;340:c2289.
90. Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.
91. Faerch K, Amadid H, Nielsen LB, Ried-Larsen M, Karstoft K, Persson F, et al. Protocol for a randomised controlled trial of the effect of dapagliflozin, metformin and exercise on glycaemic variability, body composition and cardiovascular risk in prediabetes (the PRE-D Trial). BMJ Open. 2017;7(5):e013802.
92. Marling C, Bunescu RC, editors. The OhioT1DM Dataset For Blood Glucose Level Prediction. KHD@IJCAI; 2018.
93. Martinsson J, Schliep A, Eliasson B, Mogren O. Blood Glucose Prediction with Variance Estimation Using Recurrent Neural Networks. Journal of Healthcare Informatics Research. 2020;4(1):1-18.
94. Kriventsov S, Lindsey A, Hayeri A. The Diabits App for Smartphone-Assisted Predictive Monitoring of Glycemia in Patients With Diabetes: Retrospective Observational Study. JMIR Diabetes. 2020;5(3):e18660.

# Part I

Methodological aspects of continuous glucose
monitoring and the measurement of
daily glucose variability

# Chapter 2

Glucose variability assessed with continuous glucose monitoring: reliability, reference values, and correlations with established glycemic indices – The Maastricht Study

Yuri D. Foreman, Martijn C.G.J. Brouwers, Carla J.H. van der Kallen, Demi M.E. Pagen, Marleen M.J. van Greevenbroek, Ronald M.A. Henry, Annemarie Koster, Anke Wesselius, Nicolaas C. Schaper, and Coen D.A. Stehouwer

# Abstract

**Background**

Glucose variability (GV) measured by continuous glucose monitoring (CGM) has become an accepted marker of glycemic control. Nevertheless, several methodological aspects of GV assessment require further study. We therefore investigated the minimum number of days needed to reliably measure GV, assessed GV reference values, and studied the correlation of GV with established glycemic indices (i.e., HbA$_{1c}$, seven-point OGTT-derived indices).

**Methods**

We used cross-sectional data from The Maastricht Study, an observational population-based cohort enriched with type 2 diabetes. Participants with more than 48 hours of CGM (iPro2; Medtronic) were included for analysis (n=851; age: 60±9years; 49% women; 23% type 2 diabetes). We used mean sensor glucose (MSG), standard deviation (SD), and coefficient of variation (CV) as CGM-derived indices (the latter two for GV quantification). We calculated reliability using the Spearman–Brown prophecy formula, established reference values by calculating 2.5th-97.5th percentiles, and studied correlations using Spearman's rho.

**Results**

Measurement of MSG, SD, and CV was sufficient reliability (R>0.80) with three monitoring days. The reference ranges, assessed in individuals with normal glucose metabolism (n=470), were 5.03 - 6.69 mmol/L (MSG), 0.44 - 1.37 mmol/L (SD), and 7.74-22.45% (CV). For MSG, the strongest correlation was found with fasting plasma glucose (rho=0.65 [0.61;0.69]); for SD, with the 1-hour OGTT value (rho=0.61 [0.56;0.65]); and for CV, with both the incremental glucose peak (IGP) during the OGTT (rho=0.50 [0.45;0.55]) and the 1-hour OGTT value (rho=0.50 [0.45;0.55]).

**Conclusions**

The reliability findings and reference values are relevant for studies that aim to investigate CGM-measured GV. One-hour OGTT and IGP values can be used as GV indices when CGM is unavailable.

## Introduction

Daily glucose variability (GV) is thought to contribute to the development of diabetes-related complications irrespective of mean glycemia[1], and, hence, is accepted as a valuable marker of glycemic control[2]. Continuous glucose monitoring (CGM) is the preferred method for measuring GV, as it is able to assess many of its aspects (i.e., amplitude, frequency, and duration of fluctuations)[2], whereas other measures (e.g., oral glucose tolerance test [OGTT], repeated self-measurement of blood glucose [SMBG]) only assess singular aspects. CGM is, however, a challenging technology to use in a large epidemiological setting, in part due to its costliness and relative invasiveness[3].

Despite prior efforts, several methodological aspects of GV measurement with CGM need further investigation. First, the required number of days to reliably measure daily GV is unclear for individuals with normal glucose metabolism (NGM), prediabetes or type 2 diabetes. Second, there is currently little information on GV reference values, which are ideally determined in healthy individuals, i.e., individuals with NGM[4]. Third, little is known about the relationship of GV with established glycemic indices, such as fasting plasma glucose (FPG), OGTT-derived indices and $HbA_{1c}$. Certain OGTT-derived indices may give a good approximation of GV and could thus be used as alternatives in an epidemiological setting when CGM is not available.

We therefore 1) investigated the minimum number of days needed to reliably measure GV with CGM; 2) assessed GV reference values according to glucose metabolism status (GMS); and 3) studied the correlation of CGM-measured GV with established glycemic indices, in a large population-based cohort.

## Research design and methods

### *Study population and design*

We used data from The Maastricht Study, an observational, prospective, population-based cohort study. The rationale and methodology have been described previously[5]. In brief, The Maastricht Study focuses on the etiology, pathophysiology, complications and comorbidities of type 2 diabetes, and is characterized by an extensive phenotyping approach. All individuals aged between 40 and 75 years and living in the southern part of the Netherlands were eligible for participation. Participants were recruited through mass media campaigns and from the municipal registries and the regional Diabetes Patient Registry via mailings. For reasons of efficiency, recruitment was stratified according to known type 2 diabetes status, with an oversampling of individuals with type 2 diabetes. The Maastricht Study has been approved by the institutional medical ethical committee (NL31329.068.10) and the Minister of Health, Welfare and Sports of the Netherlands (Permit 131088-105234-PG). All participants gave written informed consent.

## Continuous glucose monitoring

From 19 September 2016 until 13 September 2018, all participants were invited to undergo CGM (iPro2 and Enlite Glucose Sensor; Medtronic, Tolochenaz, Switzerland) as a part of their standard measurements (see below). To ensure inclusion of a sufficient number of participants with prediabetes and type 2 diabetes within a reasonable timeframe, we re-invited a prediabetes- and type 2 diabetes-enriched group of participants—who had visited The Maastricht Study between January 2015 until September 2016 for standard measurements—to undergo CGM as a separate research visit (further referred to as 'catch-up visit'). The CGM device was worn on the lower abdomen and recorded subcutaneous interstitial glucose values (range: 2.2 − 22.2 mmol/L) every five minutes for a seven-day period. During this time, participants were asked to perform SMBG with the use of Contour Next (Ascensia Diabetes Care, Mijdrecht, the Netherlands) four times daily (i.e., before main meals and at bedtime). These SMBG values were used to retrospectively calibrate the CGM values at the moment of data upload to CareLink iPro (Medtronic). Accuracies (expressed as the mean absolute relative difference [MARD]) of the iPro2 and Contour Next are 11.0%[6] and 3.1 - 5.6%[7-9], respectively. Participants were blinded to the CGM recording, but not to SMBG values. Diabetes medication use was allowed and no dietary or physical activity instructions were given.

We used the Glycemic Variability Research Tool (GlyVaRT; Medtronic) software to process CGM data. First, the first 24 hours of glucose values were removed, because of insufficient calibration. Next, we excluded individuals of whom less than 24 hours of recording (less than one data day) remained. Then, we calculated mean sensor glucose (MSG), standard deviation (SD), and coefficient of variation (CV; SD/MSG * 100%) for each period of 24 hours (data day) and for the total recording period. Based on international consensus, we used SD and CV as indices of GV[2]. We additionally included MSG in our analyses, as GV indices are preferably interpreted in the context of mean glucose values[3]. As described in the Supplementary methods, we also calculated other commonly used GV indices with the use of GlyVaRT.

## Oral glucose tolerance test

Participants underwent a standardized 2-hour 75 gram OGTT after fasting overnight to determine GMS, which was defined according to the World Health Organization 2006 criteria as: a) NGM, b) impaired fasting glucose and impaired glucose tolerance (combined as prediabetes), or c) type 2 diabetes[10]. For safety reasons, individuals using insulin or with a fasting glucose level above 11.0 mmol/L, as determined by a finger prick, did not undergo an OGTT. For these participants, we determined GMS based on fasting plasma glucose (FPG) and information about their diabetes medication. During the OGTT, we took venous blood glucose samples at baseline (=FPG) and at 15, 30, 45, 60, 90 and 120 minutes. In individuals with complete OGTT data, we calculated the

incremental glucose peak (IGP) by subtracting FPG from the highest OGTT glucose value (= maximum glucose peak).

## Measurement of covariates

As described previously[5], we assessed diabetes duration by questionnaire; assessed use of glucose-lowering medication as part of a medication interview; measured weight, height, and body mass index (BMI) during a physical examination; and measured $HbA_{1c}$ in fasting venous blood. Body weight and glucose-lowering medication use were re-assessed in individuals who attended a CGM catch-up visit.

## Statistical analysis

To determine the minimum number of CGM days required for a reliable assessment of MSG, SD, and CV, we calculated GMS-stratified reliability (R) with the use of the Spearman–Brown prophecy formula: $R = \sigma^2_B / (\sigma^2_B + [\sigma^2_W / n])$, where $\sigma^2_B$ is the inter-individual variance, $\sigma^2_W$ is the intra-individual variance, and $n$ is the number of consecutive monitoring days[11, 12]. We used variance component analysis via ANOVA with the different CGM-derived indices as the dependent variable to estimate the participant variance component (inter-individual; $\sigma^2_B$) and the residual variance component (intra-individual; $\sigma^2_W$). For these analyses we included only individuals with more than five complete data days of CGM to ensure optimal reliability. A R of >0.80 was considered reliable[11].

We calculated the distribution of MSG, SD, and CV values (10th - 90th, 5th - 95th, and 2.5th - 97.5th percentile) in GMS-stratified groups. We established reference ranges based on the 2.5th - 97.5th percentile of individuals with NGM[4]. As an additional analysis, we calculated the distributions for sex- and age-stratified GMS groups. In addition, MSG, SD, and CV differences between GMS groups were statistically compared with the use of ANOVA with post-hoc Bonferroni correction.

A Spearman's rank test was used to determine the correlation (rho) and 95% confidence interval of CGM-derived indices with FPG, OGTT-derived indices (i.e., 15-, 30-, 45-, 60-, 90-, and 120-minute glucose values, maximum glucose peak, and IGP), and $HbA_{1c}$. As an additional analysis, we used linear regression to calculate unadjusted regression coefficients and their corresponding 95% confidence intervals for the glycemic indices that correlated most strongly with MSG, SD and CV.

Additionally, several sensitivity analyses were performed: 1) after exclusion of individuals with less than the reliable number of CGM days; 2) after exclusion of individuals who underwent CGM as part of a catch-up visit; and 3) after exclusion of the individuals with data gaps during CGM recording. Further, the main analyses were also carried out for other commonly used GV indices (i.e., mean amplitude of glucose excursions [MAGE], interquartile range [IQR], and mean of daily differences [MODD].

Normally distributed data are presented as mean and SD, non-normally distributed data as median and IQR, and categorical data as n (%). We considered a p-value of <0.05 statistically significant. All statistical analyses were performed with the Statistical Package for Social Sciences (version 25.0; IBM, Chicago, Illinois, USA).

# Results

## Study population characteristics

The total CGM study population comprised 896 individuals, from which we excluded two participants with diabetes types other than type 2 diabetes. Next, due to the removal of the first 24h of CGM data, we excluded 27 individuals with less than 24 hours of CGM, as for them CGM-derived indices could not be calculated. Then, we excluded individuals (n=16) who had less than one data day (24 hours) of CGM remaining, which resulted in the final study population that consisted of 851 individuals (**Supplemental Figure S2.1**).

**Table 2.1** shows the GMS-stratified participant characteristics of the final study population, based on the regular study visit measurements. With deteriorating GMS, participants were older, more often male, had a higher BMI, and had higher $HbA_{1c}$, FPG, and OGTT-derived index values. Individuals with type 2 diabetes were often newly diagnosed by OGTT. More than half of the participants with type 2 diabetes used metformin as their primary diabetes medication (n=104, 53.1%); insulin use was uncommon (n=19, 9.6%).

Most participants (n=772, 90.7%) completed the seven-day recording period, resulting in six data days of usable CGM data (**Supplemental Table S2.1**). A large number of individuals (n=721, 84.7%) also had five or more consecutive, complete days of CGM. Several participants (n=66, 7.8%) experienced at least one period during which no glucose values were recorded (i.e., a data gap), the length of which varied considerably (median: 5.3 hours [interquartile range: 0.9-19.6]). SMBG adherence was high (i.e., median average daily calibration frequency: 4.0 [3.7 − 4.1], **Supplemental Table S2.1**).

Of the final study population (n = 851), 603 participants (70.9%) underwent CGM as part of their standard measurements (i.e., regular visits) and 248 participants (29.1%) during a catch-up visit. **Supplemental Table S2.2** shows the participant characteristics stratified for research visit. Participant characteristics were generally comparable after stratification for GMS. The median time between the regular and catch-up visit was 2.1 [2.0 − 2.2] years (**Supplemental Table S2.3**). Only individuals with prediabetes had gained weight between their regular and catch-up visit (from 82.9 ± 15.6 to 84.1 ± 16.2 kg, p=0.011); large differences in body weight (>10%) were rare (n=12, 4.8%). Use of glucose-lowering medication was more frequent in individuals with type 2 diabetes at the catch-up visit (71.8% versus 63.7% at the regular and catch-up visit, respectively, p=0.035).

Table 2.1    Participant characteristics of the total study population and of glucose metabolism status-stratified groups.

| Characteristic | Total (N=851) | NGM (n=470) | PreD (n=184) | T2D (n=197) |
|---|---|---|---|---|
| Age, years | 59.9 ± 8.7 | 58.2 ± 8.8 | 61.5 ± 8.1 | 62.4 ± 7.8 |
| Women, n (%) | 418 (49.1) | 266 (56.6) | 83 (45.1) | 69 (35.0) |
| BMI, kg/m$^2$ | 27.2 ± 4.4 | 25.6 ± 3.6 | 28.5 ± 4.4 | 29.7 ± 4.7 |
| Newly diagnosed T2D, n (%) | 70 (8.2) | - | - | 70 (35.5) |
| HbA$_{1c}$, % | 5.7 ± 0.8 | 5.4 ± 0.3 | 5.6 ± 0.4 | 6.7 ± 1.0 |
| HbA$_{1c}$, mmol/mol | 39.1 ± 8.3 | 35.4 ± 3.4 | 37.8 ± 4.2 | 49.2 ± 10.8 |
| Fasting plasma glucose, mmol/L | 5.4 [5.0 − 6.2] | 5.1 [4.8 − 5.4] | 6.0 [5.4 − 6.3] | 7.3 [6.5 − 8.4] |
| OGTT-derived indices | | | | |
| 15-minute, mmol/L | 7.2 [6.5 − 8.3] | 6.7 [6.2 − 7.3] | 7.7 [7.0 − 8.5] | 9.4 [8.4 − 10.3] |
| 30-minute, mmol/L | 8.8 [7.5 − 10.5] | 7.8 [7.0 − 8.8] | 9.8 [8.7 − 10.6] | 11.9 [10.7 − 13.2] |
| 45-minute, mmol/L | 9.2 [7.5 − 11.8] | 7.9 [6.7 − 9.0] | 10.7 [9.7 − 12.2] | 14.0 [12.3 − 15.7] |
| 1-hour, mmol/L | 8.9 [6.7 − 12.0] | 7.1 [5.7 − 8.6] | 11.0 [9.7 − 12.5] | 15.0 [12.9 − 17.0] |
| 90-minute, mmol/L | 7.7 [5.8 − 11.2] | 6.2 [5.2 − 7.3] | 9.9 [8.7 − 11.3] | 15.6 [13.2 − 17.9] |
| 2-hour, mmol/L | 6.7 [5.2 − 9.1] | 5.5 [4.7 − 6.4] | 8.4 [7.5 − 9.2] | 13.6 [11.7 − 16.2] |
| MGP, mmol/L | 9.8 [8.1 − 12.6] | 8.3 [7.4 − 9.3] | 11.4 [10.3 − 12.6] | 16.0 [13.4 − 18.1] |
| IGP, mmol/L | 4.3 [2.9 − 6.7] | 3.2 [2.2 − 4.2] | 5.7 [4.3 − 6.8] | 9.0 [7.2 − 10.3] |
| Diabetes medication use, n (%) | 109 (12.8) | - | - | 109 (55.6) |
| Insulin | 19 (2.2) | - | - | 19 (9.6) |
| Metformin | 104 (12.2) | - | - | 104 (53.1) |
| Sulfonylureas | 21 (2.5) | - | - | 21 (10.7) |
| Thiazolidinediones | 0 (0) | - | - | 0 (0) |
| GLP-1 analogs | 3 (0.4) | - | - | 3 (1.5) |
| DDP-4 inhibitors | 1 (0.1) | - | - | 1 (0.5) |

Data are reported as mean ± SD, median [interquartile range], or number (percentage [%]) as appropriate. Data represent the study population of participants with at least 48 hours of continuous glucose monitoring data. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; BMI, body mass index; OGTT, oral glucose tolerance test; MGP, maximum glucose peak; IGP, incremental glucose peak; HbA1c, glycated hemoglobin A1c; MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4.

## *Reliability of continuous glucose monitoring*

In the reliability analyses, individuals with more than five consecutive, complete days of CGM recording were included (n=721, 84.7%) (**Supplemental Table S2.1** and **S2.4**). **Figure 2.1** illustrates the reliability for different CGM-derived indices per number of consecutive monitoring days in individuals with NGM, prediabetes, or type 2 diabetes. In case of MSG (**Figure 2.1**; panel A), the reliability was sufficient (R > 0.80) after one day of recording for individuals with prediabetes and type 2 diabetes, and after three days for individuals with NGM. In case of SD (**Figure 2.1**; panel B), the reliability was sufficient (R>0.80) after two days of recording for individuals with prediabetes and type 2 diabetes, and after three days for individuals with NGM. In case of CV (**Figure 2.1**; panel C), the reliability was sufficient after three days of recording for individuals with NGM, prediabetes, and type 2 diabetes.
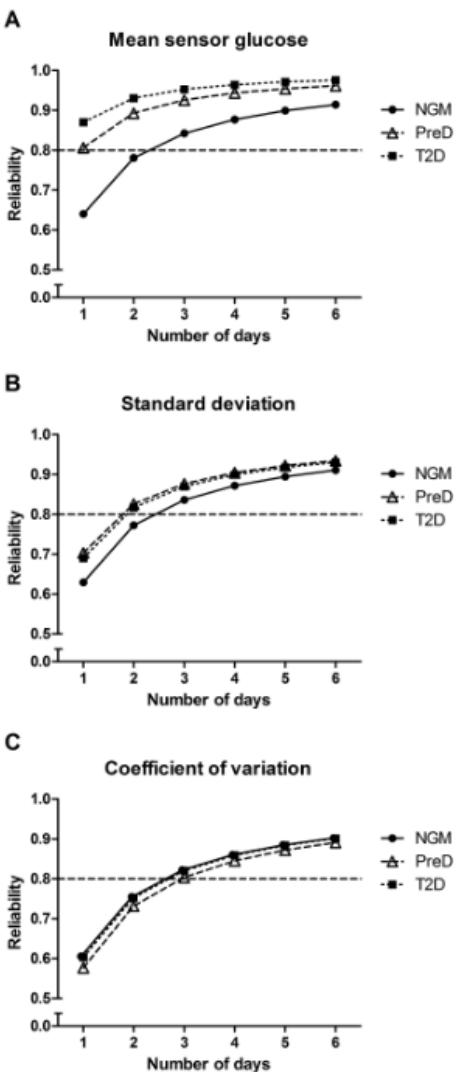
Figure 2.1    Reliability for mean sensor glucose (A), standard deviation (B), and coefficient of variation (C) per number of consecutive continuous glucose monitoring days in individuals with normal glucose metabolism, prediabetes, or type 2 diabetes. The dashed horizontal line represents the cut-off for sufficient reliability (R>0.80). NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes.

## Glucose variability according to glucose metabolism status

**Table 2.2** shows the GMS-stratified median, mean and various percentile values for MSG, SD and CV. Deteriorating GMS was associated with statistically significantly

higher MSG, SD and CV values (even after Bonferroni correction). The distributions (10th - 90th, 5th - 95th, 2.5th - 97.5th percentiles) of MSG, SD and CV were broader as GMS deteriorated. The reference values for MSG, SD and CV, based on the 2.5th - 97.5th percentiles of individuals with NGM, were 5.03 - 6.69 mmol/L, 0.44 - 1.37 mmol/L, and 7.74 - 22.45%, respectively. Distributions of MSG, SD and CV were not materially different for age- and sex-stratified GMS groups (**Supplemental Table S2.5**).

**Table 2.2**  Distribution of mean sensor glucose, standard deviation and coefficient of variation in glucose metabolism status-stratified groups.

| CGM-derived index | | NGM (n=470) | PreD (n=184) | T2D (n=197) | p-value |
|---|---|---|---|---|---|
| MSG, mmol/L | Median [IQR] | 5.8 [5.5 − 6.1] | 6.2 [5.8 − 6.6]* | 7.5 [6.8 − 8.7]*† | <0.001 |
| | Mean ± SD | 5.8 ± 0.4 | 6.3 ± 0.8 | 7.9 ± 1.7 | |
| | 10th - 90th percentile | 5.31 - 6.37 | 5.56 - 7.07 | 6.08 - 9.83 | |
| | 5th - 95th percentile | 5.15 - 6.54 | 5.29 - 7.54 | 5.78 - 11.70 | |
| | 2.5th - 97.5th percentile | 5.03 - 6.69 | 5.04 - 8.27 | 5.63 - 12.78 | |
| SD, mmol/L | Median [IQR] | 0.73 [0.62 − 0.87] | 0.90 [0.74 − 1.13]* | 1.51 [1.14 − 1.95]*† | <0.001 |
| | Mean ± SD | 0.77 ± 0.23 | 0.98 ± 0.35 | 1.64 ± 0.70 | |
| | 10th - 90th percentile | 0.53 - 1.09 | 0.65 - 1.31 | 0.88 - 2.47 | |
| | 5th - 95th percentile | 0.48 - 1.28 | 0.60 - 1.58 | 0.79 - 3.06 | |
| | 2.5th - 97.5th percentile | 0.44 - 1.37 | 0.55 - 1.90 | 0.66 - 3.44 | |
| CV, % | Median [IQR] | 12.6 [10.8 − 14.9] | 14.9 [12.2 − 17.5]* | 19.3 [15.9 − 24.0]*† | <0.001 |
| | Mean ± SD | 13.2 ± 3.6 | 15.3 ± 4.3 | 20.5 ± 6.6 | |
| | 10th - 90th percentile | 9.38 - 18.03 | 10.74 - 20.63 | 13.31 - 29.42 | |
| | 5th - 95th percentile | 8.49 - 20.66 | 10.15 - 24.04 | 12.32 - 33.35 | |
| | 2.5th - 97.5th percentile | 7.74 - 22.45 | 9.29 - 25.78 | 10.51 - 36.50 | |

CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; IQR, interquartile range. * p<0.001 versus NGM; † p<0.001 versus PreD.

### *Correlations with established glycemic indices*

**Figure 2.2** and **Supplemental Table S2.6** depict the Spearman correlation coefficients of MSG, SD and CV with FPG, OGTT-derived indices and HbA$_{1c}$. Since a complete seven-point OGTT was not available in all individuals, the number of individuals included in these correlation analyses varied from 744 to 851 (**Supplemental Table S2.6**). Correlations with MSG and SD were generally stronger than those with CV. For MSG, the strongest correlation was found with FPG (rho=0.65 [0.61; 0.69]), followed by HbA$_{1c}$ (rho=0.63 [0.58; 0.67]) and maximum glucose peak (rho=0.63 [0.59; 0.67]). For both SD and CV, a similar pattern was observed regarding the OGTT time points: the strongest correlations were found with the 1-hour OGTT value and correlations were gradually weaker for the time points that were more distant from this OGTT value. For SD, we found the highest correlation coefficient with the 1-hour OGTT value (rho=0.61 [0.56; 0.65]); for CV, with both IGP (rho=0.50 [0.45; 0.55]) and the 1-hour OGTT value (rho=0.50 [0.45; 0.55]). Of note, the 1-hour OGTT value did not necessarily represent the

glucose peak time point during the OGTT. In fact, with deteriorating GMS, the peak time point appeared to shift to a later moment during the OGTT (**Supplemental Table S2.7**).

**Supplemental Table S2.8** shows GMS-stratified Spearman correlation coefficients. We found that MSG correlated best with FPG in individuals with NGM (rho=0.37 [0.29; 0.44]), and with $HbA_{1c}$ in individuals with prediabetes (rho=0.46 [0.34; 0.57]) and type 2 diabetes (rho=0.60 [0.50; 0.68]). SD correlated best with IGP (rho=0.31 [0.22; 0.39]) and maximum glucose peak (rho=0.30 [0.22; 0.39]) in individuals with NGM, with the 90-minute OGTT value (rho=0.30 [0.16; 0.43]) and $HbA_{1c}$ (rho=0.30 [0.17; 0.43]) in individuals with prediabetes, and with the 90-minute OGTT (rho=0.54 [0.42; 0.65]) value, IGP (rho=0.54 [0.42; 0.65]) and maximum glucose peak (rho=0.54 [0.41; 0.64]) in individuals with type 2 diabetes. We observed that CV correlated best with IGP in individuals with NGM (rho=0.25 [0.16; 0.34]), with the 90-minute (rho=0.23 [0.09; 0.37]) and 2-hour (rho=0.24 [0.09; 0.37]) OGTT value in individuals with prediabetes, and with IGP (rho=0.49 [0.35; 0.60]) in individuals with type 2 diabetes.

**Supplemental Table S2.9** displays the regression coefficients of the five glycemic indices that correlated most strongly with MSG, SD and CV in the total study population. In addition, it shows GMS-stratified regression coefficients of the three glycemic indices that correlated best with MSG, SD and CV in individuals with NGM, prediabetes and type 2 diabetes.
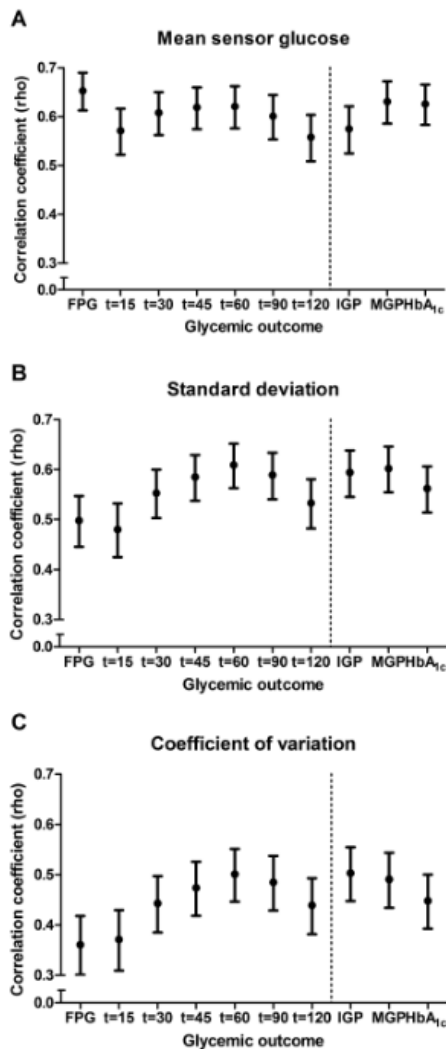
**Figure 2.2**  Correlations of mean sensor glucose (A), standard deviation (B) and coefficient of variation (C) with fasting plasma glucose, oral glucose tolerance test-derived indices and HbA$_{1c}$. A point estimate represents the Spearman's rank correlation coefficient (rho) and corresponding 95% confidence interval. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points. FPG, fasting plasma glucose; IGP, incremental glucose peak; MGP, maximum glucose peak.

### Additional analyses

The reliability for MSG, SD and CV was not materially different after exclusion of participants who returned for the catch-up visit (**Supplemental Figure S2.2**). After exclusion of individuals with less than the reliable amount of recording days, exclusion of individuals who returned for the catch-up visit or exclusion of individuals with recording data gaps, all CGM-derived indices (MSG, SD and SD) still differed statistically significantly between individuals with NGM, prediabetes or type 2 diabetes (**Supplemental Tables S2.10, S2.11** and **S2.12**). The distributions of MSG, SD and CV were generally comparable after exclusion of individuals with less than the reliable amount of recording days, exclusion of individuals who returned for the catch-up visit or exclusion of individuals with recording data gaps (**Supplemental Tables S2.10, S2.11** and **S2.12**).

No substantially different Spearman correlation coefficients were obtained after exclusion of individuals with less than the reliable number of recording days (**Supplemental Figure S2.3**) or with recording data gaps (**Supplemental Figure S2.4**). Although exclusion of individuals who returned for a catch-up visit resulted in lower correlation coefficients (0.1-0.2 on average), it did not affect the original correlation patterns that were observed for the correlations of MSG, SD and CV with FPG, OGTT-derived indices and HbA$_{1c}$ (**Supplemental Figure S2.5**). Last, the reliability, reference value, and correlation results of MAGE, IQR, and MODD are shown in the Supplemental materials (**Supplemental Figures S2.6** and **S2.7**; **Supplemental Table S2.13**).

## Discussion

We investigated—in individuals with NGM, prediabetes or type 2 diabetes—the minimum number of days needed to reliably measure CGM-derived indices, assessed their reference values, and studied their relationship with several established glycemic indices. Our study has three main findings. First, a reliable assessment of MSG, SD and CV can be made with three recording days. Second, MSG, SD and CV reference values were obtained based on individuals with NGM. Third, GV indices correlated most strongly with the 1-hour OGTT value, IGP and maximum glucose peak.

In our study population, the common seven-day recording period—with six functional CGM days—yielded high reliability, especially for MSG. In addition, a reliable assessment (R>0.80) of CGM-derived indices could be made with relatively few days: one to three days depending on the individual's GMS and the specific index used. Our conclusion contrasts with the sampling duration findings of comparable studies (i.e., 12-15 days are minimally required)[13-15]. However, comparison is complicated by the fact that all prior research was performed in individuals with type 1 diabetes. Since this group is characterized by larger day-to-day differences in their glycemic profiles (i.e., higher intra-individual variance)[16], a longer sampling duration is presumably required to achieve sufficient reliability. The differences in statistical methods (i.e., correlation,

percentage difference analyses) and cut-off values (i.e., $r^2 > 0.70$-$0.86$; overall difference <10%) used, further hamper comparison.

In line with several other studies, we found that MSG and SD values were statistically significantly higher with deteriorating GMS[17-20]; more interestingly, we observed statistically significantly higher CV values as well. Since CV quantifies GV adjusted for MSG, it can be concluded that GV is higher in individuals with type 2 diabetes as compared to prediabetes and NGM, regardless of the mean glucose value. The MSG, SD and CV values observed in the present study were generally lower than previously reported[17-19]. In contrast to the current study, some of these studies did not observe a statistically significant differences between individuals with NGM and prediabetes[17, 19], which could be attributed to their relatively small sample sizes (n=50 - 162).

We established reference values for MSG, SD and CV based on the 2.5th-97.5th percentiles of individuals with NGM, which is preferred when reporting reference intervals[4]. While Bode et al. first aimed to establish normative values in individuals with type 2 diabetes[16], Zhou et al. were the first to report the distribution (5th − 95th percentile) of MSG (4.69 - 6.60 mmol/L) and SD (0.35 - 1.40 mmol/L) in individuals with NGM[21, 22]. These values were, however, not determined under normal living conditions, as the participants were given dietary instructions as part of the study design. This could explain the lower 5th percentile values for both indices. Gude et al. also reported on SD (5th − 95th percentile) in a population-based cohort of Caucasian individuals with and without type 2 diabetes (0.72 - 4.83 mmol/L and 0.50 - 1.33 mmol/L, respectively)[23]. Although they did not specifically report reference ranges for NGM, their distributions were generally comparable to ours, except for the substantially higher 95th percentile they observed in individuals with type 2 diabetes, who had a higher mean $HbA_{1c}$ (54.0 ± 10.8 mmol/mol) and more frequently used oral glucose-lowering drugs (86%) and insulin (21%).

Of all glycemic indices, MSG correlated best, albeit moderately, with FPG, maximum glucose peak and $HbA_{1c}$. Zhou et al. previously reported similar correlations coefficients for MSG with FPG and $HbA_{1c}$[24]. We additionally found that, of all glycemic indices, the strongest correlations of SD and CV were observed with the 1-hour OGTT value, maximum glucose peak and IGP. The GMS-stratified correlation coefficients were generally lower than those for the total study population, which can be attributed to the effect of range restriction (i.e., the smaller ranges in the subgroups attenuate the correlation coefficients)[25]. Interestingly, not only SD, but also CV correlated quite moderately with indices of mean glycemia (i.e., FPG, $HbA_{1c}$). This may explain the difficulty of establishing effects of specific GV improvements while maintaining similar $HbA_{1c}$, FPG or MSG concentrations[26].

The correlations found with SD and CV legitimate the use of the 1-hour or peak glucose value during an OGTT to approximate GV if CGM is unavailable (e.g., in a large observational study[27]). This sheds an interesting light on prior research that found that

the 1-hour OGTT value was associated with cardiovascular disease and mortality[28]. Nevertheless, we observed that the strength of the correlation with GV indices was determined by the time point of the OGTT peak. The 1-hour OGTT value was often not reflective of the glucose peak, especially in individuals with NGM (15.3%) and T2D (20.3%) (**Supplemental Table S2.7** and **S2.8**). Accordingly, if the total study population and GMS-stratified groups are taken together, IGP generally correlated most strongly with SD and CV. This suggests that IGP, rather than the 1-hour OGTT value, is the preferred OGTT-derived index to use as GV proxy.

Our findings have important implications for the study of GV. First, the CGM period may be shortened to three functional days of recording for individuals with NGM, prediabetes or type 2 diabetes in order to alleviate participant burden without compromising reliability of GV assessment. Importantly, a longer recording period may be required in individuals with type 2 diabetes who are treated with multiple daily injections (not highly represented in this study). In addition, a reliable assessment of other CGM-derived indices (e.g., time in range) could require more recording days. Second, our reference ranges can be used to classify individuals as having normal or abnormal MSG, SD or CV values. However, these reference values need not necessarily represent target values, as is the case for $HbA_{1c}$[29]. Third, IGP should be viewed as the preferred OGTT-derived index to assess GV, since it correlated most strongly with SD and CV for both the total study population and the GMS-stratified groups. Last, the regression coefficients found could be used to approximate CGM-derived indices based on several established glycemic indices (**Supplemental Table S2.9**).

This study has several strengths and limitations. Strengths are 1) the large, population-based study sample; 2) the high overall SMBG adherence, ensuring adequate CGM calibration; 3) the OGTT-based assessment of GMS, allowing the calculation of reference values; 4) the availability of a seven-point OGTT, enabling comparison with multiple time points; and 5) the robustness of the results, as reflected by the consistency of multiple sensitivity analyses. Our study had certain limitations. First, the reliability analyses would ideally have been based on a longer recording period than six functional days, as this could have improved the assessment of intra-individual variation. We were, however, limited by our study design, which aimed to prevent excessive participant strain and was restricted by sensor lifespan[6]. Second, we focused mainly on SD and CV as indices of GV, while others exist (e.g., IQR, MAGE). With the use of MSG, SD, and CV we assume a normal distribution of CGM-derived glucose values, which is not necessarily true. Nevertheless, SD was shown to have a smaller random sampling error than other GV indices, is widely used and easily interpretable, and can be adapted to the preferred, MSG-adjusted GV index (i.e., CV)[2,3]. Additionally, the different CGM-derived GV indices have been found to be strongly correlated[3]. Accordingly, the MAGE, IQR, and MODD findings were generally comparable to the SD and CV results (**Supplemental Figures S2.6** and **S2.7**; **Supplemental Table S2.13**). Third, the strength of the correlation coefficients of our

total study population could have been underestimated due to the inclusion of individuals who returned for a catch-up visit, as there was a median time of 2.1 years between CGM and the measurement of the OGTT and HbA$_{1c}$ values. While the correlations were investigated in only newly recruited individuals (**Supplemental Figure S2.5**), their interpretation is hampered by range restriction (i.e., lower number of individuals with prediabetes and type 2 diabetes in the regular visit when compared to the catch-up visit)[25]. Fourth, the fact that the individuals with type 2 diabetes were relatively well-controlled (i.e., mean HbA$_{1c}$: 49.2 ± 10.8 mmol/mol) and uncommonly used insulin (n=19, 9.6%), may also have influenced our findings. Namely, a wider range of glycemic control could have altered the reliability of CGM-derived indices in individuals with type 2 diabetes by affecting the inter- and intra-individual variance ratio. Additionally, percentile distributions of the CGM-derived indices are inherently determined by the glycemic status of the included individuals. The generalizability of our results is additionally affected by the age-restricted and predominately Caucasian study population, and by the absence of individuals with type 1 diabetes.

In conclusion, we showed that a sufficiently reliable assessment of MSG, SD, and CV can be made with three recording days. In addition, we established the reference ranges for MSG, SD, and CV as 5.03 - 6.69 mmol/L, 0.44 - 1.37 mmol/L, and 7.74 - 22. 45 %, respectively. We additionally found that the 1-hour OGTT value, IGP and maximum glucose peak correlated moderately with GV indices. These findings are relevant for future studies that aim to use CGM to measure GV, or current studies that only have OGTT-derived indices at their disposal to assess GV.

# References

1.     Brownlee M, Hirsch IB. Glycemic variability: a hemoglobin A1c-independent risk factor for diabetic complications. JAMA. 2006;295(14):1707-8.
2.     Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care. 2017;40(12):1631-40.
3.     Rodbard D. The challenges of measuring glycemic variability. J Diabetes Sci Technol. 2012;6(3): 712-5.
4.     Horn PS, Pesce AJ. Reference intervals: an update. Clin Chim Acta. 2003;334(1-2):5-23.
5.     Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.
6.     Rodbard D. Continuous Glucose Monitoring: A Review of Successes, Challenges, and Opportunities. Diabetes Technol Ther. 2016;18 Suppl 2:S3-S13.
7.     Klaff LJ, Brazg R, Hughes K, Tideman AM, Schachner HC, Stenger P, et al. Accuracy evaluation of contour next compared with five blood glucose monitoring systems across a wide range of blood glucose concentrations occurring in a clinical research setting. Diabetes Technol Ther. 2015;17(1): 8-15.
8.     Halldorsdottir S, Warchal-Windham ME, Wallace JF, Pardo S, Parkes JL, Simmons DA. Accuracy evaluation of five blood glucose monitoring systems: the North American comparator trial. J Diabetes Sci Technol. 2013;7(5): 1294-304.
9.     Ekhlaspour L, Mondesir D, Lautsch N, Balliro C, Hillard M, Magyar K, et al. Comparative Accuracy of 17 Point-of-Care Glucose Meters. J Diabetes Sci Technol. 2017;11(3):558-66.
10.    WHO. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. WHO. 2006.
11.    Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc. 2005;37(11 Suppl):S531-43.
12.    Matthews CE, Ainsworth BE, Thompson RW, Bassett DR, Jr. Sources of variance in daily physical activity levels as measured by an accelerometer. Med Sci Sports Exerc. 2002;34(8):1376-81.
13.    Xing D, Kollman C, Beck RW, Tamborlane WV, Laffel L, Buckingham BA, et al. Optimal sampling intervals to assess long-term glycemic control using continuous glucose monitoring. Diabetes Technol Ther. 2011;13(3):351-8.
14.    Neylon OM, Baghurst PA, Cameron FJ. The Minimum Duration of Sensor Data From Which Glycemic Variability Can Be Consistently Assessed. J Diabetes Sci Technol. 2014;8(2):273-6.
15.    Riddlesworth TD, Beck RW, Gal RL, Connor CG, Bergenstal RM, Lee S, et al. Optimal Sampling Duration for Continuous Glucose Monitoring to Determine Long-Term Glycemic Control. Diabetes Technol Ther. 2018;20(4):314-6.
16.    Bode BW, Schwartz S, Stubbs HA, Block JE. Glycemic characteristics in continuously monitored patients with type 1 and type 2 diabetes: normative values. Diabetes Care. 2005;28(10):2361-6.
17.    Wang C, Lv L, Yang Y, Chen D, Liu G, Chen L, et al. Glucose fluctuations in subjects with normal glucose tolerance, impaired glucose regulation and newly diagnosed type 2 diabetes mellitus. Clin Endocrinol (Oxf). 2012;76(6):810-5.
18.    Hanefeld M, Sulk S, Helbig M, Thomas A, Kohler C. Differences in Glycemic Variability Between Normoglycemic and Prediabetic Subjects. J Diabetes Sci Technol. 2014;8(2):286-90.
19.    Chakarova N, Dimova R, Grozeva G, Tankova T. Assessment of glucose variability in subjects with prediabetes. Diabetes Res Clin Pract. 2019;151:56-64.
20.    Rodbard D. Increased glycemic variability at the onset and during progression of type 2 diabetes-commentary. Diabetes Technol Ther. 2013;15(6):445-7.
21.    Zhou J, Li H, Ran X, Yang W, Li Q, Peng Y, et al. Establishment of normal reference ranges for glycemic variability in Chinese subjects using continuous glucose monitoring. Med Sci Monit. 2011;17(1):CR9-13.
22.    Zhou J, Li H, Ran X, Yang W, Li Q, Peng Y, et al. Reference values for continuous glucose monitoring in Chinese subjects. Diabetes Care. 2009;32(7):1188-93.

23. Gude F, Diaz-Vidal P, Rua-Perez C, Alonso-Sampedro M, Fernandez-Merino C, Rey-Garcia J, et al. Glycemic Variability and Its Association With Demographics and Lifestyles in a General Adult Population. J Diabetes Sci Technol. 2017;11(4):780-90.

24. Zhou J, Mo Y, Li H, Ran X, Yang W, Li Q, et al. Relationship between HbA1c and continuous glucose monitoring in Chinese population: a multicenter study. PLoS One. 2013;8(12):e83827.

25. Bland JM, Altman DG. Correlation in restricted ranges of data. BMJ. 2011;342:d556.

26. Investigators F-ST. Glucose Variability in a 26-Week Randomized Comparison of Mealtime Treatment With Rapid-Acting Insulin Versus GLP-1 Agonist in Participants With Type 2 Diabetes at High Cardiovascular Risk. Diabetes Care. 2016;39(6):973-81.

27. Foreman YD, Brouwers MCGJ, Berendschot TTJM, van Dongen MCJM, Eussen SJPM, van Greevenbroek MMJ, et al. The oral glucose tolerance test-derived incremental glucose peak is associated with greater arterial stiffness and maladaptive arterial remodeling: The Maastricht Study. Cardiovascular Diabetology. 2019; 18(1):152.

28. Bergman M, Jagannathan R, Buysschaert M, Pareek M, Olsen MH, Nilsson PM, et al. Lessons learned from the 1-hour post-load glucose level during OGTT: Current screening recommendations for dysglycaemia should be revised. Diabetes Metab Res Rev. 2018;34(5):e2992.

29. American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2019. Diabetes Care. 2019;42(Suppl 1):S61-S70.

2

# Supplemental materials

## *Supplementary methods*

### *Statistical analysis*

We compared body weight and medication use of catch-up participants by use of paired T test and chi-square analysis. We used the Glycemic Variability Research Tool (GlyVaRT; Medtronic) software to additionally calculate mean amplitude of glucose excursions (MAGE)[1], interquartile range (IQR), and mean of daily differences (MODD)[2].

## References

1.  Service FJ, Molnar GD, Rosevear JW, Ackerman E, Gatewood LC, Taylor WF. Mean amplitude of glycemic excursions, a measure of diabetic instability. Diabetes. 1970;19(9):644-55.
2.  Molnar GD, Taylor WF, Ho MM. Day-to-day variation of continuously monitored glycaemia: a further measure of diabetic instability. Diabetologia. 1972;8(5):342-8.

## Supplementary results
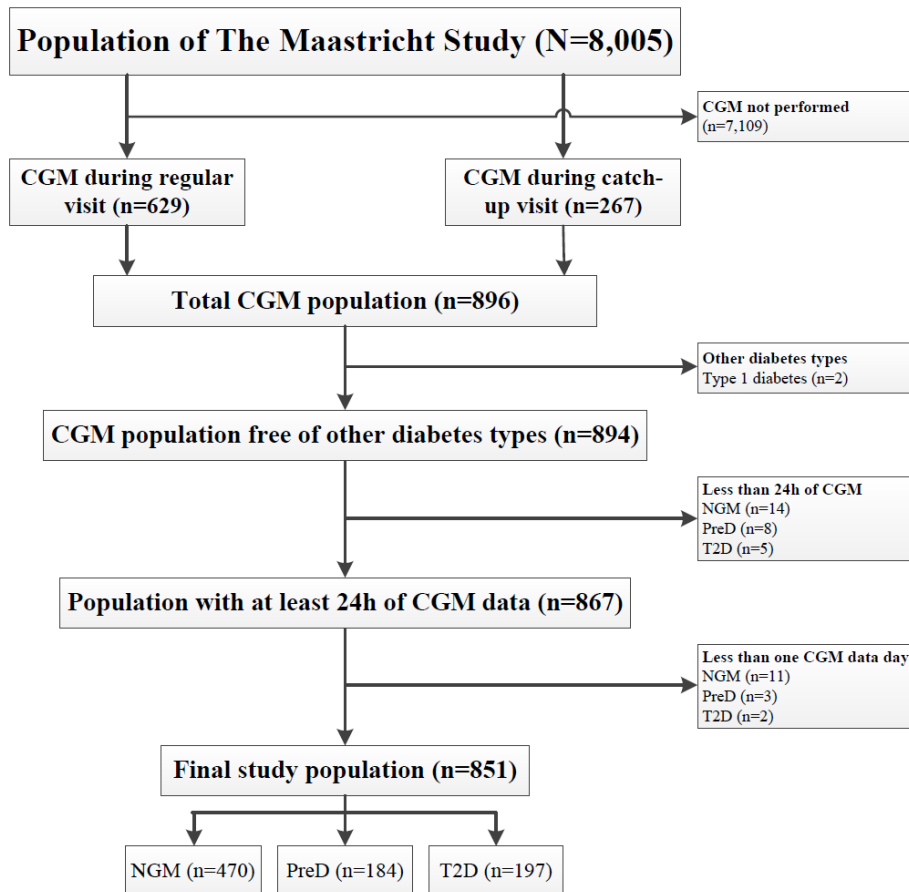


**Figure S2.1  Flowchart of the CGM study population selection process.** CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes.
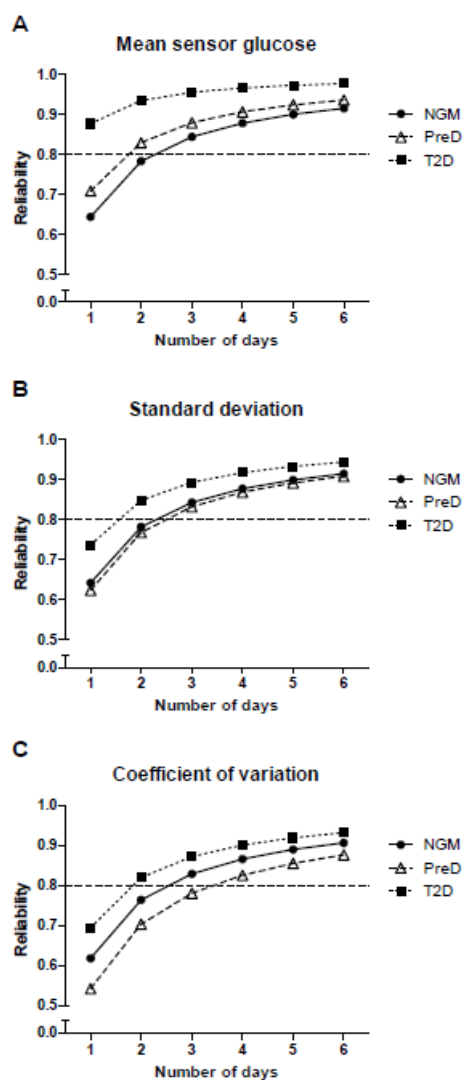
**Figure S2.2 Reliability for mean sensor glucose (A), standard deviation (B), and coefficient of variation (C), per number of consecutive monitoring days after exclusion of individuals who returned for a catch-up visit.** The dashed horizontal line represents the cut-off for sufficient reliability (R>0.80). NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes.
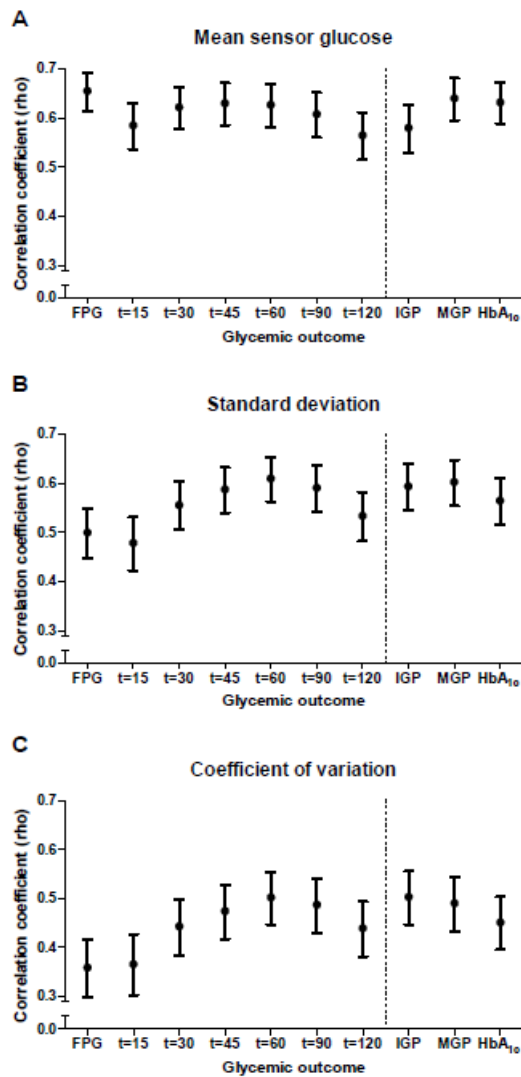
**Figure S2.3  Correlations of mean sensor glucose (A), standard deviation (B) and coefficient of variation (C) with oral glucose tolerance test-derived indices and HbA1c after exclusion of individuals with less than three recording days.** A point estimate represents the Spearman's rank correlation coefficient (rho) and corresponding 95% confidence interval. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points. FPG, fasting plasma glucose; IGP, incremental glucose peak; MGP, maximum glucose peak.
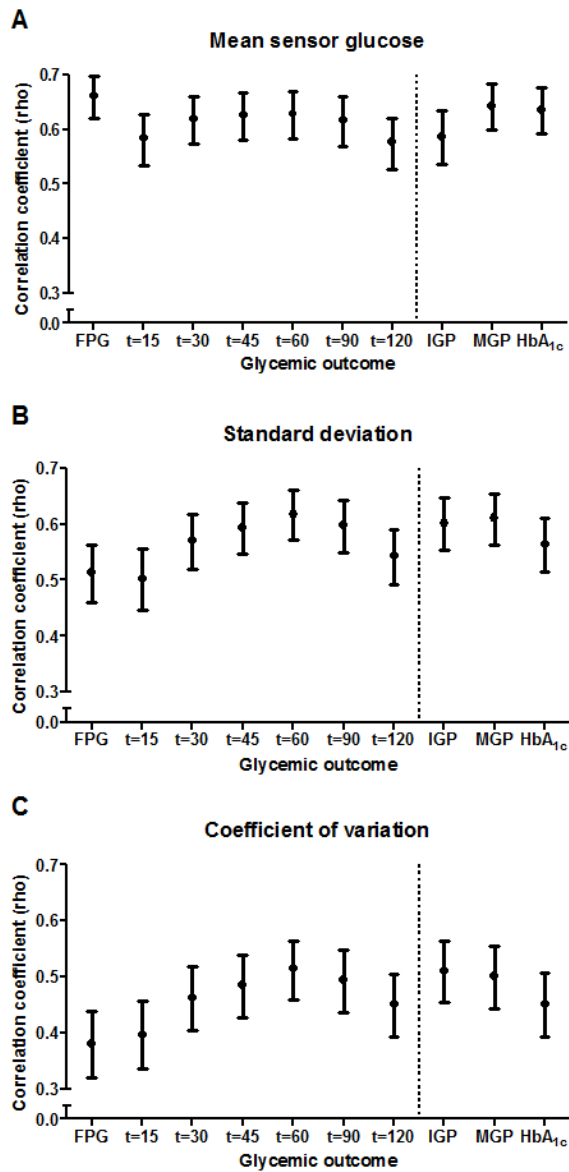
**Figure S2.4** **Correlations of mean sensor glucose (A), standard deviation (B) and coefficient of variation (C) with oral glucose tolerance test-derived indices and HbA1c after exclusion of individuals with data gaps during the continuous glucose monitoring recording.** A point estimate represents the Spearman's rank correlation coefficient (rho) and corresponding 95% confidence interval. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points. FPG, fasting plasma glucose; IGP, incremental glucose peak; MGP, maximum glucose peak.

**A**

### Mean sensor glucose



**B**

### Standard deviation



**C**

### Coefficient of variation
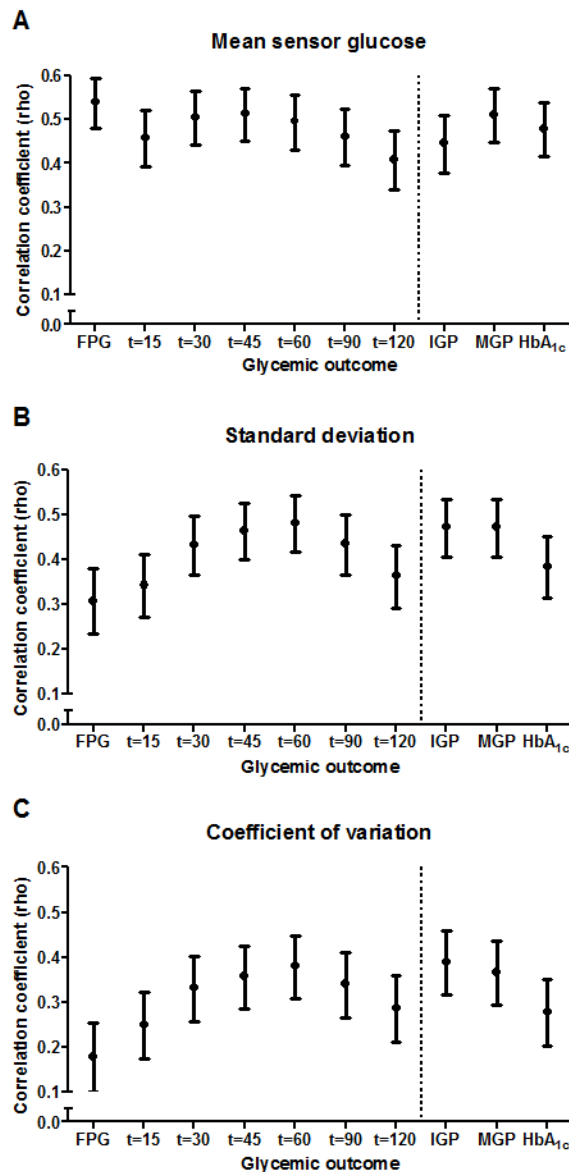


**Figure S2.5 Correlations of mean sensor glucose (A), standard deviation (B) and coefficient of variation (C) with oral glucose tolerance test-derived indices and HbA1c after exclusion of individuals who underwent continuous glucose monitoring as part of a catch-up visit.** A point estimate represents the Spearman's rank correlation coefficient (rho) and corresponding 95% confidence interval. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points. FPG, fasting plasma glucose; IGP, incremental glucose peak; MGP, maximum glucose peak.

**A**

**MAGE**

**B**

**IQR**
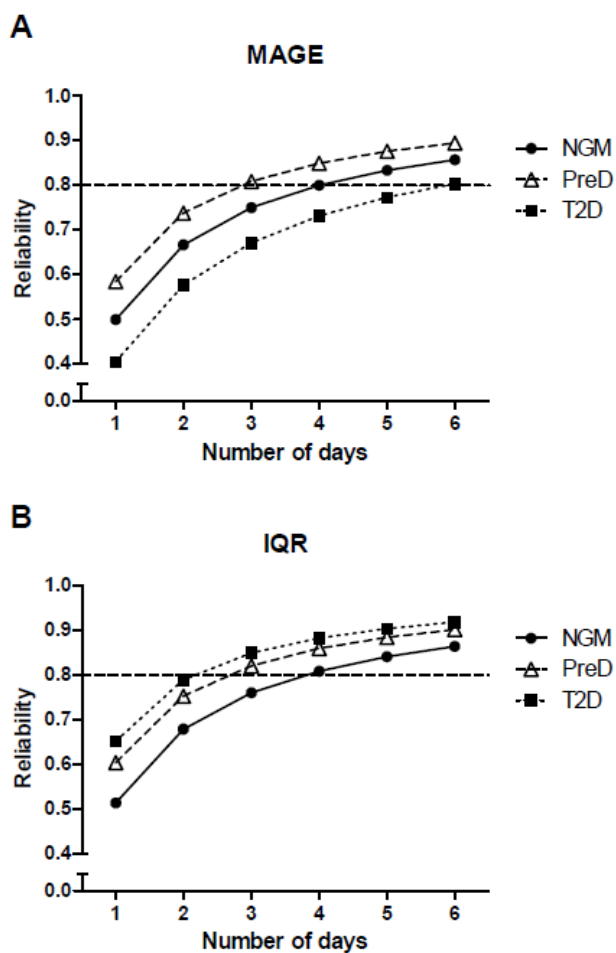
**Figure S2.6 Reliability for MAGE (A) and IQR (B) per number of consecutive monitoring days.** The dashed horizontal line represents the cut-off for sufficient reliability (R>0.80). MAGE, mean amplitude of glucose excursions; IQR, interquartile range; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes. Of note, reliability for MODD could not be calculated.
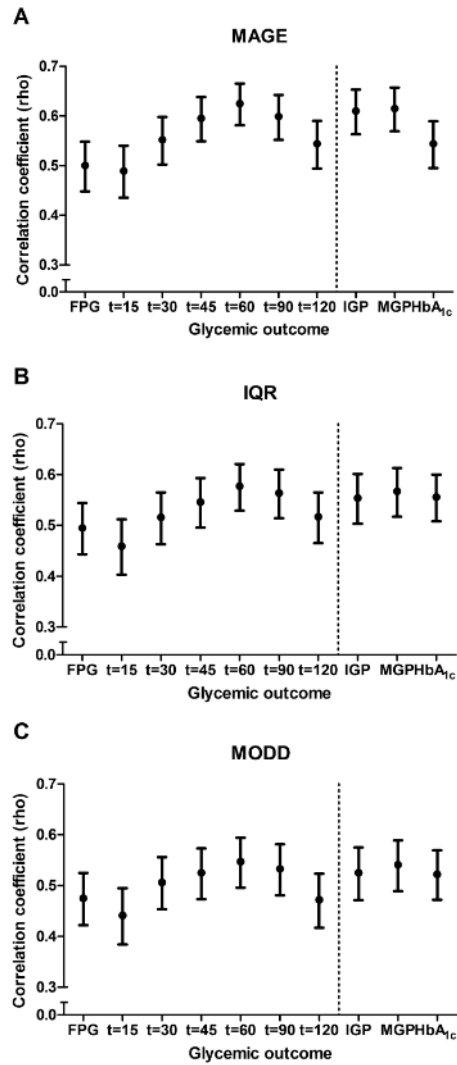
**Figure S2.7 Correlations of MAGE (A), IQR (B) and MODD (C) with oral glucose tolerance test-derived indices and HbA1c.** A point estimate represents the Spearman's rank correlation coefficient (rho) and corresponding 95% confidence interval. MAGE, mean amplitude of glucose excursions; IQR, interquartile range; MODD, mean of daily differences; NGM, normal glucose metabolism. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points. FPG, fasting plasma glucose; IGP, incremental glucose peak; MGP, maximum glucose peak.

Table S2.1 Continuous glucose monitoring properties of the total study population and of glucose metabolism status-stratified groups.

| | Total (n=851) | NGM (n=470) | PreD (n=184) | Type 2 (n=197) |
|---|---|---|---|---|
| Participants with CGM recording data on: | | | | |
| Data day 1 (CGM day 2) | 846 (99.4) | 465 (98.9) | 184 (100) | 197 (100) |
| Data day 2 (CGM day 3) | 849 (99.8) | 468 (99.6) | 184 (100) | 197 (100) |
| Data day 3 (CGM day 4) | 841 (98.8) | 465 (98.9) | 182 (98.9) | 194 (98.5) |
| Data day 4 (CGM day 5) | 821 (96.5) | 456 (97.0) | 175 (95.1) | 190 (96.4) |
| Data day 5 (CGM day 6) | 803 (94.4) | 444 (94.5) | 171 (92.9) | 188 (95.4) |
| Data day 6 (CGM day 7) | 772 (90.7) | 426 (90.6) | 164 (89.1) | 182 (92.4) |
| Consecutive, complete CGM data days | | | | |
| 0 | 24 (2.8) | 15 (3.2) | 4 (2.2) | 5 (2.5) |
| 1 | 19 (2.2) | 10 (2.1) | 5 (2.7) | 4 (2.0) |
| 2 | 26 (3.1) | 13 (2.8) | 8 (4.3) | 5 (2.5) |
| 3 | 30 (3.5) | 17 (3.6) | 6 (3.3) | 7 (3.6) |
| 4 | 31 (3.6) | 15 (3.2) | 10 (5.4) | 6 (3.0) |
| 5 or more | 721 (84.7) | 400 (85.1) | 151 (82.1) | 170 (86.3) |
| Individuals with at least one data gap during CGM | 66 (7.8) | 36 (7.7) | 16 (8.7) | 14 (7.1) |
| Data gap length, h | 5.3 [0.9-19.6] | 7.9 [0.9-24.8] | 3.1 [0.9-14.7] | 3.6 [0.8-12.9] |
| Average daily calibration frequency | | | | |
| Mean ± SD | 3.9 ± 0.5 | 3.8 ± 0.5 | 3.9 ± 0.5 | 3.9 ± 0.5 |
| Median [IQR] | 4.0 [3.7 – 4.1] | 4.0 [3.7 – 4.1] | 4.0 [3.9 – 4.1] | 4.0 [3.9 – 4.1] |

Data are reported as mean ± SD, median [interquartile range] or number (percentage [%]) as appropriate. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes

**Table S2.2   Characteristics of individuals who underwent CGM at the regular or catch-up visit.**

| Characteristic | Regular visit | | | | Catch-up visit | | | |
|---|---|---|---|---|---|---|---|---|
| | Total (n=603) | NGM (n=423) | PreD (n=108) | T2D (n=72) | Total (n=248) | NGM (n=47) | PreD (n=76) | T2D (n=125) |
| Age, years | 59.2 ± 8.7 | 58.2 ± 8.9 | 61.1 ± 8.4 | 61.9 ± 6.8 | 61.7 ± 8.3 | 58.1 ± 8.7 | 62.2 ± 7.7 | 62.7 ± 8.3 |
| Women, n (%) | 312 (51.7) | 236 (55.8) | 50 (46.3) | 26 (36.1) | 106 (42.7) | 30 (63.8) | 33 (43.4) | 43 (34.4) |
| BMI, kg/m$^2$ | 26.5 ± 4.1 | 25.6 ± 3.6 | 28.7 ± 4.4 | 28.8 ± 4.3 | 28.7 ± 4.8 | 25.6 ± 3.6 | 28.1 ± 4.3 | 30.2 ± 4.8 |
| Newly diagnosed T2D, n (%) | 36 (6.0) | - | - | 36 (50) | 34 (13.7) | - | - | 34 (27.2) |
| HbA$_{1c}$, % | 5.6 ± 0.6 | 5.4 ± 0.3 | 5.6 ± 0.4 | 6.5 ± 0.9 | 6.2 ± 1.0 | 5.3 ± 0.3 | 5.7 ± 0.4 | 6.8 ± 1.0 |
| HbA$_{1c}$, mmol/mol | 37.2 ± 6.1 | 35.5 ± 3.4 | 37.2 ± 4.3 | 47.1 ± 10.2 | 43.8 ± 10.7 | 34.9 ± 3.1 | 38.6 ± 3.9 | 50.4 ± 10.8 |
| Fasting plasma glucose, mmol/L | 5.3 [4.9 – 5.8] | 5.1 [4.8 – 5.4] | 5.9 [5.3 – 6.2] | 7.2 [6.3 – 8.4] | 6.4 [5.5 – 7.4] | 5.0 [4.8 – 5.4] | 6.1 [5.5 – 6.4] | 7.3 [6.6 – 8.1] |
| OGTT-derived indices | | | | | | | | |
| 15-minute, mmol/L | 7.0 [6.4 – 7.8] | 6.7 [6.2 – 7.3] | 7.7 [6.9 – 8.5] | 9.1 [8.1 – 10.7] | 8.4 [7.3 – 9.5] | 6.9 [6.2 – 7.8] | 7.9 [7.2 – 8.5] | 9.5 [8.7 – 10.2] |
| 30-minute, mmol/L | 8.3 [7.3 – 9.7] | 7.8 [7.0 – 8.8] | 9.8 [8.5 – 10.7] | 11.4 [10.1 – 13.5] | 10.5 [8.8 – 12.1] | 7.8 [6.7 – 8.7] | 9.7 [9.0 – 10.6] | 12.1 [10.8 – 13.1] |
| 45-minute, mmol/L | 8.6 [7.1 – 10.5] | 7.9 [6.7 – 9.0] | 10.6 [9.2 – 12.1] | 13.0 [11.8 – 15.6] | 12.0 [9.8 – 14.0] | 7.9 [6.7 – 9.3] | 10.8 [9.9 – 12.2] | 14.3 [12.5 – 15.8] |
| 1-hour, mmol/L | 8.2 [6.3 – 10.4] | 7.1 [5.7 – 8.7] | 10.8 [9.2 – 12.4] | 14.0 [12.1 – 17.0] | 12.3 [9.8 – 15.3] | 7.2 [6.0 – 8.4] | 11.2 [10.2 – 12.6] | 15.5 [13.2 – 16.9] |
| 90-minute, mmol/L | 7.1 [5.5 – 9.2] | 6.2 [5.2 – 7.4] | 9.7 [8.6 – 11.2] | 14.3 [12.4 – 17.7] | 11.6 [7.9 – 15.6] | 6.1 [5.0 – 6.9] | 10.4 [8.7 – 11.5] | 16.0 [13.8 – 18.0] |
| 2-hour, mmol/L | 6.2 [5.0 – 7.7] | 5.5 [4.7 – 6.5] | 8.4 [7.7 – 9.3] | 12.4 [11.1 – 16.6] | 9.1 [6.3 – 13.7] | 5.3 [4.2 – 5.8] | 8.3 [7.3 – 9.2] | 14.0 [11.9 – 16.1] |
| MGP, mmol/L | 9.1 [7.8 – 11.0] | 8.3 [7.4 – 9.4] | 11.2 [10.0 – 12.6] | 14.6 [12.9 – 18.2] | 12.8 [10.3 – 16.0] | 8.2 [7.5 – 9.5] | 11.7 [10.5 – 12.7] | 16.2 [14.4 – 18.0] |
| IGP, mmol/L | 3.8 [2.7 – 5.5] | 3.2 [2.2 – 4.2] | 5.5 [4.2 – 6.8] | 8.3 [6.8 – 10.2] | 6.7 [4.4 – 9.0] | 3.1 [2.4 – 4.0] | 5.8 [4.5 – 6.8] | 9.1 [7.9 – 10.5] |
| CGM-derived indices | | | | | | | | |
| MSG, mmol/L | 5.9 [5.6 – 6.3] | 5.8 [5.5 – 6.1] | 6.2 [5.7 – 6.6] | 7.3 [6.5 – 8.2] | 6.8 [6.0 – 7.9] | 5.8 [5.5 – 6.0] | 6.5 [5.9 – 6.9] | 7.8 [6.8 – 8.9] |
| SD, mmol/L | 0.78 [0.65 – 1.00] | 0.73 [0.63 – 0.88] | 0.89 [0.73 – 1.12] | 1.46 [0.97 – 1.95] | 1.14 [0.79 – 1.67] | 0.69 [0.58 – 0.80] | 0.92 [0.75 – 1.14] | 1.55 [1.23 – 1.95] |
| CV, % | 13.2 [11.2 – 16.6] | 12.7 [10.9 – 15.1] | 14.8 [12.2 – 17.5] | 19.0 [14.5 – 24.1] | 16.2 [13.1 – 20.7] | 11.6 [10.8 – 14.9] | 15.0 [12.3 – 17.4] | 19.3 [16.4 – 24.0] |
| Diabetes medication use, n (%) | 30 (5.0) | - | - | 30 (41.7) | 79 (32) | - | - | 79 (63.7) |
| Insulin | 4 (0.7) | - | - | 4 (5.6) | 15 (6.0) | - | - | 15 (12.0) |
| Metformin | 30 (5.0) | - | - | 30 (41.7) | 74 (30.0) | - | - | 74 (59.7) |
| Sulfonylureas | 7 (1.2) | - | - | 7 (9.7) | 14 (5.7) | - | - | 14 (11.3) |
| Thiazolidinediones | 0 (0) | - | - | 0 (0) | 0 (0) | - | - | 0 (0) |
| GLP-1 analogs | 1 (0.2) | - | - | 1 (1.4) | 2 (0.8) | - | - | 2 (1.6) |
| DDP-4 inhibitors | 0 (0) | - | - | 0 (0) | 1 (0.4) | - | - | 1 (0.8) |

Data are reported as mean ± SD, median [interquartile range], or number (percentage [%]) as appropriate. Data represent the study population of participants with at least one data day of continuous glucose monitoring data. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; BMI, body mass index; FPG, fasting plasma glucose; OGTT, oral glucose tolerance test; MGP, maximum glucose peak; IGP, incremental glucose peak; HbA1c, glycated hemoglobin A1c; CGM, continuous glucose monitoring; MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4.

Table S2.3 Characteristics during regular and catch-up visit of individuals who returned for a catch-up visit.

| Characteristic | During regular visit | | | | | During catch-up visit | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Total (n=248) | NGM (n=47) | PreD (n=76) | T2D (n=125) | Total | NGM | PreD | T2D |
| Time between regular and catch-up visit, years | 2.1 [2.0 – 2.2] | 2.1 [2.0 – 2.3] | 2.1 [2.0 – 2.2] | 2.1 [2.1 – 2.3] | | - | - | - | - |
| Body weight, kg | 84.6 ± 17.6 | 73.9 ± 12.1 | 82.9 ± 15.6 | 89.7 ± 18.5 | 85.1 ± 17.7 | 74.8 ± 13.2 | 84.1 ± 16.2* | 89.6 ± 18.5 |
| Participants with more than +10% weight difference | - | - | - | - | 7 (2.8) | 2 (4.3) | 1 (1.3) | 4 (3.2) |
| Participants with more than -10% weight difference | | | | | 5 (2.0) | 0 (0) | 3 (3.9) | 2 (1.6) |
| Diabetes medication use, n (%) | 79 (32) | - | - | - | 89 (36.6)† | - | - | 89 (71.8)† |
| Insulin | 15 (6.0) | - | - | - | 19 (7.7) | - | - | 19 (15.3) |
| Metformin | 74 (30.0) | - | - | - | 82 (33.7) | - | - | 82 (66.1) |
| Sulfonylureas | 14 (5.7) | - | - | - | 22 (8.9) | - | - | 22 (17.7) |
| Thiazolidinediones | 0 (0) | - | - | - | 0 (0) | - | - | 0 (0) |
| GLP-1 analogs | 2 (0.8) | - | - | - | 2 (0.8) | - | - | 2 (1.6) |
| DDP-4 inhibitors | 1 (0.4) | - | - | - | 4 (1.6) | - | - | 4 (3.2) |

Data are reported as mean ± SD, median [interquartile range], or number (percentage [%]) as appropriate. Data represent the study population of participants with at least one data day of continuous glucose monitoring data. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4. * Body weight was statistically significantly (p=0.011) higher for of individuals with prediabetes (paired samples T test); this was not the case for the total group (p=0.062), individuals with NGM (p=0.053) or individuals with type 2 diabetes (p=0.862). † Diabetes medication use was statistically significantly more frequent for the total group (p=0.021) and for individuals with type 2 diabetes (p=0.035). Differences between medication subgroups were not statistically tested.

**Table S2.4** GMS-stratified median values per data day for mean sensor glucose, standard deviation and coefficient of variation in individuals with more than five consecutive, complete days of CGM.

| CGM-derived index | Data day | NGM (n=400) | PreD (n=151) | T2D (n=170) |
|---|---|---|---|---|
| MSG, mmol/L | 1 | 5.72 [5.42 − 6.01] | 6.13 [5.70 − 6.55] | 7.36 [6.55 − 8.41] |
| | 2 | 5.76 [5.42 − 6.09] | 6.17 [5.68 − 6.74] | 7.44 [6.64 − 8.42] |
| | 3 | 5.86 [5.47 − 6.14] | 6.20 [5.80 − 6.72] | 7.48 [6.71 − 8.72] |
| | 4 | 5.86 [5.59 − 6.19] | 6.26 [5.87 − 6.64] | 7.63 [6.73 − 8.69] |
| | 5 | 5.87 [5.58 − 6.16] | 6.26 [5.85 − 6.67] | 7.55 [6.85 − 8.52] |
| | 6 | 5.90 [5.53 − 6.23] | 6.21 [5.83 − 6.80] | 7.50 [6.81 − 8.57] |
| SD, mmol/L | 1 | 0.68 [0.52 − 0.87] | 0.84 [0.67 − 1.07] | 1.46 [1.05 − 1.89] |
| | 2 | 0.63 [0.48 − 0.79] | 0.76 [0.57 − 1.02] | 1.31 [0.92 − 1.80] |
| | 3 | 0.62 [0.49 − 0.82] | 0.83 [0.64 − 1.05] | 1.32 [0.99 − 1.71] |
| | 4 | 0.62 [0.48 − 0.78] | 0.82 [0.63 − 1.05] | 1.35 [1.00 − 1.78] |
| | 5 | 0.64 [0.50 − 0.80] | 0.81 [0.60 − 1.01] | 1.36 [1.01 − 1.73] |
| | 6 | 0.63 [0.49 − 0.83] | 0.75 [0.57 − 1.00] | 1.31 [0.88 − 1.90] |
| CV, % | 1 | 11.8 [9.4 − 14.8] | 13.7 [11.3 − 17.3] | 18.4 [15.1 − 23.1] |
| | 2 | 10.9 [8.4 − 13.5] | 12.6 [9.5 − 16.5] | 17.0 [13.3 − 21.7] |
| | 3 | 10.7 [8.6 − 13.7] | 13.2 [10.8 − 16.4] | 16.4 [12.8 − 20.8] |
| | 4 | 10.5 [8.3 − 13.3] | 12.9 [10.3 − 16.3] | 17.0 [13.2 − 22.2] |
| | 5 | 11.0 [8.8 − 13.8] | 13.0 [10.4 − 15.9] | 16.8 [13.2 − 21.4] |
| | 6 | 10.9 [8.5 − 13.8] | 12.0 [9.2 − 15.3] | 16.8 [12.5 − 22.5] |

Data are reported as median [interquartile range]. Data represent the study population of participants with more than five consecutive, complete days of CGM. GMS, glucose metabolism status; CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation.

2

**Table S2.5** Distribution of mean sensor glucose, standard deviation and coefficient of variation in age- or sex-stratified glucose metabolism status groups.

| | Stratified according to age group | | | | | | | | | | | | Stratified according to sex | | | | | |
| | MSG, mmol/L | | | | SD, mmol/L | | | | CV, % | | | | MSG, mmol/L | | SD, mmol/L | | CV, % | |
| | <50y | 50-59y | 60-69y | >69y | <50y | 50-59y | 60-69y | >69y | <50y | 50-59y | 60-69y | >69y | Male | Female | Male | Female | Male | Female |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NGM n | 97 | 147 | 179 | 47 | 97 | 147 | 179 | 47 | 97 | 147 | 179 | 47 | 204 | 266 | 204 | 266 | 204 | 266 |
| Median | 5.68 | 5.78 | 5.89 | 6.05 | 0.70 | 0.72 | 0.74 | 0.74 | 12.55 | 12.53 | 12.90 | 12.28 | 5.91 | 5.76 | 0.74 | 0.72 | 12.73 | 12.55 |
| 5th percentile | 5.09 | 5.17 | 5.27 | 5.03 | 0.42 | 0.48 | 0.52 | 0.43 | 7.44 | 8.64 | 8.94 | 7.57 | 5.14 | 5.15 | 0.49 | 0.47 | 8.45 | 8.41 |
| 10th percentile | 5.18 | 5.30 | 5.36 | 5.45 | 0.50 | 0.51 | 0.54 | 0.52 | 8.98 | 9.45 | 9.59 | 8.71 | 5.36 | 5.30 | 0.54 | 0.53 | 9.48 | 9.24 |
| 90th percentile | 6.25 | 6.31 | 6.38 | 6.58 | 0.97 | 0.99 | 1.21 | 1.14 | 17.17 | 16.45 | 19.76 | 19.18 | 6.41 | 6.36 | 1.02 | 1.13 | 17.86 | 18.50 |
| 95th percentile | 6.45 | 6.47 | 6.58 | 6.69 | 1.03 | 1.19 | 1.34 | 1.55 | 17.73 | 19.74 | 22.43 | 23.81 | 6.67 | 6.50 | 1.28 | 1.28 | 20.33 | 20.92 |
| PreD n | 17 | 52 | 84 | 31 | 17 | 52 | 84 | 31 | 17 | 52 | 84 | 31 | 101 | 83 | 101 | 83 | 101 | 83 |
| Median | 5.88 | 6.23 | 6.24 | 6.34 | 0.81 | 0.87 | 0.93 | 0.87 | 13.03 | 14.71 | 15.23 | 14.96 | 6.27 | 6.23 | 0.87 | 0.97 | 13.89 | 15.29 |
| 5th percentile | - | 5.57 | 5.08 | 5.23 | - | 0.57 | 0.60 | 0.62 | - | 10.06 | 10.28 | 10.19 | 5.41 | 5.10 | 0.57 | 0.62 | 10.12 | 10.66 |
| 10th percentile | 5.24 | 5.63 | 5.47 | 5.59 | 0.57 | 0.66 | 0.63 | 0.65 | 10.28 | 10.35 | 11.04 | 10.55 | 5.57 | 5.42 | 0.62 | 0.69 | 10.50 | 11.47 |
| 90th percentile | 7.02 | 7.30 | 7.06 | 7.33 | 1.10 | 1.53 | 1.37 | 1.25 | 18.33 | 22.76 | 22.15 | 18.61 | 7.06 | 7.20 | 1.26 | 1.47 | 19.34 | 23.71 |
| 95th percentile | - | 8.13 | 7.51 | 7.98 | - | 1.64 | 1.85 | 2.00 | - | 24.02 | 25.25 | 27.68 | 7.34 | 7.65 | 1.57 | 1.85 | 21.75 | 25.34 |
| T2D n | 12 | 52 | 98 | 35 | 12 | 52 | 98 | 35 | 12 | 52 | 98 | 35 | 128 | 69 | 128 | 69 | 128 | 69 |
| Median | 6.93 | 7.56 | 7.61 | 7.59 | 1.61 | 1.61 | 1.52 | 1.41 | 21.89 | 19.08 | 19.28 | 18.52 | 7.62 | 7.47 | 1.62 | 1.47 | 19.44 | 18.99 |
| 5th percentile | - | 5.81 | 5.71 | 5.98 | - | 0.80 | 0.77 | 0.78 | - | 12.83 | 12.00 | 12.22 | 5.73 | 5.81 | 0.72 | 0.86 | 10.96 | 12.83 |
| 10th percentile | 5.56 | 6.02 | 6.06 | 6.33 | 0.58 | 0.90 | 0.87 | 0.93 | 10.31 | 13.45 | 12.93 | 14.68 | 6.08 | 6.07 | 0.83 | 0.94 | 12.96 | 14.80 |
| 90th percentile | 12.17 | 9.65 | 10.09 | 9.37 | 3.29 | 2.59 | 2.49 | 2.13 | 38.72 | 32.01 | 29.83 | 27.89 | 9.93 | 9.27 | 2.46 | 2.72 | 28.72 | 32.31 |
| 95th percentile | - | 11.97 | 12.48 | 11.59 | - | 3.11 | 3.26 | 2.49 | - | 34.35 | 33.74 | 29.79 | 11.81 | 11.84 | 3.05 | 3.30 | 32.89 | 34.78 |

MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes. In case of - , the value was not calculated due to sample size.

Table S2.6 Correlation coefficients of mean sensor glucose, standard deviation and coefficient of variation with fasting plasma glucose, oral glucose tolerance test-derived indices and HbA1c.

|  | MSG | SD | CV | n |
|---|---|---|---|---|
| FPG | 0.65 (0.61; 0.69) | 0.50 (0.45; 0.55) | 0.36 (0.30; 0.42) | 851 |
| t=15 | 0.57 (0.62; 0.52) | 0.48 (0.42; 0.53) | 0.37 (0.31; 0.43) | 791 |
| t=30 | 0.61 (0.56; 0.65) | 0.55 (0.50; 0.60) | 0.44 (0.39; 0.50) | 792 |
| t=45 | 0.62 (0.57; 0.66) | 0.59 (0.54; 0.63) | 0.47 (0.42; 0.53) | 794 |
| t=60 | 0.62 (0.58; 0.66) | 0.61 (0.56; 0.65) | 0.50 (0.45; 0.55) | 783 |
| t=90 | 0.60 (0.55; 0.64) | 0.59 (0.54; 0.63) | 0.49 (0.43; 0.54) | 765 |
| t=120 | 0.56 (0.51; 0.60) | 0.53 (0.48; 0.58) | 0.44 (0.38; 0.49) | 813 |
| IGP | 0.58 (0.52; 0.62) | 0.59 (0.55; 0.64) | 0.50 (0.45; 0.55) | 744 |
| MGP | 0.63 (0.59; 0.67) | 0.60 (0.55; 0.65) | 0.49 (0.43; 0.54) | 744 |
| HbA1c | 0.63 (0.58; 0.67) | 0.56 (0.51; 0.61) | 0.45 (0.39; 0.50) | 851 |

This table shows the Spearman's rank correlation coefficients (rho) and corresponding 95% confidence intervals. MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; MGP, maximum glucose peak; IGP, incremental glucose peak; $HbA_{1c}$, glycated hemoglobin $A_{1c}$. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points.

Table S2.7 Moment of glucose peak during the oral glucose tolerance test (OGTT) in individuals with a complete 7-point OGTT.

| Peak time point during OGTT | Total study population (n=744) | NGM (n=430) | PreD (n=166) | T2D (n=148) |
|---|---|---|---|---|
| FPG | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| t=15 | 41 (5.5) | 41 (9.5) | 0 (0) | 0 (0) |
| t=30 | 170 (22.8) | 154 (35.8) | 13 (7.8) | 3 (2.0) |
| t=45 | 198 (26.6) | 146 (34.0) | 40 (24.1) | 12 (8.1) |
| t=60 | 172 (23.1) | 66 (15.3) | 76 (45.8) | 30 (20.3) |
| t=90 | 135 (18.1) | 20 (4.7) | 33 (19.9) | 82 (55.4) |
| t=120 | 28 (3.8) | 3 (0.7) | 4 (2.4) | 21 (14.2) |

Data are reported as number (percentage %). OGTT, oral glucose tolerance test; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; FPG, fasting plasma glucose. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points.

**Table S2.8** GMS-stratified correlations of mean sensor glucose, standard deviation and coefficient of variation with oral glucose tolerance test-derived indices and HbA1c.

| | | | FPG | t=15 | t=30 | t=45 | t=60 | t=90 | t=120 | IGP | MGP | HbA1c |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NGM | MSG | rho | 0.37 | 0.24 | 0.30 | 0.31 | 0.28 | 0.26 | 0.21 | 0.27 | 0.35 | 0.29 |
| | | | (0.29; 0.44) | (0.16; 0.33) | (0.22; 0.38) | (0.22; 0.39) | (0.19; 0.36) | (0.17; 0.34) | (0.12; 0.29) | (0.18; 0.35) | (0.26; 0.43) | (0.20; 0.37) |
| | SD | rho | 0.09 | 0.16 | 0.27 | 0.27 | 0.28 | 0.21 | 0.07 | 0.31 | 0.30 | 0.19 |
| | | | (0.00; 0.18) | (0.07; 0.25) | (0.18; 0.35) | (0.19; 0.36) | (0.19; 0.36) | (0.11; 0.29) | (-0.03; 0.16) | (0.22; 0.39) | (0.22; 0.39) | (0.10; 0.28) |
| | CV | rho | -0.01 | 0.12 | 0.20 | 0.21 | 0.22 | 0.15 | 0.02 | 0.25 | 0.23 | 0.11 |
| | | | (-0.10; 0.08) | (0.02; 0.21) | (0.11; 0.29) | (0.12; 0.29) | (0.13; 0.33) | (0.06; 0.24) | (-0.07; 0.11) | (0.16; 0.34) | (0.14; 0.32) | (0.02; 0.20) |
| | | n | 470 | 457 | 456 | 458 | 449 | 438 | 465 | 430 | 430 | 470 |
| PreD | MSG | rho | 0.34 | 0.30 | 0.34 | 0.31 | 0.31 | 0.30 | 0.06 | 0.23 | 0.33 | 0.46 |
| | | | (0.20; 0.46) | (0.16; 0.43) | (0.20; 0.46) | (0.17; 0.44) | (0.17; 0.44) | (0.16; 0.43) | (-0.09; 0.20) | (0.08; 0.37) | (0.19; 0.46) | (0.34; 0.57) |
| | SD | rho | 0.05 | 0.01 | 0.08 | 0.19 | 0.25 | 0.30 | 0.22 | 0.23 | 0.21 | 0.30 |
| | | | (-0.10; 0.19) | (-0.14; 0.16) | (-0.07; 0.22) | (0.04; 0.32) | (0.11; 0.39) | (0.16; 0.43) | (0.08; 0.35) | (0.08; 0.37) | (0.06; 0.36) | (0.17; 0.43) |
| | CV | rho | -0.08 | -0.10 | -0.04 | 0.11 | 0.18 | 0.23 | 0.24 | 0.19 | 0.13 | 0.18 |
| | | | (-0.22; 0.07) | (-0.24; 0.05) | (-0.18; 0.11) | (-0.04; 0.25) | (0.03; 0.32) | (0.09; 0.37) | (0.09; 0.37) | (0.04; 0.34) | (-0.02; 0.28) | (0.03; 0.31) |
| | | n | 184 | 176 | 178 | 178 | 177 | 175 | 183 | 166 | 166 | 184 |
| T2D | MSG | rho | 0.46 | 0.43 | 0.43 | 0.47 | 0.51 | 0.53 | 0.42 | 0.43 | 0.55 | 0.60 |
| | | | (0.35; 0.57) | (0.30; 0.55) | (0.30; 0.55) | (0.33; 0.58) | (0.39; 0.62) | (0.41; 0.64) | (0.29; 0.54) | (0.29; 0.56) | (0.42; 0.65) | (0.50; 0.68) |
| | SD | rho | 0.29 | 0.39 | 0.43 | 0.45 | 0.50 | 0.54 | 0.46 | 0.54 | 0.54 | 0.50 |
| | | | (0.15; 0.41) | (0.24; 0.51) | (0.29; 0.55) | (0.32; 0.57) | (0.38; 0.61) | (0.42; 0.65) | (0.33; 0.57) | (0.42; 0.65) | (0.41; 0.64) | (0.39; 0.60) |
| | CV | rho | 0.10 | 0.30 | 0.35 | 0.36 | 0.40 | 0.44 | 0.37 | 0.49 | 0.43 | 0.33 |
| | | | (-0.04; 0.24) | (0.15; 0.43) | (0.21; 0.48) | (0.22; 0.49) | (0.26; 0.52) | (0.30; 0.56) | (0.24; 0.50) | (0.35; 0.60) | (0.28; 0.55) | (0.20; 0.45) |
| | | n | 197 | 158 | 158 | 158 | 157 | 152 | 165 | 148 | 148 | 197 |

This table shows the Spearman's rank correlation coefficients (rho) and corresponding 95% confidence intervals. MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; MGP, maximum glucose peak; IGP, incremental glucose peak; HbA1c, glycated hemoglobin A1c. The abbreviations t=15, t=30, t=45, t=60, t=90 and t=120 represent the different oral glucose tolerance test time points.

**Table S2.9** Linear regression coefficients of MSG, SD, and CV in the total study population and in GMS-stratified groups.

| Total study population | Glycaemic index | Constant | Regression coefficient |
|---|---|---|---|
| MSG, mmol/L | FPG, mmol/L | 2.828 (2.588; 3.067) | 0.609 (0.570; 0.649) |
| | MGP, mmol/L | 4.216 (4.058; 4.375) | 0.193 (0.178; 0.207) |
| | HbA$_{1c}$, mmol/mol | 2.093 (1.824; 2.362) | 0.110 (0.104; 0.117) |
| | 45-minute, mmol/L | 4.318 (4.148; 4.488) | 0.200 (0.184; 0.217) |
| | 1-hour, mmol/L | 4.589 (4.449; 4.729) | 0.174 (0.161; 0.188) |
| SD, mmol/L | 1-hour, mmol/L | 0.222 (0.157; 0.287) | 0.076 (0.070; 0.083) |
| | MGP, mmol/L | 0.075 (0.001; 0.149) | 0.083 (0.076; 0.089) |
| | IGP, mmol/L | 0.436 (0.387; 0.485) | 0.105 (0.097; 0.114) |
| | 45-minute, mmol/L | 0.108 (0.029; 0.187) | 0.087 (0.080; 0.095) |
| | 90-minute, mmol/L | 0.326 (0.271; 0.381) | 0.071 (0.065; 0.760) |
| CV, % | 1-hour, mmol/L | 8.136 (7.315; 8.958) | 0.706 (0.627; 0.785) |
| | IGP, mmol/L | 9.956 (9.329; 10.583) | 1.002 (0.891; 1.114) |
| | MGP, mmol/L | 6.918 (5.948; 7.888) | 0.750 (0.663; 0.836) |
| | 90-minute, mmol/L | 9.228 (8.516; 9.941) | 0.640 (0.567; 0.712) |
| | 45-minute, mmol/L | 7.097 (6.109; 8.084) | 0.807 (0.710; 0.903) |
| Normal glucose metabolism | | | |
| MSG, mmol/L | FPG, mmol/L | 3.840 (3.402; 4.278) | 0.389 (0.304; 0.474) |
| | MGP, mmol/L | 5.184 (4.983; 5.384) | 0.078 (0.054; 0.101) |
| | 45-minute, mmol/L | 5.345 (5.186; 5.503) | 0.063 (0.043; 0.082) |
| SD, mmol/L | IGP, mmol/L | 0.608 (0.556; 0.661) | 0.049 (0.035; 0.063) |
| | MGP, mmol/L | 0.405 (0.292; 0.519) | 0.043 (0.030; 0.560) |
| | 1-hour, mmol/L | 0.530 (0.451; 0.608) | 0.033 (0.023; 0.044) |
| CV, % | IGP, mmol/L | 10.845 (10.015; 11.675) | 0.699 (0.475; 0.924) |
| | MGP, mmol/L | 8.392 (6.584; 10.201) | 0.566 (0.357; 0.775) |
| | 1-hour, mmol/L | 10.030 (8.787; 11.274) | 0.435 (0.271; 0.600) |
| Prediabetes | | | |
| MSG, mmol/L | HbA$_{1c}$, mmol/mol | 3.120 (2.231; 4.010) | 0.084 (0.061; 0.108) |
| | FPG, mmol/L | 3.449 (2.398; 4.501) | 0.487 (0.309; 0.665) |
| | 30-minute, mmol/L | 4.450 (3.733; 5.166) | 0.191 (0.118; 0.264) |
| SD, mmol/L | 30-minute, mmol/L | 0.579 (0.226; 0.933) | 0.041 (0.005; 0.077) |
| | HbA$_{1c}$, mmol/mol | -0.039 (-0.483; 0.405) | 0.027 (0.015; 0.039) |
| | 1-hour, mmol/L | 0.443 (0.167; 0.720) | 0.048 (0.024; 0.073) |
| CV, % | 2-hour, mmol/L | 11.017 (7.554; 14.481) | 0.528 (0.113; 0.943) |
| | 90-minute, mmol/L | 10.749 (7.529; 13.970) | 0.469 (0.151; 0.786) |
| | IGP, mmol/L | 12.662 (10.273; 15.050) | 0.489 (0.082; 0.897) |
| Type 2 diabetes | | | |
| MSG, mmol/L | HbA$_{1c}$, mmol/mol | 3.440 (2.565; 4.315) | 0.090 (0.072; 0.107) |
| | MGP, mmol/L | 3.965 (2.958; 4.972) | 0.218 (0.156; 0.280) |
| | 90-minute, mmol/L | 4.686 (3.826; 5.545) | 0.181 (0.127; 0.236) |
| SD, mmol/L | IGP, mmol/L | 0.353 (0.047; 0.659) | 0.126 (0.092; 0.160) |
| | MGP, mmol/L | -0.062 (-0.489; 0.365) | 0.095 (0.069; 0.121) |
| | 90-minute, mmol/L | 0.179 (-0.180; 0.538) | 0.084 (0.061; 0.106) |
| CV, % | IGP, mmol/L | 9.095 (5.561; 12.630) | 1.169 (0.776; 1.562) |
| | 90-minute, mmol/L | 8.590 (4.498; 12.681) | 0.700 (0.440; 0.960) |
| | MGP, mmol/L | 7.166 (2.110; 12.222) | 0.760 (0.448; 1.072) |

This table shows the constants and unadjusted regression coefficients (95% confidence intervals) of the associations of mean sensor glucose (MSG), standard deviation (SD), and coefficient of variation (CV) with established glycaemic indices (i.e. FPG, fasting plasma glucose; IGP, incremental glucose peak; MGP, maximum glucose peak; several other oral glucose tolerance test time points).

Table S2.10  Distribution of CGM-derived indices after exclusion of individuals who underwent continuous glucose monitoring as part of a catch-up visit.

| CGM-derived indices | | NGM (n=423) | PreD (n=108) | T2D (n=72) | p-value |
|---|---|---|---|---|---|
| MSG, mmol/L | Median [IQR] | 5.8 [5.5 − 6.1] | 6.2 [5.7 − 6.6] | 7.3 [6.5 − 8.2] | <0.001 |
| | Mean ± SD | 5.8 ± 0.4 | 6.2 ± 0.6 | 7.6 ± 1.6 | |
| | $10^{th} − 90^{th}$ percentile | 5.32 - 6.37 | 5.46 - 7.00 | 5.82 - 9.15 | |
| | $5^{th} − 95^{th}$ percentile | 5.14 - 6.57 | 5.20 - 7.19 | 5.56 - 12.16 | |
| | $2.5^{th} − 97.5^{th}$ percentile | 5.01 - 6.71 | 5.04 - 7.41 | 5.32 - 12.52 | |
| SD, mmol/L | Median [IQR] | 0.73 [0.63 − 0.88] | 0.89 [0.73 − 1.11] | 1.46 [0.97 − 1.95] | <0.001 |
| | Mean ± SD | 0.78 ± 0.23 | 0.94 ± 0.27 | 1.54 ± 0.70 | |
| | $10^{th} − 90^{th}$ percentile | 0.53 - 1.12 | 0.62 - 1.24 | 0.78 - 2.47 | |
| | $5^{th} − 95^{th}$ percentile | 0.48 - 1.28 | 0.60 - 1.47 | 0.63 - 3.01 | |
| | $2.5^{th} − 97.5^{th}$ percentile | 0.44 - 1.38 | 0.54 - 1.65 | 0.53 - 3.23 | |
| CV, % | Median [IQR] | 12.7 [10.9 − 15.1] | 14.8 [12.2 − 17.5] | 19.0 [14.5 − 24.1] | <0.001 |
| | Mean ± SD | 13.3 ± 3.7 | 15.2 ± 3.9 | 20.0 ± 7.5 | |
| | $10^{th} − 90^{th}$ percentile | 9.40 - 18.23 | 10.66 - 20.54 | 11.47 - 29.63 | |
| | $5^{th} − 95^{th}$ percentile | 8.35 - 20.93 | 10.02 - 23.64 | 10.40 - 37.08 | |
| | $2.5^{th} − 97.5^{th}$ percentile | 7.72 - 22.46 | 9.29 - 24.46 | 9.95 - 39.62 | |

CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; IQR, interquartile range.

Table S2.11  Distribution of CGM-derived indices after exclusion of individuals with less than three recording days.

| CGM-derived indices | | NGM (n=454) | PreD (n=174) | T2D (n=190) | p-value |
|---|---|---|---|---|---|
| MSG, mmol/L | Median [IQR] | 5.8 [5.5 − 6.1] | 6.2 [5.8 − 6.6] | 7.6 [6.8 − 8.7] | <0.001 |
| | Mean ± SD | 5.8 ± 0.4 | 6.2 ± 0.8 | 7.9 ± 1.6 | |
| | $10^{th} − 90^{th}$ percentile | 5.30 - 6.36 | 5.56 - 7.07 | 6.11 - 9.87 | |
| | $5^{th} − 95^{th}$ percentile | 5.15 - 6.52 | 5.24 - 7.60 | 5.90 - 11.78 | |
| | $2.5^{th} − 97.5^{th}$ percentile | 5.02 - 6.70 | 5.03 - 8.33 | 5.63 - 12.78 | |
| SD, mmol/L | Median [IQR] | 0.72 [0.62 − 0.86] | 0.89 [0.73 − 1.12] | 1.52 [1.15 − 1.94] | <0.001 |
| | Mean ± SD | 0.77 ± 0.23 | 0.97 ± 0.54 | 1.64 ± 0.70 | |
| | $10^{th} − 90^{th}$ percentile | 0.53 - 1.08 | 0.64 - 1.30 | 0.88 - 2.46 | |
| | $5^{th} − 95^{th}$ percentile | 0.48 - 1.23 | 0.60 - 1.62 | 0.78 - 3.10 | |
| | $2.5^{th} − 97.5^{th}$ percentile | 0.44 - 1.36 | 0.54 - 1.93 | 0.66 - 3.45 | |
| CV, % | Median [IQR] | 12.6 [10.8 − 14.8] | 14.7 [12.2 − 17.5] | 19.3 [15.9 − 24.0] | <0.001 |
| | Mean ± SD | 13.2 ± 3.6 | 15.3 ± 4.4 | 20.4 ± 6.5 | |
| | $10^{th} − 90^{th}$ percentile | 9.40 - 18.00 | 10.70 - 20.63 | 13.14 - 28.62 | |
| | $5^{th} − 95^{th}$ percentile | 8.39 - 19.95 | 10.13 - 24.06 | 12.21 - 32.59 | |
| | $2.5^{th} − 97.5^{th}$ percentile | 7.70 - 22.43 | 9.29 - 25.94 | 10.48 - 36.25 | |

CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; IQR, interquartile range.

**Table S2.12** Distribution of CGM-derived indices after exclusion of individuals with data gaps during the continuous glucose monitoring recording.

| CGM-derived indices | | NGM (n=434) | PreD (n=168) | T2D (n=183) | *P*-value |
|---|---|---|---|---|---|
| MSG, mmol/L | Median [IQR] | 5.8 [5.5 − 6.1] | 6.2 [5.8 − 6.6] | 7.5 [6.8 − 8.6] | <0.001 |
| | Mean ± SD | 5.8 ± 0.4 | 6.3 ± 0.7 | 7.8 ± 1.7 | |
| | $10^{th} - 90^{th}$ percentile | 5.31 - 6.35 | 5.56 - 7.06 | 6.07 - 9.90 | |
| | $5^{th} - 95^{th}$ percentile | 5.15 - 6.58 | 5.31 - 7.51 | 5.73 - 11.87 | |
| | $2.5^{th} - 97.5^{th}$ percentile | 5.03 - 6.71 | 5.02 - 8.16 | 5.63 - 12.78 | |
| SD, mmol/L | Median [IQR] | 0.72 [0.62 − 0.86] | 0.90 [0.75 − 1.13] | 1.51 [1.12 − 1.93] | <0.001 |
| | Mean ± SD | 0.77 ± 0.23 | 0.97 ± 0.32 | 1.60 ± 0.67 | |
| | $10^{th} - 90^{th}$ percentile | 0.53 - 1.07 | 0.64 - 1.30 | 0.87 - 2.43 | |
| | $5^{th} - 95^{th}$ percentile | 0.48 - 1.23 | 0.60 - 1.58 | 0.77 - 3.04 | |
| | $2.5^{th} - 97.5^{th}$ percentile | 0.45 - 1.37 | 0.54 - 1.86 | 0.66 - 3.42 | |
| CV, % | Median [IQR] | 12.6 [10.8 − 14.7] | 15.0 [12.2 − 17.5] | 19.1 [15.8 − 23.9] | <0.001 |
| | Mean ± SD | 13.1 ± 3.5 | 15.2 ± 3.8 | 20.1 ± 6.1 | |
| | $10^{th} - 90^{th}$ percentile | 9.44 - 17.88 | 10.72 - 20.51 | 13.09 - 27.53 | |
| | $5^{th} - 95^{th}$ percentile | 8.58 - 19.91 | 10.12 - 23.70 | 12.10 - 32.44 | |
| | $2.5^{th} - 97.5^{th}$ percentile | 8.02 - 22.43 | 9.28 - 24.54 | 10.45 - 34.65 | |

CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; MSG, mean sensor glucose; SD, standard deviation; CV, coefficient of variation; IQR, interquartile range.

**Table S2.13** Distribution of MAGE, IQR and MODD.

| CGM-derived indices | | NGM (n=470) | PreD (n=184) | T2D (n=197) | *P*-value |
|---|---|---|---|---|---|
| MAGE, mmol/L | Median [IQR] | 1.47 [1.19 − 1.86] | 2.01 [1.54 − 2.47] | 3.43 [2.47 − 4.59] | <0.001 |
| | Mean ± SD | 1.60 ± 0.60 | 2.10 ± 0.84 | 3.64 ± 1.46 | |
| | $10^{th} - 90^{th}$ percentile | 1.01 - 2.36 | 1.23 - 3.06 | 1.94 - 5.73 | |
| | $5^{th} - 95^{th}$ percentile | 0.90 - 2.70 | 1.11 - 3.80 | 1.58 - 6.12 | |
| | $2.5^{th} - 97.5^{th}$ percentile | 0.83 - 3.00 | 1.01 - 4.29 | 1.22 - 6.47 | |
| IQR, mmol/L | Median [IQR] | 0.90 [0.70 − 1.10] | 1.10 [0.90 − 1.40] | 1.85 [1.38 − 2.60] | <0.001 |
| | Mean ± SD | 0.94 ± 0.30 | 1.20 ± 0.46 | 2.07 ± 1.00 | |
| | $10^{th} - 90^{th}$ percentile | 0.60 - 1.30 | 0.80 - 1.70 | 1.10 - 3.09 | |
| | $5^{th} - 95^{th}$ percentile | 0.60 - 1.57 | 0.70 - 2.08 | 0.80 - 3.80 | |
| | $2.5^{th} - 97.5^{th}$ percentile | 0.50 - 1.70 | 0.60 - 2.62 | 0.72 - 5.10 | |
| MODD, mmol/L | Median [IQR] | 0.69 [0.58 − 0.82] | 0.80 [0.68 − 1.01] | 1.36 [1.05 − 1.72] | <0.001 |
| | Mean ± SD | 0.72 ± 0.20 | 0.88 ± 0.30 | 1.47 ± 0.65 | |
| | $10^{th} - 90^{th}$ percentile | 0.50 - 0.98 | 0.56 - 1.27 | 0.75 - 2.25 | |
| | $5^{th} - 95^{th}$ percentile | 0.45 - 1.07 | 0.53 - 1.43 | 0.66 - 2.72 | |
| | $2.5^{th} - 97.5^{th}$ percentile | 0.41 - 1.21 | 0.48 - 1.70 | 0.62 - 3.47 | |

CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; MAGE, mean amplitude of glucose excursions; IQR, interquartile range; MODD, mean of daily differences.

# Chapter 3

Machine learning-based glucose prediction with use of continuous glucose and physical activity monitoring data – The Maastricht Study

William P.T.M. van Doorn*, Yuri D. Foreman*, Nicolaas C. Schaper,
Hans H.C.M. Savelberg, Annemarie Koster, Carla J.H. van der Kallen,
Anke Wesselius, Miranda T. Schram, Ronald M.A. Henry, Pieter C. Dagnelie,
Bastiaan E. de Galan, Otto Bekers, Coen D.A. Stehouwer,
Steven J.R. Meex†, and Martijn C.G.J. Brouwers†

* First authors contributed equally; † last authors contributed equally

# Abstract

**Background**

Closed-loop insulin delivery systems, which integrate continuous glucose monitoring (CGM) and algorithms that continuously guide insulin dosing, have been shown to improve glycemic control. The ability to predict future glucose values can further optimize such devices. In this study, we used machine learning to train models in predicting future glucose levels based on prior CGM and accelerometry data.

**Methods**

We used data from The Maastricht Study, an observational population-based cohort that comprises individuals with normal glucose metabolism, prediabetes, or type 2 diabetes. We included individuals who underwent >48h of CGM (n=851), most of whom (n=540) simultaneously wore an accelerometer to assess physical activity. A random subset of individuals was used to train models in predicting glucose levels at 15- and 60-minute intervals based on either CGM data or both CGM and accelerometer data. In the remaining individuals, model performance was evaluated with root-mean-square error (RMSE), Spearman's correlation coefficient (rho) and surveillance error grid. For a proof-of-concept translation, CGM-based prediction models were optimized and validated with the use of data from individuals with type 1 diabetes (OhioT1DM Dataset, n=6).

**Results**

Models trained with CGM data were able to accurately predict glucose values at 15 (RMSE: 0.19mmol/L; rho: 0.96) and 60 minutes (RMSE: 0.59mmol/L, rho: 0.72). Model performance was comparable in individuals with type 2 diabetes. Incorporation of accelerometer data only slightly improved prediction. The error grid results indicated that model predictions were clinically safe (15 min: >99%, 60 min >98%). Our prediction models translated well to individuals with type 1 diabetes, which is reflected by high accuracy (RMSEs for 15 and 60 minutes of 0.43 and 1.73 mmol/L, respectively) and clinical safety (15 min: >99%, 60 min: >91%).

**Conclusions**

Machine learning-based models are able to accurately and safely predict glucose values at 15- and 60-minute intervals based on CGM data only. Future research should further optimize the models for implementation in closed-loop insulin delivery systems.

## Introduction

The increasing prevalence of diabetes entails an increase in debilitating complications, such as retinopathy, neuropathy, and cardiovascular disease[1-3]. Maintaining plasma glucose levels within the reference range is essential for the prevention of diabetes-related complications, which are generally attributable to chronic hyperglycemia, although hypoglycemia has been suggested to contribute to cardiovascular disease risk as well[3-5]. One of the most promising developments to minimize hyperglycemia and hypoglycemia—and, hence, to increase time in range—in individuals with diabetes who require insulin treatment is a closed-loop insulin delivery system (also known as the artificial pancreas). Such a system integrates continuous glucose monitoring (CGM), insulin (with or without glucagon) infusion, and a control algorithm to continuously regulate blood glucose levels[6,7]. Multiple studies have shown the merit of incorporating the artificial pancreas into clinical care of individuals with type 1 or type 2 diabetes[8,9].

Despite prior efforts, there are still numerous points that need to be addressed in order to improve the individual components of closed-loop systems[6,10]. With regard to CGM, this includes overcoming sensor delay (i.e., the inherent ~10-minute discrepancy between interstitially measured and actual plasma glucose values), and sensor malfunctions (i.e., periods during which no glucose values are recorded)[6,10,11]. Continuous glucose prediction is a potentially viable strategy to both handle sensor delay and bridge periods of sensor malfunction. The use of machine learning has yielded encouraging glucose prediction accuracy results in relatively small study populations (mostly individuals with type 1 diabetes) or *in silico* studies, as extensively reviewed elsewhere[12]. Large, human-based study populations are now needed to reliably assess to what extent and within what time interval (i.e., prediction horizon) glucose values can be accurately predicted by use of machine learning. Additionally, incorporation of physical activity, which is considered an important factor for glucose control in daily life, could further improve glucose prediction[6].

In this study, we investigated to what extent glucose values can be accurately predicted at intervals of 15 and 60 minutes by a machine learning model that has been trained with a sliding time window of glucose values preceding the predicted values at a fixed interval. Additionally, we studied whether glucose prediction can be further improved by incorporation of accelerometer-measured physical activity, and to what extent the results differ in a subgroup analysis of individuals with type 2 diabetes only. For this, we used a large population of individuals with either normal glucose metabolism (NGM), prediabetes, or type 2 diabetes who simultaneously underwent CGM and continuous accelerometry during a one-week period. Last, we used the publicly available OhioT1DM Dataset to explore whether CGM-based prediction models would translate to individuals with type 1 diabetes, the primary target population for closed-loop insulin delivery.

## Methods

### Study population and design

We used data from The Maastricht Study, an observational, prospective, population-based cohort study. The rationale and methodology have been described previously[13]. In brief, The Maastricht Study focuses on the etiology, pathophysiology, complications and comorbidities of type 2 diabetes, and is characterized by an extensive phenotyping approach. All individuals aged between 40 and 75 years and living in the southern part of the Netherlands were eligible for participation. Participants were recruited through mass media campaigns and from the municipal registries and the regional Diabetes Patient Registry via mailings. For reasons of efficiency, recruitment was stratified according to known type 2 diabetes status, with an oversampling of individuals with type 2 diabetes. In general, the examinations of each participant were performed within a time window of three months. From 19 September 2016 until 13 September 2018, participants were invited to also undergo CGM[14]. During this period, a selected group of recently included participants were invited to return for CGM. In these participants only, there was a median time interval of 2.1 years between CGM and all other measurements. The present report includes cross-sectional data of the 851 participants who had at least 48h of CGM data available and were classified with NGM, prediabetes, or type 2 diabetes. The Maastricht Study has been approved by the institutional medical ethical committee (NL31329.068.10) and the Minister of Health, Welfare and Sports of the Netherlands (Permit 131088-105234-PG). All participants gave written informed consent.

### Continuous glucose monitoring

The rationale and methodology of CGM (iPro2 and Enlite Glucose Sensor; Medtronic, Tolochenaz, Switzerland) have been described previously[14]. In brief, the CGM device was worn abdominally and recorded subcutaneous interstitial glucose values (range: 2.2 - 22.2 mmol/L) every five minutes for a seven-day period. For calibration purposes, participants were asked to perform self-measurements of blood glucose four times daily (Contour Next; Ascensia Diabetes Care, Mijdrecht, the Netherlands). Participants were blinded to the CGM recording, but not to self-measured values. Diabetes medication use was allowed and no dietary instructions were given. We only included individuals with at least 48h of CGM, but excluded the first 24h of CGM from analysis because of insufficient calibration. For the glucose prediction analyses, all remaining glucose data points were used. We additionally calculated mean sensor glucose, standard deviation (SD), and coefficient of variation (CV) with the use of Glycemic Variability Research Tool (GlyVaRT; Medtronic) software.

## *Accelerometry*

As described previously, daily physical activity was measured with use of the triaxial activPAL3 accelerometer (PAL technologies; Glasgow, United Kingdom)[13,15]. The accelerometer was, just as the CGM device, attached during the first research visit; participants wore the accelerometer on the front of the right thigh for eight consecutive days. No physical activity instructions were given. PAL Software Suite version 8 (PAL technologies) was used to convert the event-based accelerometry data files into 15-second interval data files. We used the composite of X, Y, and Z accelerations for each 15-second interval as the measure of physical activity.

## *Assessment of participant characteristics*

As described previously[13], we classified glucose metabolism status (GMS) as either NGM, prediabetes, or type 2 diabetes based on both a standardized 2-hour 75 gram oral glucose tolerance test and use of glucose-lowering medication[16]. We assessed medication use as part of a medication interview. Additionally, we determined smoking status and history of diabetes based on questionnaires, measured weight and height—to calculate body mass index (BMI)—and office blood pressure during a physical examination, and measured $HbA_{1c}$ as well as lipid profile in fasting venous blood.

## *Dataset construction*

An overview of data preprocessing, model development, and model evaluation is given in **Figure 3.1**. In order to train our models in predicting future glucose values, we constructed two separate datasets (**Figure 3.1**, panel a). The first dataset consisted of only the participants' six-day, five-minute interval CGM data (n=851). The second dataset consisted of both CGM and accelerometry data (n=540). To synchronize CGM (determined at 5-minute intervals) and accelerometry data (determined at 15-second intervals) in the second dataset, we linearly interpolated glucose values between two glucose data points with a frequency of 15 seconds. Consistent and aligned frequency intervals across these parameters are a statistical precondition for this type of model development[17]. The study populations were randomly split into a training (70%), tuning (10%), and evaluation (20%) dataset such that data from a given individual were present only in one set. The training set was used to train the proposed models. The tuning set was used to iteratively improve the models by selecting the best model architectures and hyperparameters. Finally, the best models were evaluated on the independent evaluation set that was retained during model development.
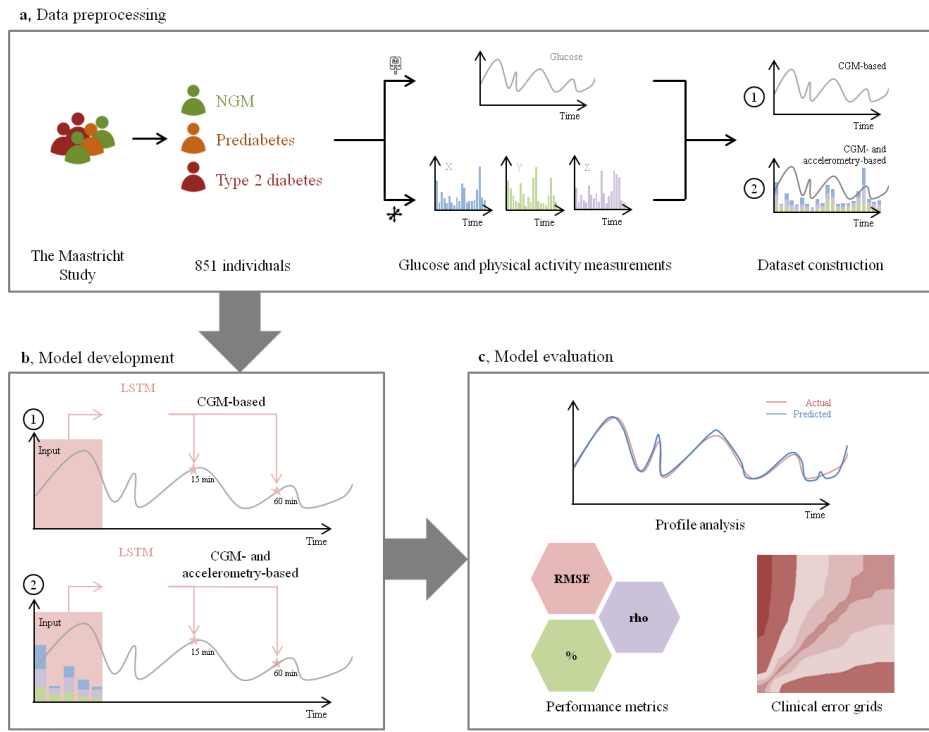
**Figure 3.1    Overview of data preprocessing, model development and evaluation.** Data was used from The Maastricht Study, an observational population-based cohort that comprises individuals with normal glucose metabolism (NGM), prediabetes, or type 2 diabetes (panel A). We included 851 individuals who underwent continuous glucose monitoring (CGM), most of whom simultaneously wore an accelerometer to assess physical activity (X, Y, and Z accelerations). Models developed with the long-short term memory (LSTM) architecture were trained in predicting glucose levels at 15- and 60-minute intervals with either CGM data only (1) or both CGM and accelerometer data (2) (panel B). Finally, model performance was evaluated by glucose profile analysis, performance metrics (root-mean-square error [RMSE]; Spearman's correlation coefficient [rho]; proportions), and clinical error grids (panel C).

## Model development and design

Our proposed predictive model operates sequentially over CGM and accelerometry data (**Figure 3.1**, panel b). At each individual time point, 30 minutes of prior time series data were provided to the statistical model (e.g., six CGM-based glucose values), based on which it predicted glucose values at specified time intervals. For this study, we set these time intervals at 15 and 60 minutes. The nature of this prediction task can be solved by a variety of statistical and machine learning models. In the current study, we assessed autoregressive integrated moving average, support vector regression,

gradient-boosting systems, shallow and deep multi-layer perceptron neural networks, and several recurrent neural network (RNN) architectures, including classical RNN[18,19], gated recurrent units[20], long-short term memory (LSTM) networks[21], and all of its bi-directional variants[22,23]. We chose to use a LSTM architecture, as this had the best performance in the tuning dataset. This architecture runs sequentially over time series data and is able to implicitly model the historical context of an individual by modifying an internal state through time. Specifically, we designed this architecture to predict both time intervals simultaneously, often referred to as "multi-task learning", which aims to share knowledge amongst prediction tasks.

## Model selection and training

We selected the multi-task LSTM network among several alternatives based on the tuning set performance (**Supporting Information S3.1** and **Table S3.1**). Next, we evaluated a broad spectrum of hyperparameter combinations for this network (**Table S3.2**). This resulted in a multi-task LSTM architecture, consisting of three layers, including a dropout layer with a total of 56-104 neurons (**Table S3.3**). During training, we used exponential learning-rate decay via the Adam optimization scheme[24]. The best validation results were achieved by use of an initial learning rate with a decay of 0.001 every 1,000 training steps, with a batch size of 1024, and a back-propagation through a time window of 30 minutes. This defines the amount of historic data the model uses, which in our case translates to six (first dataset) or 120 (second dataset) glucose data points, for the model to provide a prediction. The loss function during training was the mean average of the mean-squared error function of all predictions. The maximum amount of epochs was 50.000 with an early stopping criterion (based on 20% hold-out data) set to 250 epochs. We performed data preprocessing, model development, selection, and training using Python programming language (version 3.7.1) with the use of packages Numpy (version 1.17), Pandas (version 0.24), Keras (version 2.2.2), Scikit-learn (version 0.22.0) and Tensorflow (version 2.0.1, beta).

## Translation of the prediction models to the OhioT1DM Dataset

We used data from the OhioT1DM Dataset to explore whether our CGM-based prediction models would translate to individuals with type 1 diabetes. The OhioT1DM Dataset is freely available for scientific purposes and contains data of 6 individuals with type 1 diabetes who were all using insulin pump therapy and CGM[25]. The participants provided interstitial glucose values every five minutes for an eight-week period. First, in order to also include 30-minute prediction, we retrained our main CGM-based models on the main study population with identical hyperparameters and settings (**Table S3.3**). Then, we evaluated the main CGM-based model on the test portion of the OhioT1DM Dataset (20%). Next, we aimed to optimize our main CGM-based model by training it on the train portion of the OhioT1DM Dataset. Specifically, we trained the model using an

Adam optimizer with a learning rate of $10^{-4}$, a batch size of 1024, a maximum of 10.000 epochs and an early stopping criterion (based on 20% of the training data) set to 100 epochs. Last, we evaluated this optimized model on the test portion using performance metrics and safety error grids, as described previously.

### Model evaluation and statistical analysis

Model evaluation was performed in the independent evaluation sets of individuals that were not used during model development (**Figure 3.1**, panel c). We employed several metrics to assess the performance of our models: root-mean-square error (RMSE), proportion of predicted values within 5% or 10% of actual glucose values, and Spearman's rank correlation coefficient (rho) (**Supporting Information S3.2**). Bootstrapping was performed to obtain 95% confidence intervals for each of these metrics[26]. In addition, we used error grids that are classically used for assessment of blood glucose monitor safety (i.e., surveillance error grid, Parkes error grid) to evaluate the safety of our glucose prediction models[27,28]. Last, we performed several sensitivity analysis in our main study population by stratifying model performance for: (1) GMS (i.e., separate results for NGM and prediabetes); (2) day (06.00 to 24.00h) and night (24.00 to 06.00h); and (3) low or high glucose variability, defined as the 97.5th percentile of CGM-assessed SD in individuals with NGM (SD > 1.37 mmol/L)[14].

Normally distributed data are presented as mean ± SD, non-normally distributed data as median and interquartile range, and categorical data as n (%). Statistical analyses were performed using the Statistical Package for Social Sciences (version 25.0; IBM, Chicago, Illinois, USA) and the Python programming language (version 3.7.1).

## Results

### Main study population characteristics

In total, 896 individuals underwent CGM as part of The Maastricht Study's extensive phenotyping approach. We included participants with at least 48h of CGM data and either NGM, prediabetes, or type 2 diabetes. This resulted in the final study population of 851 individuals. Of this population, 540 participants (63.5%) simultaneously underwent CGM and accelerometry.

**Table 3.1** shows the overall and type 2 diabetes-stratified characteristics of the two study populations (CGM-based as well as CGM- and accelerometry-based glucose prediction). The overall participant characteristics of both populations were generally comparable with regard to age, sex, BMI, glycemic indices, blood pressure, and lipid profile, although the latter contained fewer participants with prediabetes or type 2 diabetes. Additionally, the participants with type 2 diabetes in the CGM- and accelerometry-based glucose prediction population were more often newly diagnosed with type 2 diabetes. Accordingly, these participants less often used glucose-lowering

medication. Participant characteristics of the NGM and prediabetes subgroups are described in **Table S3.4**.

Table 3.1 Participant characteristics of the CGM-based and CGM- and accelerometry-based glucose prediction study populations.

| Characteristic | CGM-based glucose prediction | | CGM- and accelerometry-based glucose prediction | |
|---|---|---|---|---|
| | Total (n=851) | T2D (n=197) | Total (n=540) | T2D (n=68) |
| Age, years | 59.9 ± 8.7 | 62.4 ± 7.8 | 59.1 ± 8.7 | 62.0 ± 6.9 |
| Women, n (%) | 418 (49.1) | 69 (35.0) | 276 (51.1) | 22 (32.4) |
| BMI, kg/m$^2$ | 27.2 ± 4.4 | 29.7 ± 4.7 | 26.5 ± 4.0 | 28.6 ± 4.1 |
| Newly diagnosed T2D, n (%) | 70 (8.2) | 70 (35.5) | 35 (6.5) | 35 (51.5) |
| Glucose metabolism status | | | | |
| NGM/PreD/T2D, n | 470/184/197 | - | 372/99/68 | - |
| NGM/PreD/T2D, % | 55.2/21.6/23.1 | - | 69.1/18.3/12.6 | - |
| Fasting plasma glucose, mmol/L | 5.4 [5.0–6.2] | 7.3 [6.5–8.4] | 5.3 [4.9–5.8] | 7.2 [6.3–8.4] |
| 2-h post-load glucose, mmol/L | 6.7 [5.2–9.1] | 13.6 [11.7–16.2] | 6.2 [5.0–7.7] | 12.5 [11.3–16.6] |
| HbA$_{1c}$, % | 5.7 ± 0.8 | 6.7 ± 1.0 | 5.6 ± 0.6 | 6.4 ± 0.9 |
| HbA$_{1c}$, mmol/mol | 39.1 ± 8.3 | 49.2 ± 10.8 | 37.3 ± 6.2 | 46.9 ± 10.2 |
| Sensor glucose | | | | |
| Mean, mmol/L | 6.1 [5.7–6.7] | 7.5 [6.8–8.7] | 5.9 [5.6–6.4] | 7.3 [6.5–8.2] |
| SD, mmol/L | 0.84 [0.68–1.18] | 1.51 [1.14 – 1.95] | 0.79 [0.66–1.01] | 1.46 [0.94–1.99] |
| SD >1.37 mmol/L, n (%) | 142 (16.7) | 115 (58.4) | 50 (9.3) | 36 (52.9) |
| CV, % | 14.0 [11.6–17.6] | 19.3 [15.9–24.0] | 13.3 [11.2–16.8] | 19.2 [14.5–24.1] |
| Diabetes medication use, n (%) | 109 (12.8) | 109 (55.6) | 27 (4.8) | 27 (39.7) |
| Insulin | 19 (2.2) | 19 (9.6) | 4 (0.7) | 4 (5.9) |
| Metformin | 104 (12.2) | 104 (53.1) | 27 (5.0) | 27 (39.7) |
| Sulfonylureas | 21 (2.5) | 21 (10.7) | 6 (1.1) | 6 (8.8) |
| Thiazolidinediones | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| GLP-1 analogues | 3 (0.4) | 3 (1.5) | 1 (0.2) | 1 (1.5) |
| DDP-4 inhibitors | 1 (0.1) | 1 (0.5) | 0 (0) | 0 (0) |
| SGLT-2 inhibitors | 1 (0.1) | 1 (0.5) | 0 (0) | 0 (0) |
| Office SBP, mmHg | 133.3 ± 18.0 | 139.4 ± 15.6 | 132.2 ± 17.9 | 137.7 ± 15.3 |
| Office DBP, mmHg | 75.2 ± 10.2 | 77.7 ± 10.5 | 74.7 ± 10.1 | 77.7 ± 9.6 |
| Antihypertensive medication use, n (%) | 305 (35.9) | 126 (64.3) | 162 (30.0) | 41 (60.3) |
| Total-to-HDL cholesterol ratio | 3.5 [2.8–4.3] | 3.6 [2.9–4.3] | 3.4 [2.8–4.3] | 3.7 [2.8–4.6] |
| Triglycerides, mmol/L | 1.3 [0.9–1.8] | 1.5 [1.0–2.1] | 1.2 [0.9–1.7] | 1.6 [1.0–2.3] |
| Lipid-modifying medication use, n (%) | 212 (24.9) | 115 (58.4) | 100 (18.5) | 39 (57.4) |
| Smoking status | | | | |
| Never/former/current, n | 327/415/106 | 67/104/26 | 214/253/70 | 19/36/13 |
| Never/former/current, % | 38.6/48.9/12.5 | 34.0/52.8/13.2 | 39.9/47.1/13.0 | 27.9/52.9/19.1 |

Data are reported as mean ± SD, median [interquartile range], or number (percentage [%]) as appropriate. CGM, continuous glucose monitoring; BMI, body mass index; T2D, type 2 diabetes; NGM, normal glucose metabolism; PreD, prediabetes; HbA$_{1c}$, glycated hemoglobin A1c; SD, standard deviation; CV, coefficient of variation; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4; SGLT-2, sodium-glucose cotransporter 2; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density lipoprotein.

## *Overall performance of machine learning-based glucose prediction*

We trained two machine learning models (i.e., CGM-based; CGM- and accelerometry-based) in predicting glucose levels at 15- and 60-minute intervals. Visually, both models appeared capable of accurately predicting the real glucose profiles, as illustrated by the representative examples in **Figure S3.1** and **Figure S3.2**. Next, we assessed the performance of our models in our evaluation datasets with a variety of metrics, including an average error term (RMSE), the proportion of predictions within 5% or 10% deviation of the actual value, and correlation (rho). The evaluation datasets comprise 20% of the original or stratified study populations and thus vary in sample size (n=13 - 170).

Overall, our models demonstrated high prediction accuracy, supported by low RMSE values and high proportions of predicted glucose values within 5% and 10% deviation (**Table 3.2**). Model performance in the type 2 diabetes subgroup was generally lower compared to the overall group, except for correlation coefficients, which were often higher in individuals with type 2 diabetes. This phenomenon can be largely attributed to the lower correlation coefficients of individuals with NGM and prediabetes (**Table S3.5**), which is caused by range restriction (i.e., smaller glucose ranges attenuate the correlation coefficients)[29]. Consequently, the correlation coefficients are valid for the comparison of CGM-based glucose prediction to CGM- and accelerometry-based glucose prediction, but not for comparison of the overall study population to the type 2 diabetes subgroup. In addition, we observed short-to-moderate time lags for the 15- and 60-minute predictions (**Table S3.6**).

Table 3.2  Overall performance in the main study population of CGM-based and CGM- and accelerometry-based machine learning models trained in predicting glucose values at time intervals of 15 and 60 minutes.

| | CGM-based glucose prediction | | CGM- and accelerometry-based glucose prediction | |
|---|---|---|---|---|
| | Total (n=170) | T2D (n=43) | Total (n=109) | T2D (n=13) |
| **15 minutes** | | | | |
| RMSE, mmol/L | 0.188 [0.186–0.191] | 0.288 [0.281–0.306] | 0.184 [0.177–0.189] | 0.271 [0.260–0.282] |
| <5% , % | 92.98 [92.87–93.05] | 92.02 [91.83–92.25] | 93.06 [93.03–93.09] | 92.04 [91.99–92.11] |
| <10% , % | 99.17 [99.13–99.23] | 98.88 [98.82–98.94] | 99.25 [99.21–99.28] | 98.90 [98.83–98.97] |
| Rho | 0.961 [0.959–0.962] | 0.987 [0.985–0.989] | 0.968 [0.964–0.970] | 0.990 [0.988–0.993] |
| **60 minutes** | | | | |
| RMSE, mmol/L | 0.589 [0.582–0.592] | 0.701 [0.692–0.711] | 0.582 [0.579–0.586] | 0.700 [0.693–0.708] |
| <5% , % | 70.22 [70.09–70.41] | 66.23 [66.13–66.33] | 70.11 [70.05–70.17] | 66.17 [66.09–66.22] |
| <10% , % | 87.39 [87.24–87.53] | 85.82 [85.70–85.93] | 87.44 [87.38–87.50] | 86.11 [86.01–86.20] |
| Rho | 0.721 [0.719–0.722] | 0.781 [0.779–0.782] | 0.725 [0.721–0.729] | 0.790 [0.782–0.799] |

Data are reported as mean [95% confidence interval]. CGM, continuous glucose monitoring; T2D, type 2 diabetes; RMSE, root-mean-square error; < 5%, percentage of predicted values within 5% of actual glucose values; < 10%, percentage of predicted values within 10% of actual glucose values; rho, Spearman's rank correlation coefficient.

In general, incorporation of accelerometry data in the models only slightly improved performance metrics at both prediction intervals (**Table 3.2**). **Table S3.5** shows the model performance in NGM and prediabetes subgroups. Glucose prediction was most precise in individuals with NGM. Of note, the ML-based models substantially outperformed a naive approach that used $t_0$ as predicted glucose value (**Table S3.7**, **Figures S3.3** and **S3.4**).

### Safety evaluation with clinical error grids

We assessed the safety of our machine learning-based glucose prediction using two clinical error grids (i.e., surveillance and Parkes error grids). **Figure 3.2** depicts the safety results for individuals with type 2 diabetes according to the surveillance error grid. At the 15-minute interval, almost all predictions (>99.9%) were clinically safe (i.e., a risk score between 0 and 1.0) (**Figure 3.2**, panels A and B). At the extended prediction window of 60 minutes, clinical safety was slightly lower (98.4-99.2%) (**Figure 3.2**, panels C and D). Parkes error grid assessment yielded similar results (**Figure S3.5**). Of note, less accurate predictions were more often in the vertical B-D zones than in the horizontal B-E zones (e.g., **Figure S3.4**, panel C: 11.80% versus 4.24%), which suggests a model tendency to underestimate rather than overestimate actual glucose values, the latter of which being more dangerous.

### Additional analyses

To further obtain insights into our model predictions, we assessed performance metrics stratified by day and night (**Table S3.8**). Fifteen-minute predictions did not materially differ between day and night. By contrast, accuracy of 60-minute predictions was lower during the day than at night. In addition, we stratified the results by high or low glucose variability (i.e., SD cut-off of 1.37 mmol/L) (**Table S3.9**). Model performance was slightly lower at higher glucose variability, at both time intervals of 15 and 60 minutes.

### Translation of the prediction models to the OhioT1DM Dataset

The prediction accuracy of the CGM-based model that was developed with our main study population was moderate in individuals with type 1 diabetes (RMSEs at 15, 30, and 60 min: 0.689 [0.685 − 0.693], 1.189 [1.183 − 1.195], and 1.918 [1.910 − 1.926] mmol/L), but substantially improved after being trained on data from each individual with type 1 diabetes (RMSEs at 15, 30, and 60 min: 0.426 [0.422 − 0.430], 1.046 [1.039 − 1.052], and 1.733 [1.725 − 1.741] mmol/L; **Table S3.10**). Accordingly, clinical safety was substantial as shown by the high percentages of clinically safe predictions (15-minute: >99%, 30-minute: >97%, and 60-minute: >91%; **Figure 3.3**).
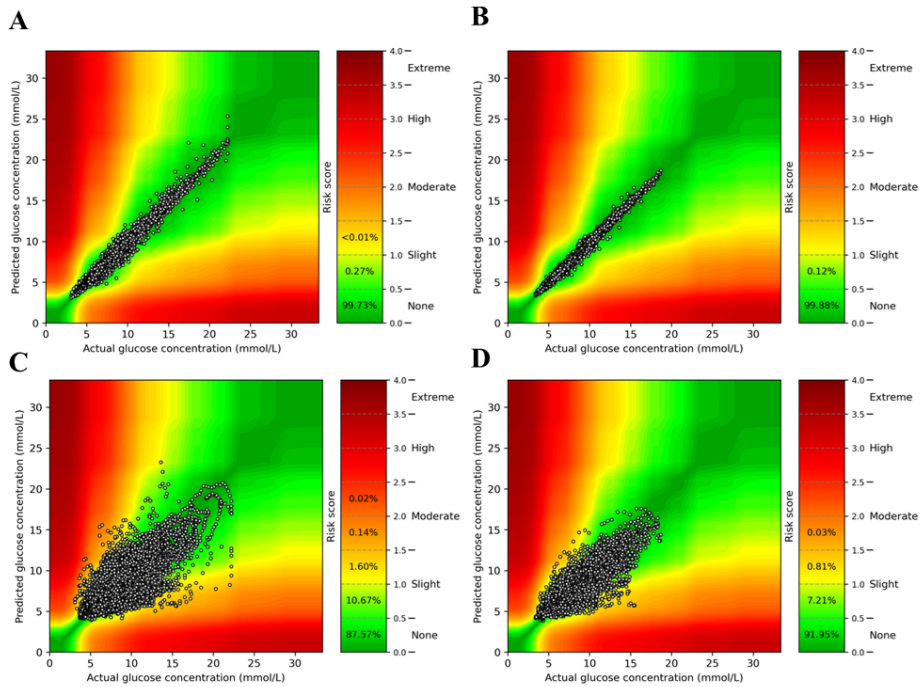
**Figure 3.2** **Surveillance error grid evaluation of glucose prediction safety at time intervals of 15 and 60 minutes in the main study population.** Assessment of CGM-based glucose prediction safety in individuals with type 2 diabetes (n=43) at 15 minutes (panel A) and 60 minutes (panel C). Assessment of CGM- and accelerometry-based glucose prediction safety in individuals with type 2 diabetes (n=13) at 15 minutes (panel B) and 60 minutes (panel D). The risk score values translate to the following degrees of risk: 0 - 0.5, none; 0.5 - 1.0, slight (lower); 1.0 - 1.5, slight (higher); 1.5 - 2.0, moderate (lower); 2.0 - 2.5, moderate (higher); 2.5 – 3.0, great (lower); 3.0 - 3.5, great (higher); >3.5 extreme[27].
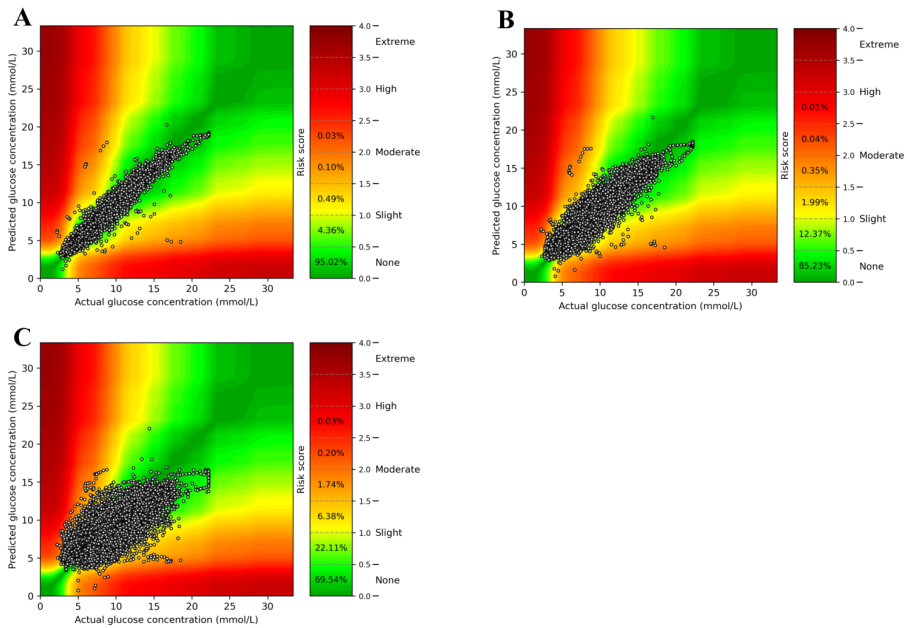
**Figure 3.3 Surveillance error grid evaluation of glucose prediction safety at time intervals of 15, 30, and 60 minutes in individuals with type 1 diabetes.** Assessment of CGM-based glucose prediction safety in individuals with type 1 diabetes (n=6) at 15 (panel A), 30 (panel B), and 60 minutes (panel C). The risk score values translate to the following degrees of risk: 0 - 0.5, none; 0.5 - 1.0, slight (lower); 1.0 - 1.5, slight (higher); 1.5 - 2.0, moderate (lower); 2.0 - 2.5, moderate (higher); 2.5 – 3.0, great (lower); 3.0 - 3.5, great (higher); >3.5 extreme[27].

## Discussion

In this study with 851 individuals and almost 1.4 million glucose measurements, we investigated whether glucose values can be accurately predicted by using machine learning-based models that utilize recently measured CGM and physical activity data with the prospect of improving closed-loop insulin delivery systems. Our study has several important findings and unique characteristics. First, the machine learning-based models are capable of accurately predicting the actual glucose profiles at 15 minutes, as reflected by several objective performance metrics (e.g., RMSE, rho; **Table 3.1**) and visual illustrations (**Figure S3.1** and **Figure S3.2**). Despite prediction accuracy being moderately lower at 60 minutes, more than 98% of the predicted values remained sufficiently accurate to be deemed clinically safe based on surveillance error grids (**Figure 3.2**). Second, glucose prediction only improved slightly when accelerometer-assessed physical activity data was incorporated in the models. Third, translation of our CGM-based glucose prediction models to individuals with

type 1 diabetes yielded encouraging results (i.e., ample prediction accuracy and clinical safety).

Although most research has thus far focused on type 1 diabetes[12], several efforts have been made to use machine learning for glucose prediction in individuals with type 2 diabetes[30-34]. Most of these studies assessed technical aspects of glucose prediction in relatively small (n=1 to 50) or even virtual, *in silico* populations. Such studies provide valuable comparisons of models, but show suboptimal and highly variable performance in predicting glucose values. To our knowledge, this is the first study to report this level of performance in a large, population-based sample of individuals with NGM, prediabetes, or type 2 diabetes. Our CGM-based models were able to accurately predict glucose values at 15 (RMSEs, overall/type 2 diabetes: 0.19/0.29 mmol/L) and 60 minutes (RMSEs, overall/type 2 diabetes: 0.59/0.70 mmol/L). These results surpass previously reported RMSE values for a sample of 50 individuals with type 2 diabetes, which were 0.65 and 1.50 mmol/L for 15- and 60-minute CGM-based glucose prediction, respectively[34]. We expect this difference to, in part, stem from our much larger sample size. To our knowledge, our exploratory translation to individuals with type 1 diabetes (**Table S3.10**) showed that our models perform equally well as recent publications in the field[12,35-38]. For example, the best performing model of the Blood Glucose Level Prediction Challenge 2018, which was also based on a LSTM architecture as well as was trained on and evaluated in the OhioT1DM Dataset, reported 30-minute and 60-minute RMSEs of 1.05 and 1.74 mmol/L[35]. Additionally, Kriventsov et al. recently described large-scale application of glucose prediction in a smartphone app (Diabits) and reported a comparable RMSE at 30 minutes (1.04 mmol/L)[36]. We anticipate that further technical development of our prediction models, while using a larger sample of individuals with type 1 diabetes, will advance performance even more.

We integrated physical activity, which we assessed via accelerometry, into our glucose prediction model, because of its short- and long-term effects on daily glucose patterns. Whereas an acute bout of physical activity can either decrease or increase serum glucose levels, prolonged exercise improves insulin sensitivity, and thus insulin-stimulated glucose uptake[39]. While it should be noted that CGM- and accelerometry-based glucose prediction yielded larger improvements relative to CGM-based glucose prediction for the 60-minute interval, most notably during the day (**Table S3.8**) and in individuals with higher glucose variability (**Table S3.9**), incorporation of physical activity generally only marginally improved glucose prediction. This can be explained by the observation that the models based on CGM data only already performed very well, which limits the ability to achieve additional improvements[40]. Also, the effect of physical activity on serum glucose levels is relatively small in people with beta-cell function that is either normal or only mildly deficient. Given the absence of pancreatic glucoregulation in individuals with type 1 diabetes, it is conceivable that incorporation of accelerometry data leads to more substantially improved model performance in this

patient group[40], which, at present, we were not able to further explore. In addition, a time interval of 15 or 60 minutes could be too short to incorporate long-term physical activity effects into the prediction model.

The closed-loop insulin delivery system has been shown to improve glycemic control in individuals with type 1 or type 2 diabetes[8,9,41]. Nevertheless, several aspects of the artificial pancreas require further enhancement[6,10]. Our results demonstrate that machine learning-based glucose prediction has the promise of being a valid and safe strategy to both overcome ~10-minute sensor delay and bridge prolonged periods of sensor malfunction. Not only are more than 99% of the predicted glucose values in clinically safe zones (i.e., Parkes error grid zone A and B), the model also tended to slightly underestimate rather than overestimate the actual glucose values. In case the prediction model were to be implemented, this would further reduce the risk of iatrogenic hypoglycemia. Nevertheless, future research is needed to assess whether incorporation of these prediction models in a closed-loop insulin delivery system safely improves glycemic control.

This proof-of-principle study has several strengths and limitations. Strengths are 1) the largest well-characterized, population-based study sample thus far, which ensured sufficient statistical power; 2) the unique large-scale combination of CGM and continuous accelerometry, which enabled us to study to what extent incorporation of data on physical activity would improve prediction in this population; 3) result robustness, as reflected by the consistency of several statistical and clinical performance metrics; and 4) the gold-standard assessment of GMS, which allowed for the comparison of performance in NGM, prediabetes and type 2 diabetes.

Our research had certain limitations. First, the main study population comprised individuals with NGM, prediabetes, or type 2 diabetes, who are generally not the target population for closed-loop insulin delivery systems. We, therefore, exploratively investigated whether our prediction models would translate to individuals with type 1 diabetes using the OhioT1DM Dataset, which yielded encouraging results. Nevertheless, we underscore the importance of extensive evaluation of the models in a larger sample of individuals with type 1 diabetes, insulin-treated type 2 diabetes, or both. Second, we were unable to factor in other important elements pertaining to glycemic control (e.g., diet or medication use)[6]. In automated, self-regulatory closed-loop systems, utilization of these kinds of data requires manual input, which is less convenient and reliable than CGM. In addition, since glucose prediction was only slightly improved by incorporating physical activity, we expect relatively little gain from including such factors into our models, at least in individuals with type 2 diabetes. However, given the results of several small studies that have incorporated diet and medication use[12], we acknowledge that this may not hold true for individuals with type 1 diabetes. In this regard, large-scale studies are required to reach more definitive conclusions. If diet, medication use, or other factors were to be incorporated,

it is necessary to evaluate whether LSTM remains the best-performing machine learning architecture.

## Conclusion

In this study, we show that our machine learning-based models are able to accurately and safely predict glucose values for up to 60 minutes in individuals with, NGM, prediabetes, or type 2 diabetes. In addition, translation of our prediction models to individuals with type 1 diabetes showed encouraging results. We observed particularly high precision at a 15-minute prediction window, which is a clinically relevant timespan to align interstitially measured glucose values by continuous glucose measurement systems with actual plasma glucose values. As such, the prediction model can be used to improve closed-loop insulin delivery systems by overcoming sensor delay. In addition, longer prediction intervals may be used to safely bridge periods of sensor malfunction. Last, our current findings question the use of accelerometry to substantially improve prediction. Future research should validate our findings by replicating the results in a larger sample of individuals with type 1 diabetes and studying the effects of implementing the prediction model in a closed-loop insulin delivery system.

# References

1. Collaboration NCDRF. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 2016;387(10027):1513-30.
2. Emerging Risk Factors C, Sarwar N, Gao P, Seshasai SR, Gobin R, Kaptoge S, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet. 2010;375(9733):2215-22.
3. Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev. 2013;93(1):137-88.
4. American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2019. Diabetes Care. 2019;42(Suppl 1):S61-S70.
5. International Hypoglycaemia Study G. Hypoglycaemia, cardiovascular disease, and mortality in diabetes: epidemiology, pathogenesis, and management. Lancet Diabetes Endocrinol. 2019;7(5): 385-96.
6. Cobelli C, Renard E, Kovatchev B. Artificial pancreas: past, present, future. Diabetes. 2011;60(11): 2672-82.
7. Bruttomesso D. Toward Automated Insulin Delivery. N Engl J Med. 2019;381(18):1774-5.
8. Weisman A, Bai JW, Cardinez M, Kramer CK, Perkins BA. Effect of artificial pancreas systems on glycaemic control in patients with type 1 diabetes: a systematic review and meta-analysis of outpatient randomised controlled trials. Lancet Diabetes Endocrinol. 2017;5(7):501-12.
9. Kumareswaran K, Thabit H, Leelarathna L, Caldwell K, Elleri D, Allen JM, et al. Feasibility of closed-loop insulin delivery in type 2 diabetes: a randomized controlled study. Diabetes Care. 2014;37(5):1198-203.
10. Blauw H, Keith-Hynes P, Koops R, DeVries JH. A Review of Safety and Design Requirements of the Artificial Pancreas. Ann Biomed Eng. 2016;44(11):3158-72.
11. Rodbard D. Continuous Glucose Monitoring: A Review of Successes, Challenges, and Opportunities. Diabetes Technol Ther. 2016;18 Suppl 2:S3-S13.
12. Woldaregay AZ, Arsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, et al. Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes. Artif Intell Med. 2019;98:109-34.
13. Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.
14. Foreman YD, Brouwers M, van der Kallen CJH, Pagen DME, van Greevenbroek MMJ, Henry RMA, et al. Glucose variability assessed with continuous glucose monitoring: reliability, reference values and correlations with established glycaemic indices - The Maastricht Study. Diabetes Technol Ther. 2019.
15. van der Berg JD, Stehouwer CD, Bosma H, van der Velde JH, Willems PJ, Savelberg HH, et al. Associations of total amount and patterns of sedentary behaviour with type 2 diabetes and the metabolic syndrome: The Maastricht Study. Diabetologia. 2016;59(4):709-18.
16. WHO. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. WHO. 2006.
17. Staudemeyer RC, Rothstein Morris E. Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks. arXiv e-prints [Internet]. 2019 September 01, 2019. Available from: https://ui.adsabs.harvard.edu/abs/2019arXiv190909586S.
18. Sherstinsky A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network. arXiv e-prints. 2018:arXiv:1808.03314.
19. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986;323(6088):533-6.
20. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. arXiv e-prints. 2014:arXiv:1412.3555.
21. Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput. 1997;9(8):1735-80.

3

22. Graves A, Fernández S, Schmidhuber J. Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition2005. 799-804 p.

23. Schuster M, Paliwal K. Bidirectional recurrent neural networks. Signal Processing, IEEE Transactions on. 1997;45:2673-81.

24. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. arXiv e-prints [Internet]. 2014 December 01, 2014:[arXiv:1412.6980 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2014arXiv1412.6980K.

25. Marling C, Bunescu RC, editors. The OhioT1DM Dataset For Blood Glucose Level Prediction. KHD@IJCAI; 2018.

26. Efron B, Tibshirani RJ. An introduction to the bootstrap. New York, N.Y.; London: Chapman & Hall; 1993.

27. Klonoff DC, Lias C, Vigersky R, Clarke W, Parkes JL, Sacks DB, et al. The surveillance error grid. J Diabetes Sci Technol. 2014;8(4):658-72.

28. Pfutzner A, Klonoff DC, Pardo S, Parkes JL. Technical aspects of the Parkes error grid. J Diabetes Sci Technol. 2013;7(5):1275-81.

29. Bland JM, Altman DG. Correlation in restricted ranges of data. BMJ. 2011;342:d556.

30. Sudharsan B, Peeples M, Shomali M. Hypoglycemia prediction using machine learning models for patients with type 2 diabetes. J Diabetes Sci Technol. 2015;9(1):86-90.

31. Georga E, Protopappas V, Fotiadis D. Glucose Prediction in Type 1 and Type 2 Diabetic Patients Using Data Driven Techniques. 2011.

32. Faruqui SHA, Du Y, Meka R, Alaeddini A, Li C, Shirinkam S, et al. Development of a Deep Learning Model for Dynamic Forecasting of Blood Glucose Level for Type 2 Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial. JMIR Mhealth Uhealth. 2019;7(11):e14452.

33. Albers DJ, Levine M, Gluckman B, Ginsberg H, Hripcsak G, Mamykina L. Personalized glucose forecasting for type 2 diabetes using data assimilation. PLoS Comput Biol. 2017;13(4):e1005232.

34. Mohebbi A, Johansen AR, Hansen N, Christensen PE, Tarp JM, Jensen ML, et al. Short Term Blood Glucose Prediction based on Continuous Glucose Monitoring Data. arXiv e-prints [Internet]. 2020 February 01, 2020:[arXiv:2002.02805 p.]. Available from: https://ui.adsabs. harvard edu/abs/2020arXiv200202805M.

35. Martinsson J, Schliep A, Eliasson B, Mogren O. Blood Glucose Prediction with Variance Estimation Using Recurrent Neural Networks. Journal of Healthcare Informatics Research. 2020;4(1):1-18.

36. Kriventsov S, Lindsey A, Hayeri A. The Diabits App for Smartphone-Assisted Predictive Monitoring of Glycemia in Patients With Diabetes: Retrospective Observational Study. JMIR Diabetes. 2020;5(3):e18660.

37. Li K, Liu C, Zhu T, Herrero P, Georgiou P. GluNet: A Deep Learning Framework for Accurate Glucose Forecasting. IEEE J Biomed Health Inform. 2020;24(2):414-23.

38. Chen J, Li K, Herrero P, Zhu T, Georgiou P, editors. Dilated Recurrent Neural Network for Short-time Prediction of Glucose Concentration. KHD@IJCAI; 2018.

39. Stanford KI, Goodyear LJ. Exercise and type 2 diabetes: molecular mechanisms regulating glucose uptake in skeletal muscle. Adv Physiol Educ. 2014;38(4):308-14.

40. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. Am J Epidemiol. 2012;176(6):473-81.

41. Blauw H, Onvlee AJ, Klaassen M, van Bon AC, DeVries JH. Fully Closed Loop Glucose Control With a Bihormonal Artificial Pancreas in Adults With Type 1 Diabetes: An Outpatient, Randomized, Crossover Trial. Diabetes Care. 2021.

# Supplemental materials

## *Supporting information S3.1. Background information on machine learning models reviewed in current study*

We conducted a comparison of available algorithms on the continuous glucose prediction task. We considered the following algorithms:

- **ARIMA:** an autoregressive integrated moving average (ARIMA) model is an adaptation of the autoregressive moving average (ARMA) model. These models aim to estimate the time series using two polynomials, one for autoregression (AR), and the other for the moving average (MA). ARIMA models for glucose prediction have widely been described in the literature[1-4]. For the current comparison, we used a ARIMA (p=3, d=0, q=0) model based on previous findings by Otoom et al[1].

- **Support vector regression:** support vector regression (SVR) is a generalization of support vector machines (SVM) that work for regression problems. SVR models for the prediction of glucose values have been described in the literature[5, 6]. For the current comparison, we build the SVR model with a Gaussian RFB kernel that was optimized by the differential evolution algorithm as described by Georga et al[5].

- **Gradient-boosting trees:** gradient-boosting trees have shown superior performances in the medical domain, but lack the capability to aggregate information over time. For baseline comparison, we used the LightGBM[7] implementation with a learning rate of 0.01, a maximum number of trees of 500, and a maximum depth of each base learner to be 50.

- **Feed-forward neural networks:** feed-forward neural networks do not have the capacity to aggregate the information about past glucose status of an individual over time. Yet, they are relatively simple networks which have previously been employed for glucose[8-10]. We considered shallow and deep multi-layer perceptron (MLP) neural networks. The shallow MLP consisted of one hidden layer (ReLU) with 16 neurons. The deep MLP consisted of three hidden layers (ReLU) with 64, 32, and 16 neurons, respectively. Both networks were trained using the Adam optimizer scheme with a learning rate of 0.001.

- **Recurrent neural networks:** recurrent neural networks are a type of neural networks with the capacity to explicitly model information over time. These type of neural networks are designed to work with temporal data such as glucose and physical activity data. Although we acknowledge the broad availability of neural network architectures[11-15], we only considered recurrent neural networks (RNN), and long-short term memory (LSTM) networks for the baseline comparison. Both models consisted of one RNN layer (either RNN or LSTM) with 32 neurons, followed by a Dense layer of 8 neurons. Both networks were trained using the Adam optimizer scheme with a learning rate of 0.001.

## *Supporting information S3.2. Background information on metrics used in the current study*

In the current study we used several metrics to assess the performance of our models. In this paragraph we will describe each of them briefly and also provide their mathematical definition.

- Root-mean-square error (RMSE): the RMSE is an average error term which is in the order of the predicted / measured variable. It also can be interpreted as the standard deviation of the prediction errors. The formula of RMSE is as follows:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

*with $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_n$ depicting the predicted values; $y_1$, $y_2$, $y_n$ depicting the real values and $n$ representing the total number of predictions*

- Correlation: correlation is a measure of how well the relationship between two variables is. In this study, we deal with non-parametrically distributed data and therefore use Spearman's rank correlation coefficient (rho) which tries to describe the relationship of two variables using a monotonic function:

$$\rho = \frac{S_{xy}}{S_x S_y} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left((R(x_i) - \overline{R(x)}) \cdot (R(y_i) - \overline{R(y)})\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left((R(x_i) - \overline{R(x)})^2\right) \cdot \left(\frac{1}{n}\sum_{i=1}^{n}(R(y_i) - \overline{R(y)})^2\right)}}$$

*with $x_1$, $x_2$, $x_n$ depicting the predicted values; $y_1$, $y_2$, $y_n$ depicting the real values and $n$ representing the total number of predictions*

- Time lag: time lag is a measure of the time shift between the actual and predicted glucose profile which results in the highest cross correlation coefficient between them[10,16]:

$$\tau_{delay} = \arg\max_{k} \left(\hat{y}_k(k|k - PH) \cdot y(k)\right)$$

*with $y$ depicting the real values; $\hat{y}_k$ depicting the predicted values and PH depicting the prediction horizon*

## Figures
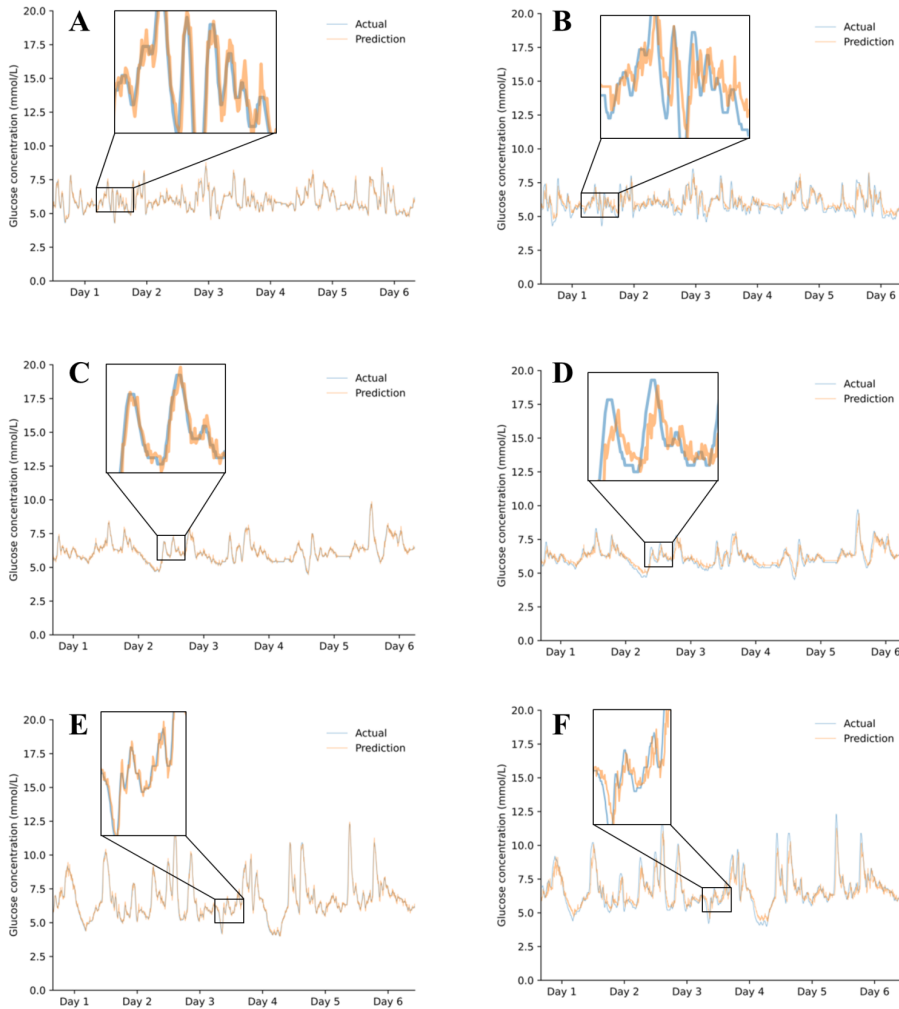


**Figure S3.1  Illustrative examples of continuous glucose monitoring-based machine learning model predictions compared to actual values.** Predictions in an individual with normal glucose metabolism (NGM) on a time interval of 15 (A) and 60 (B) minutes. Predictions in an individual with prediabetes (PreD) on a time interval of 15 (C) and 60 (D) minutes. Predictions in an individual with type 2 diabetes (T2D) on a time interval of 15 (E) and 60 (F) minutes.
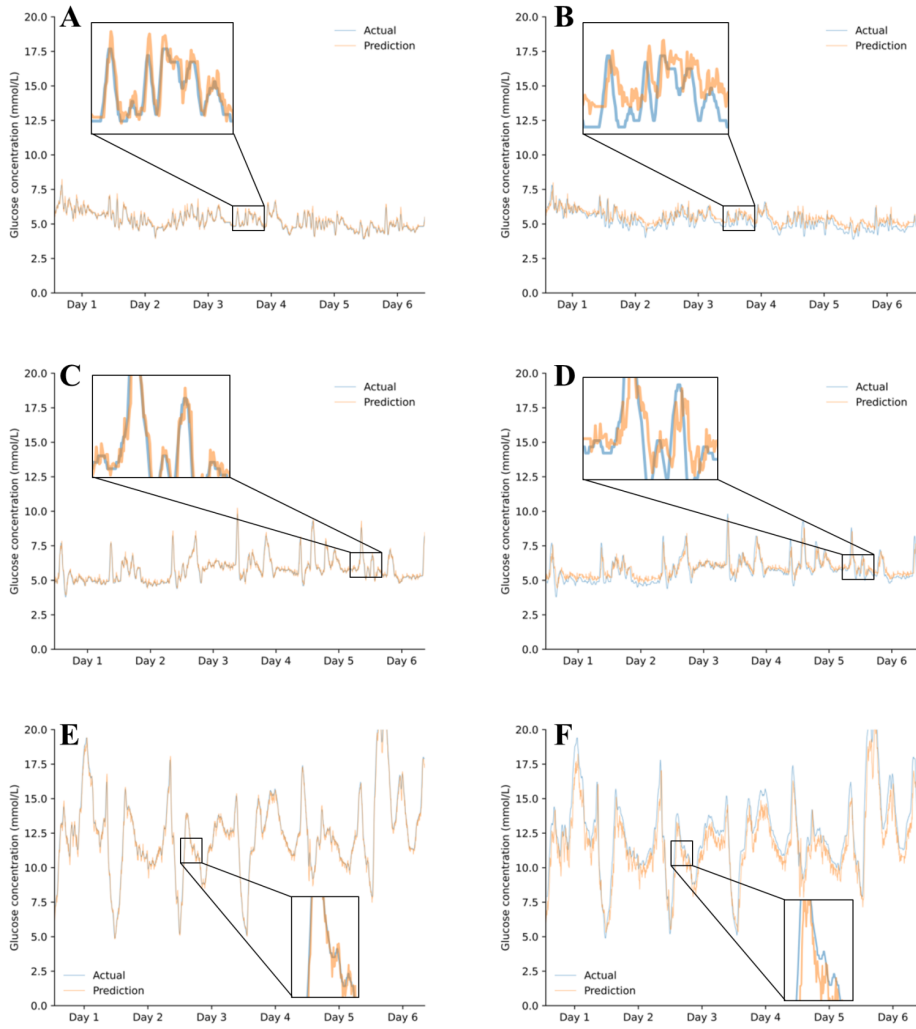
Figure S3.2 **Illustrative examples of continuous glucose monitoring- and accelerometry-based machine learning model predictions compared to actual values.** Predictions in an individual with normal glucose metabolism (NGM) on a time interval of 15 (A) and 60 (B) minutes. Predictions in an individual with prediabetes (PreD) on a time interval of 15 (C) and 60 (D) minutes. Predictions in an individual with type 2 diabetes (T2D) on a time interval of 15 (E) and 60 (F) minutes.

**A** **B**
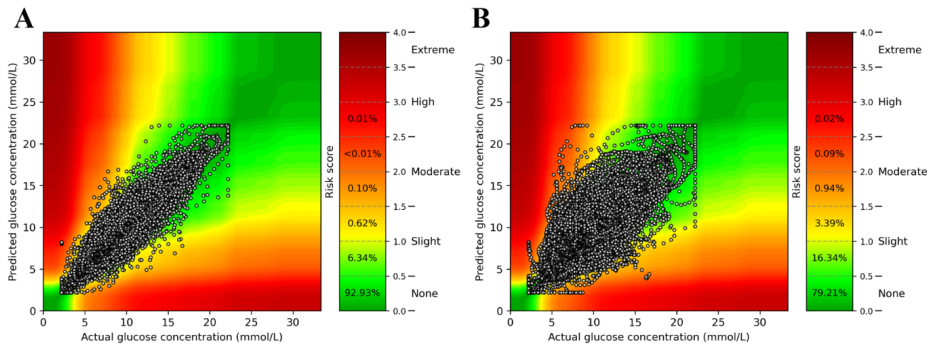


**Figure S3.3** **Surveillance error grid evaluation of glucose prediction safety at time intervals of 15 and 60 minutes using glucose value t0 as predictor.** Assessment of glucose prediction safety in individuals with type 2 diabetes (n=43) at 15 minutes (panel A) and 60 minutes (panel B) using a naïve approach with t0 as predictor. The risk score values translate to the following degrees of risk: 0-0.5, none; 0.5-1.0, slight (lower); 1.0-1.5, slight (higher); 1.5-2.0, moderate (lower); 2.0-2.5, moderate (higher); 2.5–3.0, great (lower); 3.0-3.5, great (higher); >3.5 extreme.
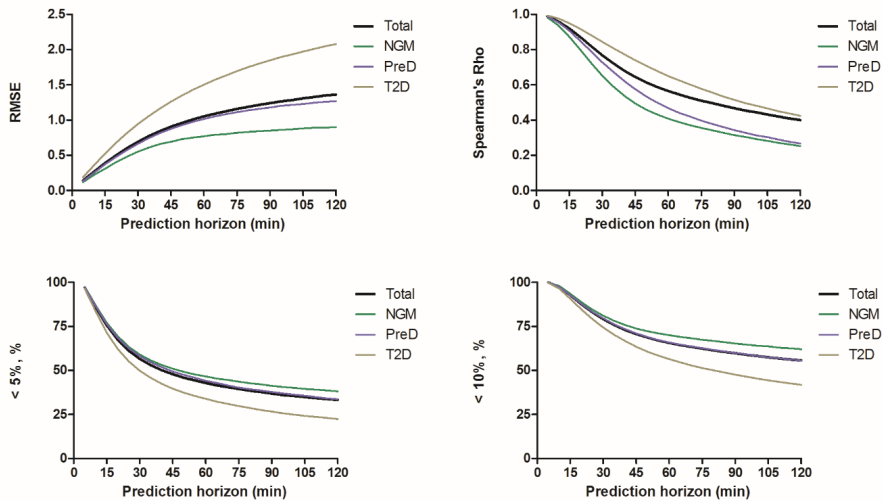


**Figure S3.4** **Performance characteristics of a prediction model using t0 as predictor across time horizons between 0 and 120 minutes.** An extended analysis of model performance using t0 glucose value as predictor was carried out for normal glucose metabolism (NGM), prediabetes (PreD) and individuals with type 2 diabetes (T2D). Models were evaluated using RMSE (root-mean-square error; upper-left), Spearman's rank correlation coefficient (upper-right) and percentage of predicted values within 5% and 10% of actual glucose values, respectively (lower-left and right).
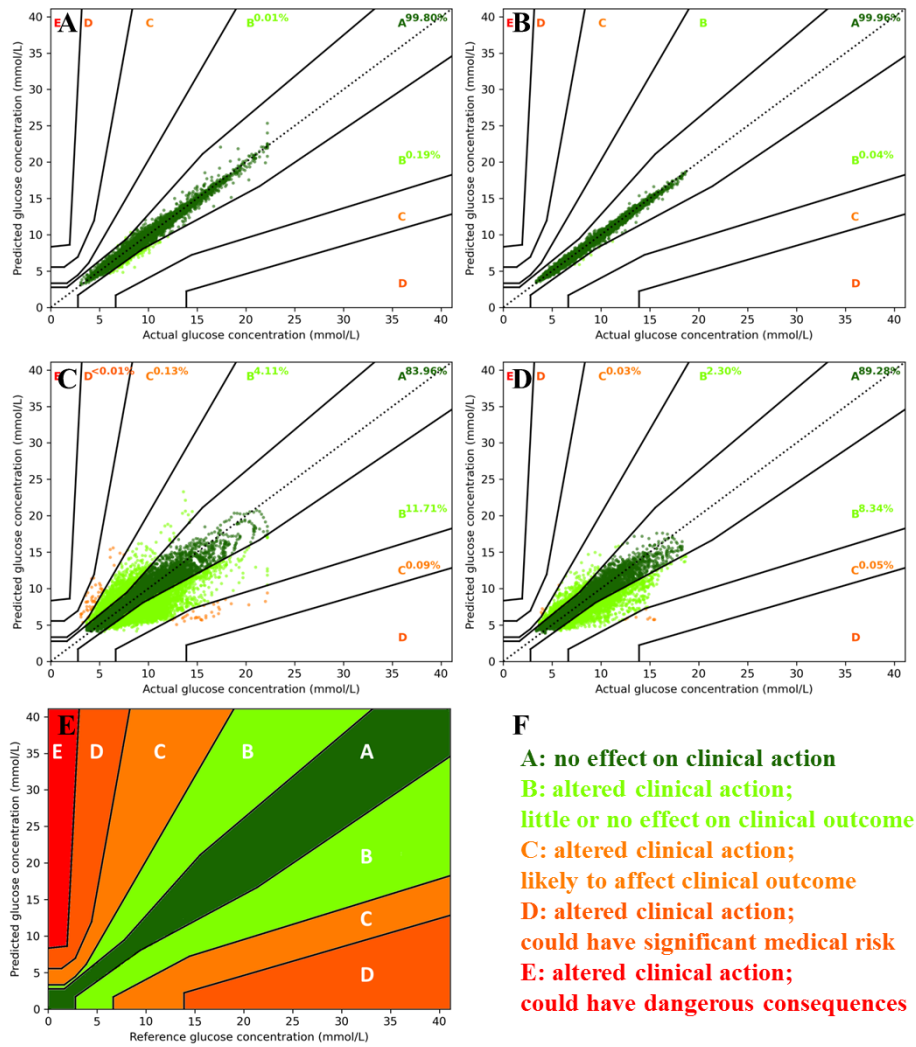
**Figure S3.5** **Parkes error grid evaluation of glucose prediction safety at time intervals of 15 and 60 minutes.** Assessment of CGM-based glucose prediction safety in individuals with type 2 diabetes (n=43) at 15 minutes (panel A) and 60 minutes (panel C). Assessment of CGM- and accelerometry-based glucose prediction safety in individuals with type 2 diabetes (n=13) at 15 minutes (panel B) and 60 minutes (panel D). Each error zone (panel E) has a different clinical interpretation and consequence (panel F). A horizontal shift towards zone E represents overestimation by the algorithm (higher predicted than actual glucose values); a vertical shift towards zone D represents underestimation (lower predicted than actual glucose values).

## Tables

Table S3.1    Baseline statistical and machine learning model comparison for predicting glucose values.

| Prediction window and baseline model | CGM-based glucose prediction | | Combined glucose prediction | |
|---|---|---|---|---|
| | Rho | RMSE, mmol/L | Rho | RMSE, mmol/L |
| 15 minutes | | | | |
| ARIMA | 0.842 | 0.504 | 0.834 | 0.498 |
| | [0.837 − 0.848] | [0.490 − 0.518] | [0.829 − 0.840] | [0.492 − 0.505] |
| SVR | 0.791 | 0.558 | 0.703 | 0.612 |
| | [0.781 − 0.802] | [0.549 − 0.567] | [0.694 − 0.712] | [0.601 − 0.622] |
| LightGBM | 0.783 | 0.589 | 0.783 | 0.497 |
| | [0.767 − 0.795] | [0.577 − 0.601] | [0.771 − 0.794] | [0.582 − 0.613] |
| Shallow MLP | 0.810 | 0.517 | 0.763 | 0.592 |
| | [0.804 − 0.816] | [0.506 − 0.529] | [0.754 − 0.772] | [0.581 − 0.603] |
| Deep MLP | 0.807 | 0.511 | 0.828 | 0.510 |
| | [0.797 − 0.818] | [0.504 − 0.518] | [0.819 − 0.837] | [0.503 − 0.517] |
| RNN | 0.894 | 0.485 | 0.890 | 0.477 |
| | [0.887 − 0.902] | [0.481 − 0.490] | [0.882 − 0.898] | [0.472 − 0.482] |
| LSTM | 0.872 | 0.482 | 0.884 | 0.501 |
| | [0.865 − 0.879] | [0.477 − 0.487] | [0.878 − 0.890] | [0.496 − 0.506] |
| 60 minutes | | | | |
| ARIMA | 0.307 | 1.543 | 0.303 | 1.502 |
| | [0.284 − 0.329] | [1.489 − 1.623] | [0.283 − 0.322] | [1.455 − 1.568] |
| SVR | 0.388 | 1.386 | 0.394 | 1.412 |
| | [0.376 − 0.398] | [1.322 − 1.452] | [0.382 − 0.405] | [1.350 − 1.475] |
| LightGBM | 0.500 | 1.118 | 0.498 | 1.128 |
| | [0.491 − 0.508] | [1.098 − 1.136] | [0.485 − 0.511] | [1.107 − 1.148] |
| Shallow MLP | 0.503 | 1.081 | 0.483 | 1.081 |
| | [0.495 − 0.511] | [1.074 − 1.088] | [0.470 − 0.495] | [1.070 − 1.092] |
| Deep MLP | 0.496 | 1.108 | 0.515 | 1.108 |
| | [0.484 − 0.509] | [1.100 − 1.115] | [0.502 − 0.528] | [1.099 − 1.017] |
| RNN | 0.591 | 0.989 | 0.596 | 0.992 |
| | [0.581 − 0.600] | [0.983 − 0.995] | [0.589 − 0.603] | [0.984 − 0.998] |
| LSTM | 0.605 | 0.941 | 0.602 | 0.922 |
| | [0.593 − 0.616] | [0.937 − 0.945] | [0.595 − 0.609] | [0.919 − 0.926] |

Comparison of statistical and machine learning models for the continuous glucose prediction. We trained models on the training dataset and evaluated their performance on the tuning dataset. Performance was assessed by Spearman's rank correlation coefficient (rho) and root-mean-square error (RMSE). 95% confidence intervals were calculated using 1,000 bootstraps.

3

Table S3.2    Hyperparameter combinations evaluated in current experiments.

| Hyperparameter | Values considered |
|---|---|
| Data preprocessing | |
|   Normalization to [0, 1] | On, off |
|   Back-propagation window | 15, 30, 60, 120 minutes |
| Neural Network architecture | |
|   RNN cell type | LSTM, RNN, GRU |
|   RNN cell type, bi-directional structure | On, off |
|   RNN number of hidden layers | 1, 2, 3 |
|   RNN cell size | 128, 64, 32, 16, 8, 4 |
|   RNN, activation function | ReLU, leaky ReLU, tanh, sigmoid, ELU |
|   Dropout, presence | On, off |
|   Dropout | 0.05, 0.1, 0.2, 0.25 |
| Model training | |
|   Learning rate | $1^e$-2, $1^e$-3, $1^e$-4, $1^e$-5 |
|   Learning rate scheduling | On, off |
|   Learning rate scheduling decay | 0.5, 0.6, 0.7, 0.8, 0.9, 0.95 |
|   Batch size | 64, 128, 256, 512, 1024, 2048, 4096 |
|   Optimizer scheme | Adam, NAdam, RMSprop, SGD |

Table S3.3    Final set of hyperparameters for each of the machine learning models.

| Hyperparameter | CGM-based glucose prediction | Combined glucose prediction |
|---|---|---|
| Data preprocessing | | |
|   Normalization to [0, 1] | On | On |
|   Back-propagation window | 30 minutes | 30 minutes |
| Neural Network architecture | | |
|   RNN cell type | LSTM, LSTM, Dense | LSTM, LSTM, Dense |
|   RNN cell type, bi-directional structure | Off, off, off | On, off, off |
|   RNN number of hidden layers | 3 | 3 |
|   RNN cell size | 32, 16, 8 | 64, 32, 8 |
|   RNN, activation function | ReLU, ReLU, ReLU | ReLU, ReLU, ReLU |
|   Dropout, presence | Off, off, on | Off, off, on |
|   Dropout | 0.1 | 0.1 |
| Model training | | |
|   Learning rate | 0.001 | 0.001 |
|   Learning rate scheduling | On | On |
|   Learning rate scheduling decay | 0.005 every 1,000 steps | 0.005 every 1,000 steps |
|   Batch size | 1024 | 1024 |
|   Optimizer scheme | Adam | Adam |

**Table S3.4    Extended baseline characteristics.**

| Characteristic | CGM-based glucose prediction | | | Combined glucose prediction | | |
|---|---|---|---|---|---|---|
| | Total (n=851) | NGM (n=470) | PreD (n=184) | Total (n=540) | NGM (n=373) | PreD (n=99) |
| Age, years | 59.9 ± 8.7 | 58.2 ± 8.8 | 61.5 ± 8.1 | 59.1 ± 8.7 | 58.1 ± 8.9 | 60.7 ± 8.5 |
| Women, n (%) | 418 (49.1) | 266 (56.6) | 83 (45.1) | 276 (51.1) | 207 (55.5) | 47 (47.5) |
| BMI, kg/m² | 27.2 ± 4.4 | 25.6 ± 3.6 | 28.5 ± 4.4 | 26.5 ± 4.0 | 25.6 ± 3.5 | 28.8 ± 4.4 |
| Newly diagnosed T2D, n (%) | 70 (8.2) | - | - | 35 (6.5) | - | - |
| Fasting plasma glucose, mmol/L | 5.4 [5.0 – 6.2] | 5.1 [4.8 – 5.4] | 6.0 [5.4 – 6.3] | 5.3 [4.9 – 5.8] | 5.1 [4.8 – 5.4] | 5.9 [5.3 – 6.2] |
| 2-h post-load glucose, mmol/L | 6.7 [5.2 – 9.1]* | 5.5 [4.7 – 6.4] | 8.4 [7.5 – 9.2] | 6.2 [5.0 – 7.7] | 5.5 [4.7 – 6.4] | 8.4 [7.8 – 9.3] |
| HbA$_{1c}$, % | 5.7 ± 0.8 | 5.4 ± 0.3 | 5.6 ± 0.4 | 5.6 ± 0.6 | 5.4 ± 0.3 | 5.5 ± 0.4 |
| HbA$_{1c}$, mmol/mol | 39.1 ± 8.3 | 35.4 ± 3.4 | 37.8 ± 4.2 | 37.3 ± 6.2 | 35.5 ± 3.4 | 37.1 ± 4.4 |
| Sensor glucose | | | | | | |
| Mean, mmol/L | 6.1 [5.7 – 6.7] | 5.8 [5.5 – 6.1] | 6.2 [5.8 – 6.6] | 5.9 [5.6 – 6.4] | 5.9 [5.5 – 6.1] | 6.2 [5.7 – 6.6] |
| SD, mmol/L | 0.84 [0.68 – 1.18] | 0.73 [0.62 – 0.87] | 0.90 [0.74 – 1.13] | 0.79 [0.66 – 1.01] | 0.73 [0.63 – 0.89] | 0.89 [0.73 – 1.11] |
| SD >1.37 mmol/L, n (%) | 142 (16.7) | 11 (2.3) | 16 (8.7) | 50 (9.3) | 9 (2.4) | 5 (5.1) |
| CV, % | 14.0 [11.6 – 17.6] | 12.6 [10.8 – 14.9] | 14.9 [12.2 – 17.5] | 13.3 [11.2 – 16.8] | 12.8 [10.9 – 15.2] | 14.7 [12.1 – 17.5] |
| Diabetes medication use, n (%) | 109 (12.8)† | - | - | 27 (4.8) | - | - |
| Insulin | 19 (2.2) | - | - | 4 (0.7) | - | - |
| Metformin | 104 (12.2) | - | - | 27 (5.0) | - | - |
| Sulfonylureas | 21 (2.5) | - | - | 6 (1.1) | - | - |
| Thiazolidinediones | 0 (0) | - | - | 0 (0) | - | - |
| GLP-1 analogs | 3 (0.4) | - | - | 1 (0.2) | - | - |
| DDP-4 inhibitors | 1 (0.1) | - | - | 0 (0) | - | - |
| SGLT-2 inhibitors | 1 (0.1) | - | - | 0 (0) | - | - |
| Office SBP, mmHg | 133.3 ± 18.0 | 129.3 ± 17.5 | 137.2 ± 19.2 | 132.2 ± 17.9 | 129.6 ± 17.6 | 138.1 ± 18.6 |
| Office DBP, mmHg | 75.2 ± 10.2 | 73.5 ± 9.8 | 76.7 ± 10.3 | 74.7 ± 10.1 | 73.5 ± 9.8 | 77.2 ± 10.8 |
| Antihypertensive medication use, n (%) | 305 (35.9)† | 105 (22.3) | 74 (40.2) | 162 (30.0) | 84 (22.5) | 37 (37.4) |
| Total-to-HDL cholesterol ratio | 3.5 [2.8 – 4.3] | 3.3 [2.8 – 4.4] | 3.8 [3.1 – 4.7] | 3.4 [2.8 – 4.3] | 3.3 [2.8 – 4.3] | 3.6 [2.9 – 4.5] |
| Triglycerides, mmol/L | 1.3 [0.9 – 1.8] | 1.1 [0.8 – 1.5] | 1.4 [1.0 – 2.0] | 1.2 [0.9 – 1.7] | 1.1 [0.8 – 1.5] | 1.4 [1.0 – 1.9] |
| Lipid-modifying medication use, n (%) | 212 (24.9)† | 52 (11.1) | 45 (24.5) | 100 (18.5) | 38 (10.2) | 23 (23.2) |
| Smoking status | | | | | | |
| Never/former/current, n | 327/415/106‡ | 198/210/60 | 62/101/20 | 214/253/70 | 160/164/47 | 35/52/10 |
| Never/former/current, % | 38.6/48.9/12.5 | 42.3/44.9/12.8 | 33.9/55.2/10.9 | 39.9/47.1/13.0 | 43.1/44.2/12.7 | 35.7/54.1/10.2 |

Data are reported as mean ± SD, median [interquartile range], or number (percentage [%]) as appropriate. CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; BMI, body mass index; HbA1c, glycated hemoglobin A1c; S D, standard deviation; CV, coefficient of variation; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4; SGLT-2, sodium-glucose cotransporter 2; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL, high-density lipoprotein. * Missing in 38 participants; † missing in one participant; ‡ missing in three participants.

Table S3.5   Extended analysis of model performance in normal glucose metabolism and prediabetes subgroups.

| | CGM-based glucose prediction | | Combined glucose prediction | |
|---|---|---|---|---|
| | NGM (n=92) | PreD (n=35) | NGM (n=75) | PreD (n=21) |
| **15 minutes** | | | | |
| RMSE, mmol/L | 0.158 [0.154–0.161] | 0.238 [0.233–0.343] | 0.151 [0.149–0.153] | 0.232 [0.227–0.237] |
| <5% , % | 93.56 [93.50–93.62] | 92.58 [92.51–92.66] | 93.76 [93.71–93.80] | 92.65 [92.57–97.73] |
| <10% , % | 99.47 [99.42–99.52] | 99.01 [98.98–99.05] | 99.65 [99.62–99.68] | 99.08 [99.04–99.12] |
| Rho | 0.951 [0.948–0.954] | 0.973 [0.967–0.980] | 0.953 [0.950–0.957] | 0.974 [0.968–0.979] |
| **60 minutes** | | | | |
| RMSE, mmol/L | 0.501 [0.498–0.505] | 0.602 [0.594–0.610] | 0.503 [0.495–0.510] | 0.599 [0.594–0.604] |
| <5% , % | 74.19 [74.11–74.26] | 69.48 [69.39–69.56] | 75.02 [74.95–75.08] | 70.01 [69.95–70.07] |
| <10% , % | 89.89 [89.82–89.97] | 87.43 [87.36–87.50] | 89.25 [89.19–89.31] | 88.20 [88.09–88.30] |
| Rho | 0.699 [0.697–0.702] | 0.732 [0.727–0.738] | 0.701 [0.697–0.705] | 0.739 [0.732–0.747] |

Data are reported as mean [95% confidence interval]. CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; RMSE, root-mean-square error; <5%, percentage of predicted values within 5% of actual glucose values; <10%, percentage of predicted values within 10% of actual glucose values; rho, Spearman's rank correlation coefficient.

Table S3.6   Extended analysis on time lag between predicted and actual glucose values.

| Time lag | Total (n=170) | NGM (n=92) | PreD (n=35) | T2D (n=43) |
|---|---|---|---|---|
| **15 minutes** | | | | |
| CGM-based | 0.12 ± 0.18 | 0.08 ± 0.12 | 0.07 ± 0.11 | 0.41 ± 0.92 |
| Combined | 0.17 ± 0.11 | 0.11 ± 0.18 | 0.10 ± 0.18 | 0.44 ± 0.38 |
| **60 minutes** | | | | |
| CGM-based | 12.28 ± 6.84 | 7.02 ± 3.46 | 9.03 ± 3.77 | 14.92 ± 11.18 |
| Combined | 11.95 ± 7.32 | 7.19 ± 2.87 | 8.75 ± 3.91 | 14.28 ± 9.96 |

Data are reported as mean ± SD. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes.

| Time lag | Total (n=170) | NGM (n=92) | PreD (n=35) | T2D (n=43) |
|---|---|---|---|---|
| **15 minutes** | | | | |
| CGM-based | 0 [0 − 5] | 0 [0 − 0] | 0 [0 − 0] | 0 [0 − 0] |
| Combined | 0.50 [0.25 − 0.75] | 0.25 [0.0 − 0.75] | 0.25 [0 − 0.50] | 0.50 [0.25 − 0.75] |
| **60 minutes** | | | | |
| CGM-based | 10 [5 − 15] | 10 [0 − 15] | 10 [5 − 15] | 15 [5 − 20] |
| Combined | 9.50 [4.25 − 16.50] | 6.75 [4.25 − 9.50] | 7.50 [4.50 − 10.75] | 14.50 [6.75 − 21.50] |

Data are reported as median [IQR]. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes.

Table S3.7   Extended analysis of model performance with $t_0$ glucose value as predictor.

| Prediction with $t_0$ | Total (n=170) | NGM (n=92) | PreD (n=35) | T2D (n=43) |
|---|---|---|---|---|
| 15 minutes | | | | |
| RMSE, mmol/L | 0.158 [0.154–0.161] | 0.238 [0.233–0.343] | 0.151 [0.149 – 0.153] | 0.232 [0.227 – 0.237] |
| <5% , % | 93.56 [93.50–93.62] | 92.58 [92.51–92.66] | 93.76 [93.71 – 93.80] | 92.65 [92.57 – 97.73] |
| <10% , % | 99.47 [99.42–99.52] | 99.01 [98.98–99.05] | 99.65 [99.62–99.68] | 99.08 [99.04–99.12] |
| Rho | 0.951 [0.948–0.954] | 0.973 [0.967–0.980] | 0.953 [0.950–0.957] | 0.974 [0.968–0.979] |
| 60 minutes | | | | |
| RMSE, mmol/L | 0.501 [0.498–0.505] | 0.602 [0.594–0.610] | 0.503 [0.495–0.510] | 0.599 [0.594–0.604] |
| <5% , % | 74.19 [74.11–74.26] | 69.48 [69.39–69.56] | 75.02 [74.95–75.08] | 70.01 [69.95–70.07] |
| <10% , % | 89.89 [89.82–89.97] | 87.43 [87.36–87.50] | 89.25 [89.19–89.31] | 88.20 [88.09–88.30] |
| Rho | 0.699 [0.697–0.702] | 0.732 [0.727–0.738] | 0.701 [0.697–0.705] | 0.739 [0.732–0.747] |

Data are reported as mean [95% confidence interval]. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; RMSE, root-mean-square error; <5%, percentage of predicted values within 5% of actual glucose values; <10%, percentage of predicted values within 10% of actual glucose values; rho, Spearman's rank correlation coefficient.

TableS3.8   Model performance stratified by day and night.

| 15 minutes | CGM-based glucose prediction | | Combined glucose prediction | |
|---|---|---|---|---|
| | Total (n=170) | T2D (n=43) | Total (n=109) | T2D (n=13) |
| Day | | | | |
| RMSE, mmol/L | 0.199 [0.196–0.202] | 0.300 [0.295–0.305] | 0.197 [0.193–0.201] | 0.287 [0.283–0.291] |
| <5% , % | 92.92 [92.85–92.99] | 91.97 [91.87–92.07] | 92.94 [92.90–92.98] | 91.95 [91.92–91.98] |
| <10% , % | 99.11 [99.07–99.15] | 98.84 [98.72–98.95] | 99.17 [99.13–99.21] | 98.82 [98.78–98.86] |
| Rho | 0.955 [0.952–0.958] | 0.984 [0.981–0.987] | 0.964 [0.962–0.966] | 0.986 [0.984–0.988] |
| Night | | | | |
| RMSE, mmol/L | 0.182 [0.175–0.189] | 0.278 [0.273 – 0.283] | 0.178 [0.173 – 0.183] | 0.257 [0.252 – 0.262] |
| <5% , % | 93.08 [93.01–93.15] | 92.07 [91.99 – 92.15] | 93.11 [93.07 – 93.15] | 92.14 [92.10 – 92.18] |
| <10% , % | 99.28 [99.20–99.36] | 98.94 [98.88–99.00] | 99.34 [99.30–99.38] | 99.03 [98.98–99.08] |
| Rho | 0.967 [0.964–0.970] | 0.989 [0.986–0.992] | 0.974 [0.970–0.978] | 0.994 [0.992–0.996] |

| 60 minutes | CGM-based glucose prediction | | Combined glucose prediction | |
|---|---|---|---|---|
| | Total (n=170) | T2D (n=43) | Total (n=109) | T2D (n=13) |
| Day | | | | |
| RMSE, mmol/L | 0.687 [0.683–0.691] | 0.775 [0.768–0.783] | 0.536 [0.532–0.540] | 0.768 [0.760–0.776] |
| <5% , % | 68.40 [68.35–68.45] | 63.69 [63.57–63.81] | 68.87 [68.80–68.94] | 64.03 [63.95–64.11] |
| <10% , % | 85.27 [85.20–85.35] | 83.82 [83.77–83.87] | 86.12 [86.07–86.17] | 84.08 [84.01–84.15] |
| Rho | 0.640 [0.629–0.651] | 0.703 [0.697–0.709] | 0.658 [0.654–0.663] | 0.710 [0.705–0.715] |
| Night | | | | |
| RMSE, mmol/L | 0.497 [0.491–0.503] | 0.634 [0.629–0.639] | 0.512 [0.503–0.521] | 0.633 [0.627–0.639] |
| <5% , % | 72.61 [72.56–72.66] | 69.42 [69.33–69.51] | 71.58 [71.47–71.69] | 69.28 [69.21–69.35] |
| <10% , % | 89.44 [89.31–89.57] | 87.10 [87.01–87.20] | 88.78 [88.68–88.88] | 87.30 [87.19–87.41] |
| Rho | 0.793 [0.784–0.802] | 0.854 [0.848–0.860] | 0.783 [0.780–0.786] | 0.861 [0.855–0.867] |

Data are reported as mean [95% confidence interval]. CGM, continuous glucose monitoring; T2D, type 2 diabetes; RMSE, root-mean-square error; <5%, percentage of predicted values within 5% of actual glucose values; <10%, percentage of predicted values within 10% of actual glucose values; rho, Spearman's rank correlation coefficient.

3

Table S3.9   Model performance stratified by low versus high glucose variability.

| | CGM-based glucose prediction | | Combined glucose prediction | |
|---|---|---|---|---|
| | SD ≤1.37 mmol/L (n=142) | SD >1.37 mmol/L (n=28) | SD ≤1.37 mmol/L (n=101) | SD >1.37 mmol/L (n=8) |
| 15 minutes | | | | |
| RMSE, mmol/L | 0.179 [0.176–0.181] | 0.301 [0.289–0.313] | 0.180 [0.177–0.183] | 0.288 [0.276–0.299] |
| <5% , % | 93.02 [92.99–93.05] | 91.88 [91.77–92.05] | 93.21 [93.18–93.24] | 92.00 [91.92–92.08] |
| <10% , % | 99.25 [99.22–99.28] | 98.79 [98.76–98.84] | 99.30 [99.29–99.32] | 98.82 [98.73–98.91] |
| Rho | 0.960 [0.959–0.962] | 0.983 [0.980–0.986] | 0.965 [0.964–0.967] | 0.992 [0.988–0.996] |
| 60 minutes | | | | |
| RMSE, mmol/L | 0.549 [0.542–0.555] | 0.711 [0.699–0.724] | 0.559 [0.552–0.565] | 0.710 [0.700–0.719] |
| <5% , % | 71.04 [70.89–71.21] | 65.33 [65.19–66.46] | 71.81 [71.77–71.85] | 66.17 [66.09–66.23] |
| <10% , % | 89.19 [89.15–89.22] | 84.42 [84.28–84.56] | 90.01 [89.95–90.06] | 85.25 [85.11–85.39] |
| Rho | 0.701 [0.700–0.702] | 0.741 [0.737–0.745] | 0.723 [0.719–0.728] | 0.801 [0.784–0.817] |

Data are reported as mean [95% confidence interval]. CGM, continuous glucose monitoring; SD, standard deviation; RMSE, root-mean-square error; <5%, percentage of predicted values within 5% of actual glucose values; <10%, percentage of predicted values within 10% of actual glucose values; rho, Spearman's rank correlation coefficient.

Table S3.10   Extended analysis of model performance in the OhioT1DM Dataset.

| | 15 minutes | 30 minutes | 60 minutes |
|---|---|---|---|
| Main model | | | |
| RMSE, mmol/L | 0.689 [0.685 − 0.693] | 1.189 [1.183 − 1.195] | 1.918 [1.910 − 1.926] |
| <5% , % | 61.84 [61.64 − 61.04] | 39.10 [38.86 − 39.34] | 22.28 [22.01 − 22.65] |
| <10% , % | 86.02 [85.83 − 86.21] | 64.72 [64.45 − 64.99] | 40.19 [39.89 − 40.50] |
| Rho | 0.908 [0.905 − 0.911] | 0.792 [0.789 − 0.795] | 0.605 [0.602 − 0.608] |
| Optimized model | | | |
| RMSE, mmol/L | 0.426 [0.422 − 0.430] | 1.046 [1.039 − 1.052] | 1.733 [1.725 − 1.741] |
| <5% , % | 72.22 [71.95 − 72.47] | 48.99 [48.77 − 49.22] | 39.85 [39.55 − 40.15] |
| <10% , % | 91.22 [91.01 − 91.43] | 71.48 [71.24 − 71.73] | 49.81 [49.48 − 50.14] |
| Rho | 0.948 [0.946 − 0.950] | 0.886 [0.884 − 0.888] | 0.689 [0.686 − 0.692] |

Data are reported as mean [95% confidence interval]. Main model: CGM-based model trained on main study population; Optimized model; main CGM-based model trained on main study population and portion of data from individuals with type 1 diabetes; RMSE, root-mean-square error; < 5%, percentage of predicted values within 5% of actual glucose values; < 10%, percentage of predicted values within 10% of actual glucose values; rho, Spearman's rank correlation coefficient.

# Supplementary references

1. Otoom M, Alshraideh H, Almasaeid HM, Lopez-de-Ipina D, Bravo J. Real-Time Statistical Modeling of Blood Sugar. J Med Syst. 2015;39(10):123.
2. Yang J, Li L, Shi Y, Xie X. An ARIMA Model With Adaptive Orders for Predicting Blood Glucose Concentrations and Hypoglycemia. IEEE J Biomed Health Inform. 2019;23(3):1251-60.
3. Sun Q, Jankovic MV, Bally L, Mougiakakou SG. Predicting Blood Glucose with an LSTM and Bi-LSTM Based Deep Neural Network. arXiv e-prints [Internet]. 2018 September 01, 2018. Available from: https://ui.adsabs.harvard.edu/abs/2018arXiv180903817S.
4. Rodriguez-Rodriguez I, Chatzigiannakis I, Rodriguez JV, Maranghi M, Gentili M, Zamora-Izquierdo MA. Utility of Big Data in Predicting Short-Term Blood Glucose Levels in Type 1 Diabetes Mellitus Through Machine Learning Techniques. Sensors (Basel). 2019;19(20).
5. Georga EI, Protopappas VC, Ardigo D, Polyzos D, Fotiadis DI. A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions. Diabetes Technol Ther. 2013;15(8):634-43.
6. Hamdi T, Ben Ali J, Di Costanzo V, Fnaiech F, Moreau E, Ginoux J-M. Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm. Biocybernetics and Biomedical Engineering. 2018;38(2):362-72.
7. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 2017:3146--54.
8. Allam F, Nossair Z, Gomma H, Ibrahim I, Salam MA-e, editors. Prediction of subcutaneous glucose concentration for type-1 diabetic patients using a feed forward neural network. The 2011 International Conference on Computer Engineering & Systems; 2011 29 Nov.-1 Dec. 2011.
9. Zarkogianni K, Mitsis K, Litsa E, Arredondo MT, Ficomicron G, Fioravanti A, et al. Comparative assessment of glucose prediction models for patients with type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring. Med Biol Eng Comput. 2015;53(12):1333-43.
10. Perez-Gandia C, Facchinetti A, Sparacino G, Cobelli C, Gomez EJ, Rigla M, et al. Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring. Diabetes Technol Ther. 2010;12(1):81-8.
11. Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput. 1997;9(8):1735-80.
12. Sherstinsky A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network. arXiv e-prints. 2018:arXiv:1808.03314.
13. Staudemeyer RC, Rothstein Morris E. Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks. arXiv e-prints [Internet]. 2019 September 01, 2019. Available from: https://ui.adsabs.harvard.edu/abs/2019arXiv190909586S.
14. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. arXiv e-prints. 2014:arXiv:1412.3555.
15. Schuster M, Paliwal K. Bidirectional recurrent neural networks. Signal Processing, IEEE Transactions on. 1997;45:2673-81.
16. Li K, Liu C, Zhu T, Herrero P, Georgiou P. GluNet: A Deep Learning Framework for Accurate Glucose Forecasting. IEEE J Biomed Health Inform. 2020;24(2):414-23.

3

# Chapter 4

## Towards precision medicine in diabetes? A critical review of glucotypes

Adam Hulman, Yuri D. Foreman, Martijn C.G.J. Brouwers, Abraham A. Kroon,
Koen D. Reesink, Pieter C. Dagnelie, Carla J.H. van der Kallen,
Marleen M.J. van Greevenbroek, Kristine Færch, Dorte Vistisen,
Marit E. Jørgensen, Coen D.A. Stehouwer[†], and Daniel R. Witte[†]

[†] last authors contributed equally

# Abstract

In response to a study previously published in PLoS Biology, this Formal Comment thoroughly examines the concept of 'glucotypes' with regard to its generalizability, interpretability and relationship to more traditional measures used to describe data from continuous glucose monitoring.

Although the promise of precision medicine has led to advances in the recognition and treatment of rare monogenic forms of diabetes, its impact on prevention and treatment of more common forms of diabetes has been underwhelming[1]. Several approaches to the subclassification of individuals with, or at high risk of, type 2 diabetes have been published recently[2-4]. Hall and colleagues introduced the concept of 'glucotypes' in a research article[3] that has received enormous attention in the top high-impact scientific journals[5-8], mostly in relation to precision medicine. The authors developed an algorithm to identify patterns of glucose fluctuations based on continuous glucose monitoring (CGM). They named the three identified patterns 'low variability', 'moderate variability', and 'severe variability' glucotypes. Each individual was characterized by the proportion of time spent in the three glucotypes and was assigned to an overall glucotype based on the highest proportion. They argued that glucotypes provide the advantage of taking into account a more detailed picture of glucose dynamics, in contrast to commonly used single time point or average-based measures, revealing subphenotypes within traditional diagnostic categories of glucose regulation. Even though the study was based on data from only 57 individuals without a prior diabetes diagnosis, others have interpreted the results as indicating that glucotypes might identify individuals at an early stage of glucose dysregulation, suggesting a potential role in diabetes risk stratification and prevention[5]. However, before glucotypes can become 'an important tool in early identification of those at risk for type 2 diabetes'[3], the concept requires thorough validation. Therefore, we explore the generalizability and interpretability of glucotypes, and their relationship to traditional CGM-based measures.

We used data from The Maastricht Study[9, 10] and the PRE-D Trial[11] comprising a total number of 770 diabetes-free individuals with a seven-day CGM registration. We observed that the average proportion of time spent in the low variability glucotype was low both in The Maastricht Study (6%) and the PRE-D Trial (4%), compared to 20% in the original study. A reason for the difference may be that our study populations were on average 11-12 years older, and that the PRE-D Trial (n=116) included only overweight and obese individuals with prediabetes. In The Maastricht Study, the median (interquartile range) body mass index was 25.9 kg/m$^2$ (23.4-28.7) and 72% had normal glucose tolerance. As a logical consequence, the severe glucotype was most common in the PRE-D Trial (55%). Regardless, our data show that the initial estimates of the different glucotype prevalences do not necessarily generalize to other populations, especially in age groups at increased risk of type 2 diabetes.

Hall and colleagues described glucotypes as a new measure of glucose variability, a clinically relevant metric of glycemic patterns[3]. In the figures that accompanied the original publication, the low variability pattern was characterized by both the lowest mean glucose level and variation, while the severe pattern had both the highest mean glucose level and variation. As such, these examples did not give an intuition whether

glucotypes were predominantly driven by glucose variability or by mean glucose levels. We therefore present three examples from the PRE-D trial (**Figure 4.1**).
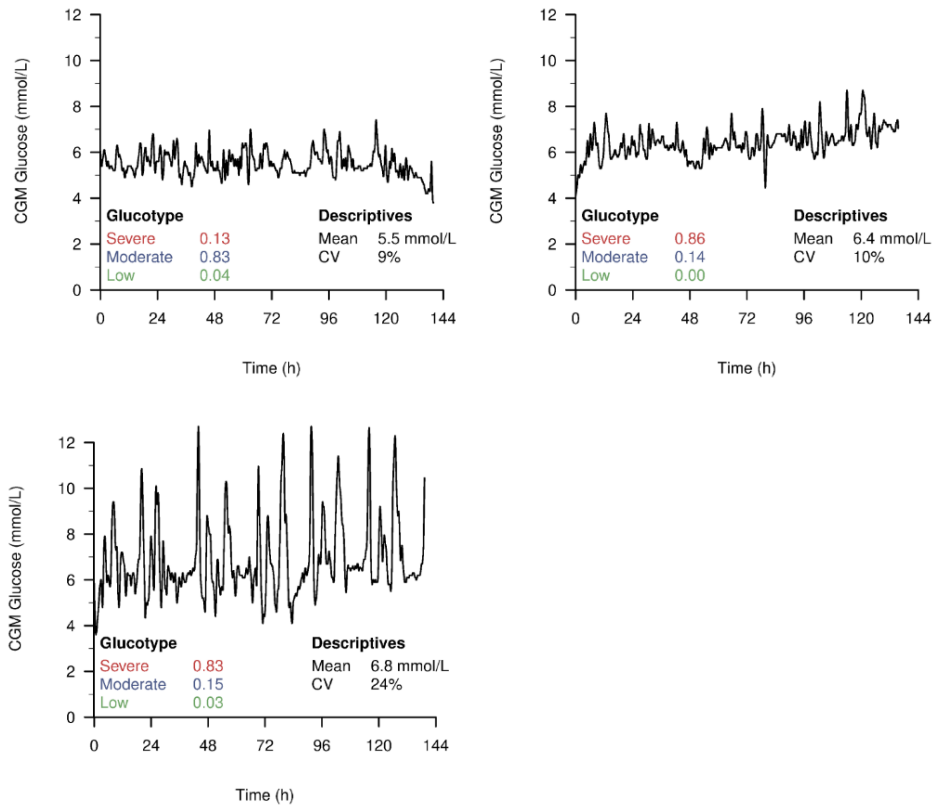


Figure 4.1    Example CGM profiles of participants in the PRE-D Trial with corresponding proportion of time spent in different glucotypes and conventional measures (mean CGM glucose and CV). CGM, continuous glucose monitoring; CV, coefficient of variation

The first two profiles are very similar with regard to glucose variability. Thus, the driver of the most severe glucotype of the second participant is clearly the slightly higher mean glycemic level. Also, even though the third participant has a much larger variation than the first two, the proportion of time in the severe glucotype is not higher than for the second participant as one would expect from a classical measure of glucose variability. To investigate this further, we assessed the association between glucotypes and classical CGM measures, i.e., the mean CGM glucose level (**Figure 4.2A**) and the coefficient of variation (**Figure 4.2B**) in The Maastricht Study. The scatterplots

show a clear association between the mean CGM glucose and glucotypes. They also suggest that participants with a high proportion of time in the moderate glucotype do not have high variation in glucose. Rather than a biological feature, this may well be a methodological consequence of being assigned to the middle cluster. If large fluctuations were present, glucose levels would reach either low or high values, resulting in a higher proportion of time spent in the low or severe glucotypes respectively (assuming a strong association between glucotypes and mean CGM glucose). Therefore, we decided to quantify this association using regression analysis where glucotype proportions were the outcomes and the mean CGM glucose concentration was the independent variable modelled with natural cubic splines (more details on the specification of the models are given in the supporting information S1 Code, S2 Code, S3 Code). Then, we used the equation estimated in The Maastricht Study to predict glucotypes in the external validation sample (PRE-D Trial, **Figure 4.2C**). First, similarly to Hall and colleagues, we assigned individuals to the pattern with the highest proportion of time and then compared the predicted and the observed glucotypes. We found that in 107 out of 116 individuals, the glucotype was predicted correctly when using only the mean CGM glucose value. When the glucotypes were considered as continuous proportions of time, the root mean squared errors (RMSE) were 0.05, 0.09 and 0.07 for the low, moderate and severe variability glucotypes, respectively, which indicates good predictive ability. These results demonstrate that glucotypes either mainly reflect the mean CGM glucose level or do not translate to external datasets (e.g., due to overfitting).

To investigate this further, we conducted the same analyses as described for the PRE-D trial in the original data from Hall et al. and found a slightly weaker, but still strong association between mean CGM glucose levels and glucotypes. Using the regression model from The Maastricht Study, we could correctly predict 79% of the glucotypes, while the RMSEs were 0.11, 0.15 and 0.13. Scatterplots are shown in the electronic supplement.

Although the transformation of continuous measures into categorical ones is a common procedure in clinical research, assigning individuals to the glucotype with the highest proportion of time runs very much against the 'precision' tenet of precision medicine. In line with this, a recent study has demonstrated how simple clinical features outperformed clusters in predicting relevant clinical outcomes[12]. This is especially problematic when a method does not provide clear separation between clusters, which can be quantified by calculating relative entropy[13]. A relative entropy of zero would mean that all individuals spend one third of the time in each of the three glucotypes, while a value of one would indicate that each individual spends the entire time period in only one of the three glucotypes. In the original cohort of Hall et al.[3], we calculated a relative entropy of 0.24, which indicates that cluster separation is far from optimal and—together with the previous results—questions the claim that the

glucotype is really a 'more comprehensive measure of the pattern of glucose excursions than the standard laboratory tests in current use'[3].
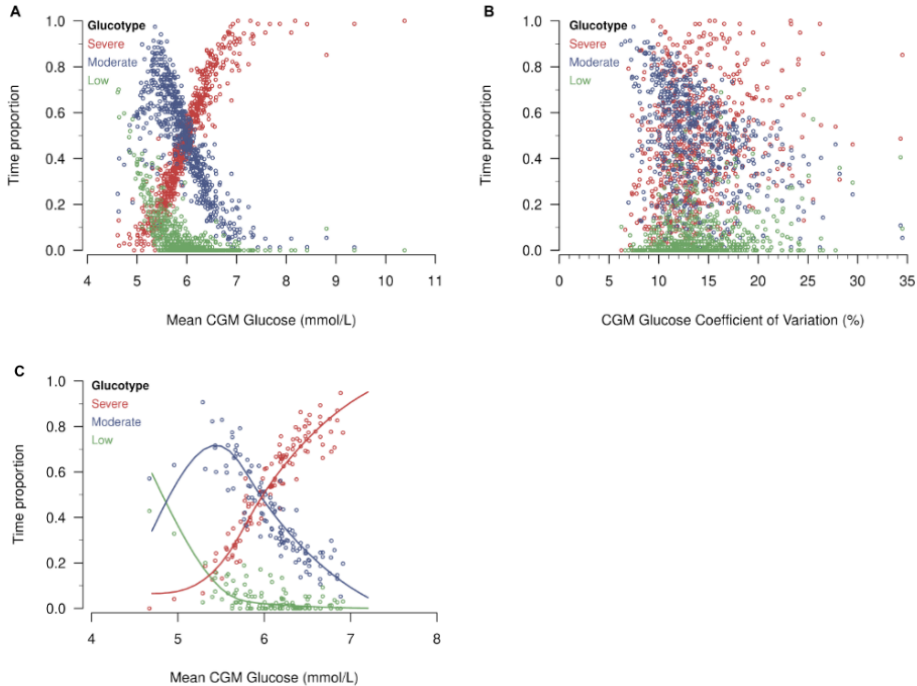


**Figure 4.2** Observed proportion of time spent in the three glucotypes by mean CGM glucose (panel A) and coefficient of variation (panel B) in The Maastricht Study, and by mean CGM glucose in the PRE-D Trial (panel C) alongside predicted proportions based on the regression analysis in The Maastricht Study. CGM, continuous glucose monitoring

In conclusion, we demonstrate in two large, external datasets that the assessment of glucotypes does not offer more novel insights than the mean CGM glucose, highlighting the importance of large development datasets and external validation for data-driven algorithms. As CGM is becoming more widely used in large clinical studies (also among individuals without diabetes), glucose patterns derived from CGMs will be an important focus area in future diabetes research. However, it is important that scientific scrutiny precedes the introduction of emerging tools with a promise of identifying individuals at high risk of type 2 diabetes and its late complications at an earlier stage of disease progression, especially in an observational setting. Furthermore, future efforts towards precision medicine for diabetes prevention and treatment should go beyond the glucocentric approach we have seen so far. We know

that hyperglycemia is a late feature of diabetes development and that patients benefit most from a multifactorial treatment approach[14]. A multifactorial approach, with relevance to the etiology of micro- and macrovascular complications, may also yield a more clinically useful risk stratification of non-diabetic individuals[15]. Even so, if we aim for precision medicine, we should strive to retain as much precision as possible—at every step of the process—by treating determinants and outcomes as continuous measures, if possible, and by retaining information on the uncertainty of any hard classification such as cluster membership.

4

# References

1.  Fitipaldi H, McCarthy MI, Florez JC, Franks PW. A Global Overview of Precision Medicine in Type 2 Diabetes. Diabetes. 2018;67(10):1911-22.
2.  Ahlqvist E, Storm P, Karajamaki A, Martinell M, Dorkhan M, Carlsson A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol. 2018;6(5):361-9.
3.  Hall H, Perelman D, Breschi A, Limcaoco P, Kellogg R, McLaughlin T, et al. Glucotypes reveal new patterns of glucose dysregulation. PLoS Biol. 2018;16(7):e2005143.
4.  Hulman A, Vistisen D, Glumer C, Bergman M, Witte DR, Faerch K. Glucose patterns during an oral glucose tolerance test and associations with future diabetes, cardiovascular disease and all-cause mortality rate. Diabetologia. 2018;61(1):101-7.
5.  Zaccardi F, Khunti K. Glucose dysregulation phenotypes - time to improve outcomes. Nat Rev Endocrinol. 2018;14(11):632-3.
6.  Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
7.  Hamideh D, Arellano B, Topol EJ, Steinhubl SR. Your digital nutritionist. Lancet. 2019;393(10166):19.
8.  Li J, Li X, Zhang S, Snyder M. Gene-Environment Interaction in the Era of Precision Medicine. Cell. 2019;177(1):38-44.
9.  Foreman YD, Brouwers MCGJ, van der Kallen CJH, Pagen DME, van Greevenbroek MMJ, Henry RMA, et al. Glucose Variability Assessed with Continuous Glucose Monitoring: Reliability, Reference Values, and Correlations with Established Glycemic Indices-The Maastricht Study. Diabetes Technol Ther. 2020;22(5):395-403.
10. Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.
11. Faerch K, Amadid H, Nielsen LB, Ried-Larsen M, Karstoft K, Persson F, et al. Protocol for a randomised controlled trial of the effect of dapagliflozin, metformin and exercise on glycaemic variability, body composition and cardiovascular risk in prediabetes (the PRE-D Trial). BMJ Open. 2017;7(5):e013802.
12. Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT. Clusters provide a better holistic view of type 2 diabetes than simple clinical features &#x2013; Authors' reply. The Lancet Diabetes & Endocrinology. 2019;7(9):669.
13. Lennon H, Kelly S, Sperrin M, Buchan I, Cross AJ, Leitzmann M, et al. Framework to construct and interpret latent class trajectory modelling. BMJ Open. 2018;8(7):e020683.
14. Gaede P, Lund-Andersen H, Parving HH, Pedersen O. Effect of a multifactorial intervention on mortality in type 2 diabetes. N Engl J Med. 2008;358(6):580-91.
15. Vas PRJ, Alberti KG, Edmonds ME. Prediabetes: moving away from a glucocentric definition. Lancet Diabetes Endocrinol. 2017;5(11):848-9.

# Supplemental materials

## *Supplementary methods*

**S1 Code**. This R script demonstrates the assessment of glucotypes in the PRE-D Trial. It uses files from Alessandra Breschi's GitHub page (https://github.com/abreschi/shinySpecClust). These files were accessed and downloaded on 5th of July, 2019. Glucotypes in The Maastricht Study were assessed with the same method (code not shown). The script also includes the code to create **Figure 4.1** displaying individual glucose trajectories. To avoid plotting individual data points, we calculate a 10-minute moving average for the glucose values.

**S2 Code**. The regression model is fitted with glucotypes as outcomes and mean CGM glucose exposure. R code is also available for Figure 4.2 showing the association between glucotypes and mean CGM glucose (Figure 4.2A) and the coefficient of variation (Figure 4.2B).

**S3 Code**. We are using the regression model developed in The Maastricht Study cohort (see S2 Code) to predict glucotypes based on mean CGM glucose in the PRE-D Trial. Then, the root mean-squared error and the number of correctly classified individuals are calculated. Also, relative entropy is calculated in the Stanford cohort from the original paper (Hall et al., PLoS Biol, 2018).

4

# Part II

Potential consequences of
daily glucose variability

# Chapter 5

The oral glucose tolerance test-derived incremental glucose peak is associated with greater arterial stiffness and maladaptive arterial remodeling – The Maastricht Study

Yuri D. Foreman, Martijn C.G.J. Brouwers, Tos T.J.M. Berendschot,
Martien C.J.M. van Dongen, Simone J.P.M. Eussen,
Marleen M.J. van Greevenbroek, Ronald M.A. Henry, Alfons J.H.M. Houben,
Carla J.H. van der Kallen,  Abraham A. Kroon, Koen D. Reesink,
Miranda T. Schram, Nicolaas C. Schaper, and Coen D.A. Stehouwer

# Abstract

## Background
Daily glucose variability may contribute to vascular complication development irrespective of mean glucose values. The incremental glucose peak (IGP) during an oral glucose tolerance test (OGTT) can be used as a proxy of glucose variability. We investigated the association of IGP with arterial stiffness, arterial remodeling, and microvascular function, independent of HbA1c and other confounders.

## Methods
IGP was calculated as the peak minus baseline plasma glucose value during a seven-point OGTT in 2,758 participants (age: 60±8years; 48% women) of The Maastricht Study, an observational population-based cohort. We assessed the cross-sectional associations between IGP and arterial stiffness (carotid-femoral pulse wave velocity [cf-PWV], carotid distensibility coefficient [carDC]); arterial remodeling (carotid intima-media thickness [cIMT]; mean [$CWS_{mean}$] and pulsatile [$CWS_{puls}$] circumferential wall stress), and microvascular function (retinal arteriolar average dilatation; heat-induced skin hyperemia) via multiple linear regression with adjustment for age, sex, HbA1c, cardiovascular risk factors, lifestyle factors, and medication use.

## Results
Higher IGP was independently associated with higher cf-PWV (regression coefficient [B]: 0.054 m/s [0.020; 0.089]) and with higher $CWS_{mean}$ (B: 0.227 kPa [0.008; 0.446]). IGP was not independently associated with carDC (B: -0.026 $10^{-3}$/kPa [-0.112; 0.060]), cIMT (B: -2.745 μm [-5.736; 0.245]), $CWS_{puls}$ (B: 0.108 kPa [-0.054; 0.270]), retinal arteriolar average dilatation (B: -0.022 % [-0.087; 0.043]), or heat-induced skin hyperemia (B: -1.380% [-22.273; 19.513]).

## Conclusions
IGP was independently associated with aortic stiffness and maladaptive carotid remodeling, but not with carotid stiffness, cIMT, and microvascular function measures. Future studies should investigate whether glucose variability is associated with cardiovascular disease.

# Background

Chronic hyperglycemia is a key factor in the development of type 2 diabetes-related macrovascular and microvascular complications[1,2]. In the macrovasculature, elevated mean blood glucose levels contribute to arterial stiffening[3,4], atherosclerosis[1], and large artery endothelial dysfunction[5]. In the microvasculature, hyperglycemia and endothelial dysfunction are considered to be bidirectionally related, potentially entering a vicious cycle that could lead to microvascular complications[6]. Of note, these pathophysiologic processes have been shown to already occur in the prediabetic state[7,8].

Importantly, chronic hyperglycemia per se does not fully explain the incidence of complications[9]. Daily glucose variability could play a role in vascular complication development irrespective of mean glucose values[10]. While relatively small observational studies have found conflicting results regarding the association between glucose variability and classic diabetic complications[11-13], experimental studies have shown that greater glucose variability can be harmful independent of mean glucose values[14,15].

Continuous glucose monitoring, the gold standard for glucose variability assessment[16], is a challenging technology to use in a large epidemiological setting. The incremental glucose peak (IGP), i.e. the glucose increase from baseline during an oral glucose tolerance test (OGTT), can be used as an index of glucose variability[17]. In view of the aforementioned, we investigated, in a large population-based cohort, whether IGP is associated with arterial stiffness, arterial remodeling, and microvascular function, independent of $HbA_{1c}$.

# Methods

## Study population and design

We used data from The Maastricht Study, an observational prospective population-based cohort study. The rationale and methodology have been described previously[18]. In brief, The Maastricht Study focuses on the etiology, pathophysiology, complications and comorbidities of type 2 diabetes, and is characterized by an extensive phenotyping approach. All individuals aged between 40 and 75 years and living in the southern part of the Netherlands were eligible for participation. We recruited participants through mass media campaigns and from the municipal registries and the regional Diabetes Patient Registry via mailings. For reasons of efficiency, we stratified recruitment according to known type 2 diabetes status, with an oversampling of individuals with type 2 diabetes. The present report includes cross-sectional data from the first 3,451 participants who completed the baseline survey between November 2010 and September 2013. All examinations were performed within a three-month time window; the OGTT and vascular measurements were performed during different

5

research visits. The Maastricht Study has been approved by the institutional medical ethical committee (NL31329.068.10) and the Minister of Health, Welfare and Sports of the Netherlands (Permit 131088-105234-PG). All participants gave written informed consent.

## Assessment of glucose metabolism status and incremental glucose peak

Participants underwent a standardized 2-hour 75 gram OGTT after fasting overnight to determine glucose metabolism status (GMS), which was defined according to the World Health Organization 2006 criteria as normal glucose metabolism (NGM), impaired fasting glucose, impaired glucose tolerance (combined as prediabetes), or type 2 diabetes[19]. For safety reasons, participants using insulin or with a fasting plasma glucose (FPG) value above 11.0 mmol/L (determined by finger prick) did not undergo the OGTT. For these individuals, we used FPG and information about their diabetes medication to determine GMS. During the OGTT, we took venous blood glucose samples at baseline and 15, 30, 45, 60, 90 and 120 minutes; we calculated IGP by subtracting FPG from the absolute glucose peak (AGP) value.

## Assessment of arterial stiffness, intima-media thickness and circumferential wall stress

The rationale and methodology of the macrovascular measurements have been described previously[20, 21]. We determined carotid-femoral pulse wave velocity (cf-PWV) with the use of applanation tonometry (SphygmoCor, Atcor Medical, Sydney, Australia)[22], and used the median of three consecutive cf-PWV recordings in our analyses.

We measured the left common carotid artery using an ultrasound scanner equipped with a 7.5-MHz linear probe (MyLab 70, Esaote Europe B.V., Maastricht, the Netherlands) to assess local carotid distension, intima-media thickness (cIMT), and interadventitial diameter (IAD)[23]. We quantified local arterial stiffness by calculating the carotid distensibility coefficient (carDC) based on the following formula: carDC = $(2*\Delta D*IAD + \Delta D^2)/(braPP*IAD^2)$, where IAD is interadventitial arterial diameter, $\Delta D$ distension, and braPP brachial pulse pressure[24].

We defined cIMT as the distance between the lumen-intima and media-adventitia interfaces of the far (posterior) wall[23], and IAD as the distance between the media-adventitia interfaces of the near and far wall. The median carDC, cIMT and IAD of three consecutive measurements were used. We calculated carotid lumen diameter (LD) according to the following formula[25]: LD = IAD − (2*cIMT). In parallel with the vascular measurements, we determined mean heart rate (HR) and mean arterial pressure (MAP) every five minutes with an oscillometric device (Accutorr Plus, Datascope Inc., Montvale, NJ, USA). Mean ($CWS_{mean}$) and pulsatile ($CWS_{puls}$) carotid circumferential wall stress were calculated by use of the Lamé equation as $CWS_{mean}$ = (MAP*(LD/2))/cIMT

and $CWS_{puls}$ = (carPP*(LD/2))/cIMT, where carotid pulse pressure (carPP) was obtained from carotid pressure waveform calibration[20].

## Assessment of microvascular function

The rationale and methodology of assessing the microcirculation of the retina and skin have been described previously[8]. In short, we measured the retinal microvascular dilation response to flicker light during a 50-second baseline, 40-second flicker-light provocation, and 60-second recovery phase, by use of the Dynamic Vessel Analyzer (DVA; Imedos, Jena, Germany). The integrated DVA software (version 4.51; Imedos) automatically calculated average baseline diameter size (expressed in measurement units; MUs) during the 20-50 seconds of baseline recording, and percentage dilation at time points 10 and 40 seconds during the flicker stimulation period. Two regression lines were drawn (at the 0–10-second and 10–40-second intervals) and averaged to assess average percentage dilation. We measured skin blood flow with a laser-Doppler system (Periflux 5000; Perimed, Järfalla, Sweden) equipped with a thermostatic laser-Doppler probe (PF457; Perimed) at the dorsal side of the left wrist. After a two-minute baseline recording, the probe temperature was rapidly increased to 44 °C and kept constant until the end of the registration. The heat-induced skin hyperemic response was expressed as the percentage increase in average perfusion units (PUs) during the 23-minute heating phase over the average baseline PU.

## Measurement of covariates

As described previously[18], we assessed history of cardiovascular disease (CVD), physical activity, and smoking status (never, former, current) by questionnaire; calculated Mediterranean diet adherence according to Trichopoulou et al. based on a food frequency questionnaire[26]; assessed lipid-modifying, antihypertensive, and glucose-lowering medication use as part of a medication interview; measured weight, height, body mass index (BMI), and waist circumference, during a physical examination; measured office and 24-hour ambulatory blood pressure (BP); measured $HbA_{1c}$, fasting plasma insulin and lipid profile in fasting venous blood samples; quantified insulin resistance (IR) based on the updated Homeostatic Model Assessment (HOMA2-IR); measured albumin excretion in two 24-hour urine collections; calculated the estimated glomerular filtration rate (eGFR) based on both serum creatinine and cystatin C[27]; and assessed retinopathy presence in both eyes via fundus photography.

## Statistical analysis

Normally distributed data are presented as mean and standard deviation (SD), non-normally distributed data as median and interquartile range (IQR), and categorical data as n (%). We used multivariable linear regression to study the associations between IGP and arterial stiffness (cf-PWV, carDC), arterial remodeling (cIMT, $CWS_{mean}$, $CWS_{puls}$), and

microvascular function (retinal arteriolar average dilatation, heat-induced skin hyperemia). Model 1 was the crude model, which included only IGP as a determinant; model 2 was adjusted for age and sex; model 3 was additionally adjusted for HbA$_{1c}$; model 4 was additionally adjusted for MAP (or alternatively office systolic BP; or carPP) and HR in case of cf-PWV only; model 5 was additionally adjusted for cardiovascular risk and lifestyle factors (i.e. BMI, smoking status, physical activity, Mediterranean diet score, antihypertensive and lipid-modifying drug use, fasting triglycerides and total-to-high-density lipoprotein cholesterol levels). The results are presented as: regression coefficient (B) (corresponding 95% confidence interval [CI]), p-value. We considered a p-value of <0.05 statistically significant. To test the robustness of our findings, we performed multiple sensitivity analyses by: 1) additionally adjusting for history of CVD, retinopathy, eGFR, and urinary albumin excretion; 2) additionally adjusting for fasting plasma insulin or HOMA2-IR; 3) replacing HbA$_{1c}$ with GMS or FPG; 4) replacing IGP with AGP or percentage increase from baseline (IGP$_{percentage}$ = IGP/FPG*100%); 5) adjusting for alternative BP measurements (e.g. ambulatory 24-hour systolic BP); and 6) replacing IGP with time to glucose peak. We incorporated interaction terms in the fully adjusted regression models to test for interactions between IGP and sex[28], as well as IGP and age, as previously advocated[29]. We considered a p-value for interaction of <0.10 statistically significant. We performed all statistical analyses with the Statistical Package for Social Sciences (Version 25.0; IBM, Chicago, IL).

## Results

### *Study population characteristics*

The total study population comprised 3,451 individuals, from which we excluded 41 participants with diabetes types other than type 2 diabetes. Some participants had incomplete data on the seven-point OGTT, either because of missing glucose samples (n=368) or an OGTT contraindication (n=238; i.e. insulin use or plasma glucose levels >11.0 mmol/L before initiation of the OGTT), resulting in a study population of 2,804 individuals. Those with missing glucose samples were generally comparable to the final study population (**Supplementary Table S5,1**); as expected, those with an OGTT contraindication differed statistically significantly from the final study population with regard to almost all characteristics (**Supplementary Table S5.1**). Finally, for 46 participants all outcome data was missing. These individuals were similar to the final study population (**Supplementary Table S5.1**), which consisted of 2,758 individuals. Since outcome and covariate data could not be obtained in all individuals (**Supplementary Table S5.2**), the number of individuals included in the different regression analyses varied (n=1,134-1,978) (**Figure 5.1**).
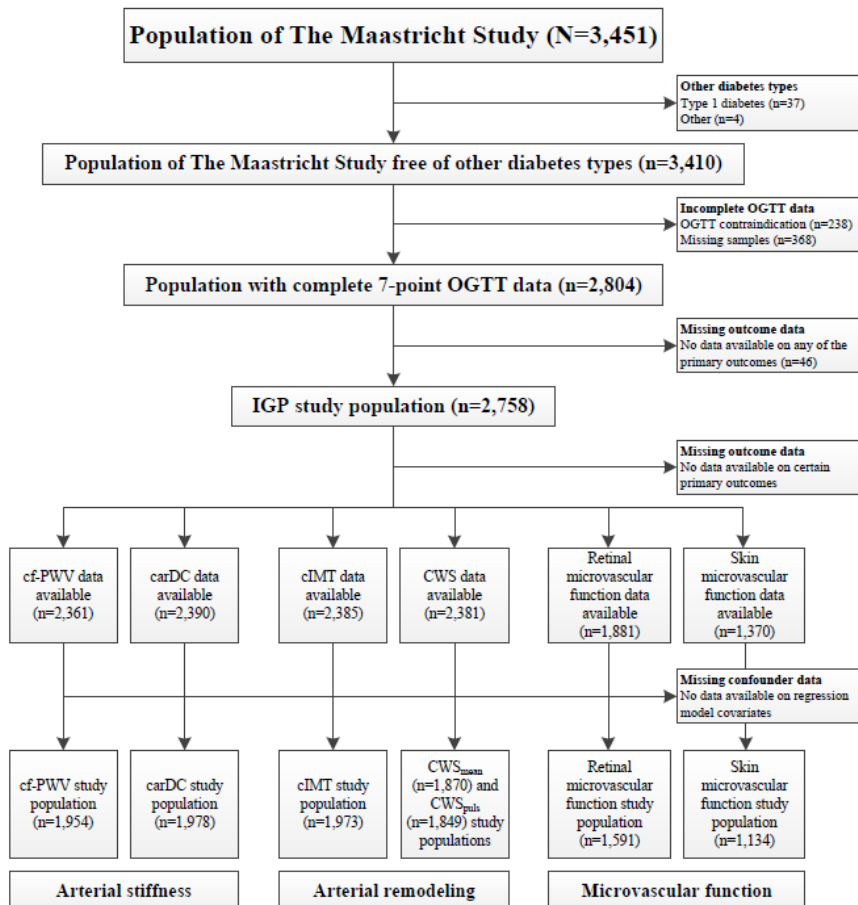
**Figure 5.1** Flowchart of the IGP study population selection process. OGTT, oral glucose tolerance test; IGP, incremental glucose peak; cf-PWV, carotid-femoral pulse wave velocity; carDC, carotid distensibility coefficient; cIMT, carotid intima-media thickness; CWSmean, mean circumferential wall stress; CWSpuls, pulsatile circumferential wall stress.

**Table 5.1** shows the general characteristics of the final study population, stratified according to IGP tertiles. Participants in the highest tertile were older, predominantly male, and had a worse cardiometabolic profile, i.e. higher BMI, waist circumference, systolic BP, and fasting glucose, 2-hour post-load glucose, $HbA_{1c}$, fasting plasma insulin, and triglycerides levels. They were also less physically active, more often smoker, more frequently used lipid-modifying, antihypertensive or glucose-lowering medication, and more often had a history of CVD, decreased eGFR, albuminuria, and retinopathy. Of

note, IGP tertiles did not fully correspond with GMS. Several individuals with type 2 diabetes (n=47; 17 newly diagnosed) were not in the highest IGP tertile. Individuals with prediabetes were distributed equally among the second and third tertile. Heterogeneity existed regarding the OGTT glucose peak time point (**Supplementary Table S5.3**); for the first, second and third IGP tertile, the most frequently occurring time points were 30, 45 and 90 minutes, respectively.

Table 5.1    Participant characteristics according to incremental glucose peak (IGP) tertiles.

| Characteristic | first tertile (n=924) | second tertile (n=909) | third tertile (n=925) |
|---|---|---|---|
| Age, y | 57.3±8.2 | 60.0±8.0 | 62.1±7.7 |
| Women | 569 (61.6) | 431 (47.4) | 330 (35.7) |
| Body mass index, kg/m$^2$ | 25.2±3.6 | 26.4±3.9 | 28.8±4.6 |
| Waist circumference, cm | | | |
| Men | 95.1±9.1 | 98.4±10.0 | 105.2±11.3 |
| Women | 85.4±10.5 | 89.1±11.4 | 96.5±13.9 |
| Office SBP, mmHg | 129.3±16.6 | 134.2±17.3 | 140.2±17.7 |
| Office DBP, mmHg | 74.7±9.5 | 76.7±10.3 | 78.1±9.6 |
| Ambulatory 24-h SBP, mmHg | 116.6±10.9 | 119.0±10.9 | 121.9±12.4 |
| Ambulatory 24-h DBP, mmHg | 73.6±7.0 | 74.5±7.2 | 74.3±7.4 |
| Mean arterial pressure, mmHg | 94.3±10.3 | 96.9±10.4 | 98.8±10.1 |
| Carotid pulse pressure, mmHg | 46.3±14.1 | 49.7±14.7 | 52.3±16.0 |
| Mean heart rate, beats/minute | 60.7±8.1 | 61.8±9.0 | 64.5±10.0 |
| Physical activity, hours/week | 14.0 [9.5-19.0] | 13.5 [8.3-19.0] | 11.5 [7.5-17.4] |
| Mediterranean diet score, (range: 0-9) | 4.6±1.7 | 4.6±1.7 | 4.3±1.6 |
| Smoking | | | |
|   Never/former/current | 368/451/91 | 315/463/125 | 261/511/135 |
|   Never/former/current, % | 40.4/49.6/10.0 | 34.9/51.3/13.8 | 28.8/56.3/14.9 |
| Fasting plasma glucose (FPG), mmol/L | 5.2±0.5 | 5.5±0.7 | 6.8±1.3 |
| 2-hour post-load glucose, mmol/L | 5.1±1.1 | 6.3±1.7 | 12.3±4.6 |
| Glucose metabolism status | | | |
| NGM/prediabetes/type 2 diabetes | 858/53/13 | 672/203/34 | 132/192/601 |
| NGM/prediabetes/type 2 diabetes, % | 92.9/5.7/1.4 | 73.9/22.3/3.7 | 14.3/20.8/65.0 |
| Newly diagnosed type 2 diabetes | 10 (1.1) | 7 (0.8) | 92 (9.9) |
| Incremental glucose peak (IGP), mmol/L | 2.2 [1.8-2.7] | 4.1 [3.6-4.7] | 8.1 [6.5-10.0] |
| HbA$_{1c}$, % | 5.4±0.3 | 5.6±0.4 | 6.3±0.7 |
| HbA$_{1c}$, mmol/mol | 35.8±3.7 | 37.4±4.5 | 45.2±7.8 |
| Fasting plasma insulin, pmol/L | 52.4 [38.3-71.0] | 59.6 [41.7-86.8] | 81.9 [51.5-125.5] |
| HOMA2-IR | 1.2 [0.9-1.5] | 1.3 [1.0-2.0] | 2.0 [1.2-2.9] |
| Triglycerides, mmol/L | 1.0 [0.8-1.4] | 1.2 [0.9-1.6] | 1.5 [1.1-2.1] |
| Total-to-HDL cholesterol ratio | 3.3 [2.8-4.1] | 3.5 [2.9-4.4] | 3.6 [3.0-4.5] |
| Total cholesterol, mmol/L | 5.5±1.0 | 5.5±1.1 | 4.9±1.2 |
| LDL cholesterol, mmol/L | 3.4±0.9 | 3.3±1.0 | 2.8±1.1 |
| HDL cholesterol, mmol/L | 1.7±0.4 | 1.6±0.5 | 1.3±0.4 |
| Lipid-modifying medication use | 145 (15.7) | 219 (24.1) | 558 (60.5) |

Table 5.1     (continued)

| Characteristic | first tertile (n=924) | second tertile (n=909) | third tertile (n=925) |
|---|---|---|---|
| Antihypertensive medication use | 197 (21.3) | 276 (30.4) | 546 (59.2) |
| Diabetes medication use | 1 (0.1) | 21 (2.3) | 449 (48.6) |
| Insulin | 0 (0) | 0 (0) | 0 (0) |
| Metformin | 1 (0.1) | 20 (2.2) | 423 (45.8) |
| Sulfonylureas | 0 (0) | 3 (0.3) | 149 (16.1) |
| Thiazolidinediones | 0 (0) | 1 (0.1) | 8 (0.9) |
| GLP-1 analogs | 0 (0) | 0 (0) | 6 (0.7) |
| DDP-4 inhibitors | 0 (0) | 1 (0.1) | 50 (5.4) |
| History of CVD | 110 (12.3) | 116 (13.1) | 174 (19.4) |
| eGFR, mL/min/1.73 m$^2$ | 90.7±13.2 | 88.5±13.3 | 86.0±15.6 |
| eGFR < 60 mL/min/1.73 m$^2$ | 14 (1.5) | 20 (2.2) | 58 (6.3) |
| (Micro)albuminuria | 31 (3.4) | 49 (5.4) | 120 (13.0) |
| Retinopathy | 0 (0) | 2 (0.2) | 14 (1.6) |
| Carotid-femoral pulse wave velocity (cf-PWV), m/s | 8.3±1.7 | 8.8±1.9 | 9.7±2.2 |
| Carotid distensibility coefficient (carDC), 10$^{-3}$/kPa | 15.6±5.4 | 14.4±5.1 | 13.2±4.8 |
| Carotid intima-media thickness (cIMT), μm | 846.3±150.3 | 854.4±155.8 | 876.1±161.1 |
| Mean circumferential wall stress (CWS$_{mean}$), kPa | 43.8 [37.8-50.9] | 46.5 [40.8-53.0] | 47.5 [41.0-56.0] |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa | 20.9 [16.3-26.4] | 23.1 [18.6-28.9] | 24.2 [19.0-31.3] |
| Retinal arteriolar average dilatation, % | 3.1 [1.1-5.3] | 2.8 [1.1-5.2] | 2.1 [0.5-4.4] |
| Heat-induced skin hyperemia, % | 1,110.5 [666.3-1,592.3] | 1,027.6 [633.3-1,587.3] | 868.6 [521.3-1,318.0] |

Data are reported as mean ± SD, median [interquartile range], or number (percentage %) as appropriate. Data represent the study population of participants with complete oral glucose tolerance test data and results of at least one primary outcome. CVD, cardiovascular disease; SBP, systolic blood pressure; DBP, diastolic blood pressure; NGM, normal glucose metabolism; HbA$_{1c}$, glycated hemoglobin A$_{1c}$; HOMA2-IR, updated homeostasis model assessment; HDL, high-density lipoprotein; LDL, low-density lipoprotein; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4; eGFR, estimated glomerular filtration rate.

## Incremental glucose peak and arterial stiffness

**Figure 5.2** (panels A-B) and **Supplementary Table S5.4** show the associations of IGP with cf-PWV and carDC. Higher IGP was statistically significantly associated with higher cf-PWV in the crude analysis. This association persisted after adjustment for age, sex, and HbA$_{1c}$ (model 3), and additional adjustment for MAP and HR (model 4). After further adjustment for cardiovascular risk and lifestyle factors, the association of IGP with cf-PWV remained statistically significant (model 5, B: 0.054 m/s [0.020; 0.089], p=0.002).

Higher IGP was statistically significantly associated with lower carDC in the crude analysis. This association persisted after adjustment for age, sex, and HbA$_{1c}$ (model 3). The association did not remain statistically significant after adjustment for MAP, and cardiovascular risk factors and lifestyle factors (model 5, B: -0.026 10$^{-3}$/kPa [-0.112; 0.060], p=0.551).
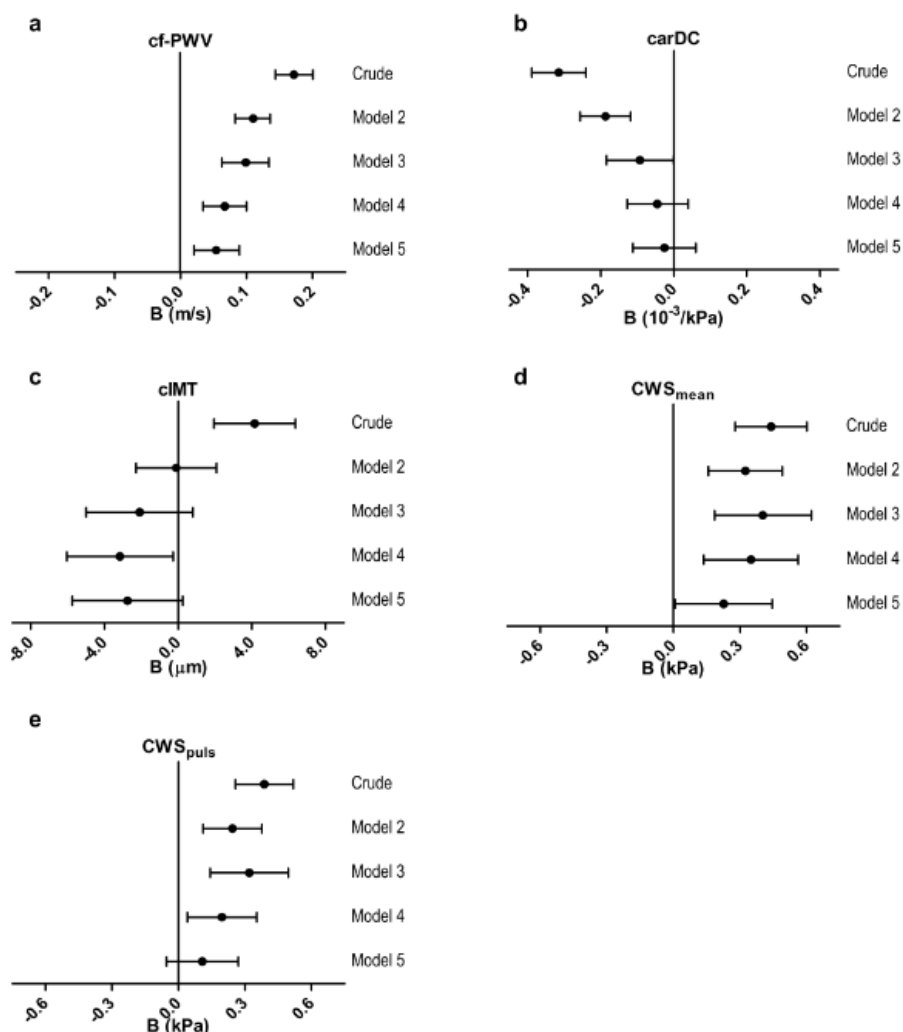
**Figure 5.2 Multivariable-adjusted associations of incremental glucose peak (IGP) and arterial stiffness and arterial remodeling.** Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. The panels depict the (a) associations between IGP and carotid-femoral pulse wave velocity (cf-PWV); (b) associations between IGP and carotid distensibility coefficient (carDC); (c) associations between IGP and carotid intima-media thickness (cIMT); (d) associations between IGP and mean circumferential wall stress (CWS$_{mean}$); (e) associations between IGP and pulsatile circumferential wall stress (CWS$_{puls}$). Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for HbA$_{1c}$. Model 4: additionally adjusted for mean arterial pressure and mean heart rate (cf-PWV), mean arterial pressure (carDC, CWS$_{puls}$), office systolic blood pressure (cIMT) or carotid pulse pressure (CWS$_{mean}$). Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

### Incremental glucose peak, intima-media thickness and circumferential wall stress

**Figure 5.2** (panels C-E) and **Supplementary Table S5.4** show the associations of IGP with cIMT, $CWS_{mean}$ and $CWS_{puls}$. There was a statistically significant, positive association between IGP and cIMT in the crude analysis (crude, B: 4.157 μm [1.944; 6.370], p<0.001). Of note, this association became negative after correction for age, sex, and $HbA_{1c}$ (model 3). Conversely, $HbA_{1c}$ was positively associated with cIMT (data not shown). Higher IGP was not statistically significantly associated with lower cIMT in the fully adjusted model (model 5, B: -2.745 μm [-5.736; 0.245], p=0.072).

Higher IGP was associated with higher $CWS_{mean}$ in the crude model, and remained associated after adjustment for age, sex, and $HbA_{1c}$ (model 3). The association between IGP and $CWS_{mean}$ remained statistically significant after further correction for cardiovascular risk factors and lifestyle factors (model 5, B: 0.227 kPa [0.008; 0.446], p=0.043). IGP was positively associated with $CWS_{puls}$ in the crude model, and after additional adjustment for age, sex, and $HbA_{1c}$ (model 3). The association between IGP and $CWS_{puls}$ did not remain statistically significant after correction for cardiovascular risk and lifestyle factors (model 5, B: 0.108 kPa [-0.054; 0.270], p=0.192).

### Incremental glucose peak and microvascular function

IGP was not associated with retinal arteriolar baseline diameter or skin baseline blood flow (**Table 5.2**). Higher IGP was statistically significantly associated with lower retinal arteriolar average dilatation and lower heat-induced skin hyperemia (crude models, **Table 5.2**). These associations did not remain statistically significant after adjustment for age, sex, and $HbA_{1c}$ (retinal arteriolar average dilatation), and age and sex (heat-induced skin hyperemia).

### Additional analyses

Additional adjustment for history of CVD, retinopathy, eGFR, and urinary albumin excretion did not materially alter the results (**Supplementary Table S5.5** and **S5.6**), although statistical significance was not retained in the associations of IGP with $CWS_{mean}$ (model 6). Additional adjustment for fasting plasma insulin or HOMA2-IR did not materially affect the results (**Supplementary Table S5.7**). The GMS- or FPG-adjusted models yielded results that were mostly comparable with the main (i.e. $HbA_{1c}$-adjusted) models (**Supplementary Table S5.8**). Differences from the main models with regard to statistical significance were observed for the GMS-adjusted association of IGP and cf-PWV (5a, B: 0.031 m/s [-0.008; 0.071], p=0.120), the GMS-adjusted association of IGP and cIMT (5a, B: -4.282 μm [-7.706; -0.857], p=0.014), and the FPG-adjusted association of IGP and $CWS_{mean}$ (5b, B: 0.156 kPa [-0.065; 0.377], p=0.167). In general, the results were not materially different when AGP or $IGP_{percentage}$ were used as

**5**

determinant instead of IGP (**Supplementary Tables S5.9** and **S5.10**). Adjustment for alternative BP measurements did not materially affect the results either (**Supplementary Tables S5.11, S5.12** and **S5.13**). When time to glucose peak was used as determinant instead of IGP, only cf-PWV was found to be statistically significantly associated (B: 0.005 m/s [0.001; 0.008], p-value=0.007) (**Supplementary Table S5.14**).

The association between IGP and cf-PWV was stronger with higher age (p-value for interaction <0.001; **Supplementary Table S5.15**). The association between IGP and carDC was weaker with higher age (p-value for interaction <0.001; **Supplementary Table S5.16**). Age statistically significantly modified the association between IGP and $CWS_{puls}$ (p-value for interaction = 0.013; **Supplementary Table S5.17**). Sex did not modify the associations of IGP with arterial stiffness, arterial remodeling, or microvascular function.

**Table 5.2** Multivariable-adjusted associations of incremental glucose peak (IGP) and microvascular function.

| Model | B (95%CI) | p-value |
|---|---|---|
| Retinal arteriolar baseline diameter, MU (n=1,591) | | |
| Crude | -0.035 (-0.295; 0.225) | 0.792 |
| Model 2 | 0.068 (-0.204; 0.339) | 0.626 |
| Model 3 | -0.145 (-0.503; 0.213) | 0.428 |
| Model 4 | -0.092 (-0.451; 0.267) | 0.614 |
| Model 5 | -0.157 (-0.528; 0.214) | 0.406 |
| Retinal arteriolar average dilatation, % (n=1,591) | | |
| Crude | -0.088 (-0.134; -0.043) | < 0.001 |
| Model 2 | -0.073 (-0.121; -0.026) | 0.002 |
| Model 3 | -0.038 (-0.101; 0.024) | 0.229 |
| Model 4 | -0.042 (-0.105; 0.020) | 0.184 |
| Model 5 | -0.022 (-0.087; 0.043) | 0.506 |
| Skin baseline blood flow, PU (n=1,134) | | |
| Crude | 0.016 (-0.110; 0.142) | 0.799 |
| Model 2 | -0.024 (-0.155; 0.107) | 0.722 |
| Model 3 | 0.025 (-0.149; 0.198) | 0.780 |
| Model 4 | 0.049 (-0.126; 0.224) | 0.581 |
| Model 5 | 0.065 (-0.117; 0.246) | 0.485 |
| Heat-induced skin hyperemia, % (n=1,134) | | |
| Crude | -28.109 (-42.778; -13.440) | < 0.001 |
| Model 2 | -12.503 (-27.509; 2.504) | 0.102 |
| Model 3 | -3.311 (-23.208; 16.586) | 0.744 |
| Model 4 | -5.332 (-25.420; 14.756) | 0.603 |
| Model 5 | -1.380 (-22.273; 19.513) | 0.897 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for office systolic blood pressure. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

## Discussion

In the present study, we investigated cross-sectional associations of IGP with arterial stiffness, arterial remodeling and microvascular function. Our study has two main findings. First, higher IGP was independently associated with higher aortic stiffness (cf-PWV) and higher $CWS_{mean}$, but not with carotid stiffness (carDC), cIMT and $CWS_{puls}$. Second, IGP was not independently associated with measures of microvascular function.

Our study shows that IGP measured during an OGTT provides additional information on top of established glycemic indices (i.e. $HbA_{1c}$, GMS, and FPG). We found that IGP not fully corresponds with GMS, as individuals with prediabetes were equally distributed among the second and third IGP tertile (**Table 5.1**). Furthermore, we showed that the associations of IGP with cf-PWV and $CWS_{mean}$ were independent of $HbA_{1c}$.

### *Current observations in the perspective of prior research*

This is the first study to report on the association of IGP with arterial stiffness measures. Our findings are in concordance with studies using comparable determinants or outcomes. Hulman et al., for example, showed that individuals with the highest glucose peak during an OGTT were characterized by a worse cardiometabolic risk factor profile (i.e. age, sex, smoking status, BP, plasma lipids)[30]. Moreover, the 1-hour OGTT value has previously been found to be independently associated with cf-PWV[31] and brachial-ankle PWV[29]. Still, in our study IGP was observed at this time point in only 20.8% of the participants (**Supplementary Table S5.3**). Our independent association between time to glucose peak and arterial stiffness (**Supplementary Table S5.14**), is also in line with Hulman et al.'s findings on glucose peak time point and cardiometabolic risk[30]. By contrast, in a study by the same research group no independent association between OGTT glucose peak and incident CVD was found[32]. However, their use of just a three-point OGTT entails a major limitation.

We observed a negative, albeit not statistically significant, association of IGP with cIMT, which is in contrast with two studies that have found a positive association between IGP and cIMT[33,34]. These associations, however, were not adjusted for $HbA_{1c}$ or other glycaemic indices. Adjustment for $HbA_{1c}$, GMS or FPG consistently resulted in a negative association between IGP and cIMT.

The absence of a $HbA_{1c}$-independent association between IGP and measures of microvascular function is in line with current literature to the extent that the association of glucose variability, as assessed by continuous glucose monitoring, with macroalbuminuria disappeared after adjustment for mean sensor glucose[11]. This could imply that mean glucose values, rather than glucose peaks, are an important determinant of microvascular function.

## *Mechanistic explanations*

The biological mechanism underlying the relationship between IGP and aortic stiffness remains to be elucidated. Previous research has shown that the glucose peak during an OGTT correlates well with glucose variability based on a self-determined ten-point home glucose profile[17]. Greater daily glucose variability may lead to greater oxidative stress[14,15], which in turn could lead to advanced glycation end product (AGE) formation[2]. AGEs are thought to induce arterial stiffening by accumulating in the arterial wall and forming cross-links between elastin and collagen[3,35]. Of interest, a previous study by our group showed that AGE precursor levels peaked in parallel with glucose values during an OGTT[36]. This supports the mechanistic concept that the mean glucose (reflected by HbA$_{1c}$) and glucose variability (reflected by IGP) both contribute to arterial stiffness, mediated by AGEs. Alternatively, elevated IGP could be a hallmark of higher IR, which may, just as hyperinsulinemia, cause arterial stiffening[3,37]. Indeed, glucose peak height and time point were associated with higher indices of IR in our study, as recently reported by Wang et al.[38]. However, additional adjustment for fasting plasma insulin or HOMA2-IR did not substantially alter the results (Supplementary Table S5.7).

## *Reflections on unexpected findings*

Our analyses yielded several interesting findings. First, IGP was independently associated with cf-PWV, but not with carDC. This difference could be due to structural differences between the aorta (mixed elastic and muscular) and carotid artery (predominately elastic)[22]. Indeed, while an association of tissue and circulating AGEs with cf-PWV has been reported[39], no link has been established for carDC[40]. Second, the inverse association between IGP and cIMT was unexpected, in particular because we observed a positive association of cIMT with HbA$_{1c}$, and found a statistically significant positive association between IGP and CWS$_{mean}$, which normally should stimulate arterial remodeling to increase arterial wall thickness[25]. Our findings could therefore imply that individuals with high IGP values experience maladaptive arterial remodeling of the carotid artery, as has previously been demonstrated in patients with type 2 diabetes[25]. Alternatively, although the assumptions of linear regression were met and the sensitivity analyses showed comparable results, our findings could still be spurious (type 1 error). Third, IGP was statistically significantly associated with CWS$_{mean}$, but not with CWS$_{puls}$. Based on this finding and on the notion that the regression coefficients decreased more from model 3 to 4 for CWS$_{puls}$ compared to CWS$_{mean}$, we conclude that IGP is more strongly associated with MAP than with carPP, and thus that in our study population IGP corresponds more with mean than pulsatile vascular stress. Fourth, while we observed an independent association of IGP with the macrovasculature, no such association was found with the microvasculature, which might be attributable to sample size differences. Future research should focus on further elucidating this discrepancy. Fifth, the regression coefficients of the HbA$_{1c}$-adjusted models differed in

magnitude from the GMS- and FPG-adjusted models, which were performed as sensitivity analyses. This could be a result of using a categorical (i.e. GMS) instead of a continuous (i.e. HbA$_{1c}$ and FPG) confounder[41]. Still, these associations were not statistically significantly dissimilar, as the 95% CIs strongly overlap and include each other's regression coefficients[42]. The loss of statistical significance after adjustment for additional variables (e.g. history of CVD, eGFR) could be due to smaller sample size or overadjustment bias[43].

## Clinical relevance

Aortic stiffness, as measured by cf-PWV, is an independent determinant of CVD, cardiovascular mortality, and all-cause mortality[22]. We found that after adjustment for HbA$_{1c}$ and all other relevant confounders cf-PWV was 0.054 m/s higher per IGP unit (mmol/L). This corresponds with six months of vascular aging per 1 mmol/L higher IGP[44]. Accordingly, the 5.9 mmol/L difference in median IGP between the first and third IGP tertile reflects a three year vascular aging difference. Our results may imply that, even in case of well-controlled HbA$_{1c}$, the harmful effects of glucose peaks on aortic stiffness are still present. Future studies should investigate whether these findings translate to daily glucose fluctuations. If they are replicated using continuous glucose monitoring data, it would further justify therapeutic interventions that specifically target glucose variability.

## Strengths and limitations

This study has several strengths and limitations. Strengths are 1) the use of multiple, state-of-the-art measurements to study arterial stiffness, arterial remodeling, and microvascular function; 2) the study sample size and the extensive participant characterization, allowing adjustment for a broad array of possible confounders; and 3) the robustness of the results, reflected by the consistency of several sensitivity analyses. Our study had certain limitations. First, we could only calculate the main determinant using one OGTT, which is known for its moderate reproducibility[45]. The consequent random measurement error in IGP may have resulted in underestimated associations (i.e. attenuation bias)[46]. Second, a relatively large number of individuals were excluded due to missing determinant, outcome and/or confounder data (**Figure 5.1**). Still, individuals with missing data were generally comparable to the final study population (**Supplementary Table S5.1**), except for the participants with an OGTT contraindication, who were characterized by a more adverse cardiometabolic profile. The inability to calculate IGP in this relatively unhealthy subgroup might have affected the precision of the associations. Third, the cross-sectional design renders us unable to rule out reverse causality. Arterial stiffness, which has been associated with incident diabetes, could theoretically influence glucose values[47]. Fourth, our study population is mostly Caucasian, which limits the generalizability of our results. Fifth, although the

models were adjusted for a large number of cardiovascular risk and lifestyle factors, residual confounding may still be present.

## Conclusions

We show that higher IGP is independently associated with greater aortic stiffness and maladaptive carotid remodeling, but not with carotid stiffness, cIMT, or microvascular function. Taken together, these findings support the concept that glucose peaks have harmful macrovascular effects, regardless of mean glucose levels. Further research is needed to elucidate how these findings translate to daily glucose fluctuations and to what extent CVD could be prevented by reducing glucose variability.

# References

1. Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev. 2013;93(1):137-88.
2. Brownlee M. The pathobiology of diabetic complications: a unifying mechanism. Diabetes. 2005;54(6):1615-25.
3. Stehouwer CD, Henry RM, Ferreira I. Arterial stiffness in diabetes and the metabolic syndrome: a pathway to cardiovascular disease. Diabetologia. 2008;51(4):527-39.
4. Swoboda PP, Erhayiem B, Kan R, McDiarmid AK, Garg P, Musa TA, et al. Cardiovascular magnetic resonance measures of aortic stiffness in asymptomatic patients with type 2 diabetes: association with glycaemic control and clinical outcomes. Cardiovasc Diabetol. 2018;17(1):35.
5. van Sloten TT, Henry RM, Dekker JM, Nijpels G, Unger T, Schram MT, et al. Endothelial dysfunction plays a key role in increasing cardiovascular risk in type 2 diabetes: the Hoorn study. Hypertension. 2014;64(6):1299-305.
6. Stehouwer CDA. Microvascular Dysfunction and Hyperglycemia: A Vicious Cycle With Widespread Consequences. Diabetes. 2018;67(9):1729-41.
7. Henry RM, Kostense PJ, Spijkerman AM, Dekker JM, Nijpels G, Heine RJ, et al. Arterial stiffness increases with deteriorating glucose tolerance status: the Hoorn Study. Circulation. 2003;107(16):2089-95.
8. Sorensen BM, Houben AJ, Berendschot TT, Schouten JS, Kroon AA, van der Kallen CJ, et al. Prediabetes and Type 2 Diabetes Are Associated With Generalized Microvascular Dysfunction: The Maastricht Study. Circulation. 2016;134(18):1339-52.
9. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. N Engl J Med. 1993;329(14):977-86.
10. Brownlee M, Hirsch IB. Glycemic variability: a hemoglobin A1c-independent risk factor for diabetic complications. JAMA. 2006;295(14):1707-8.
11. Jin SM, Kim TH, Oh S, Baek J, Joung JY, Park SM, et al. Association between the extent of urinary albumin excretion and glycaemic variability indices measured by continuous glucose monitoring. Diabet Med. 2015;32(2):274-9.
12. Xu F, Zhao LH, Su JB, Chen T, Wang XQ, Chen JF, et al. The relationship between glycemic variability and diabetic peripheral neuropathy in type 2 diabetes with well-controlled HbA1c. Diabetol Metab Syndr. 2014;6(1):139.
13. Sartore G, Chilelli NC, Burlina S, Lapolla A. Association between glucose variability as assessed by continuous glucose monitoring (CGM) and diabetic retinopathy in type 1 and type 2 diabetes. Acta Diabetol. 2013;50(3):437-42.
14. Monnier L, Mas E, Ginet C, Michel F, Villon L, Cristol JP, et al. Activation of oxidative stress by acute glucose fluctuations compared with sustained chronic hyperglycemia in patients with type 2 diabetes. JAMA. 2006;295(14):1681-7.
15. Ceriello A, Esposito K, Piconi L, Ihnat MA, Thorpe JE, Testa R, et al. Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in normal and type 2 diabetic patients. Diabetes. 2008;57(5):1349-54.
16. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care. 2017;40(12):1631-40.
17. Meier JJ, Baller B, Menge BA, Gallwitz B, Schmidt WE, Nauck MA. Excess glycaemic excursions after an oral glucose tolerance test compared with a mixed meal challenge and self-measured home glucose profiles: is the OGTT a valid predictor of postprandial hyperglycaemia and vice versa? Diabetes Obes Metab. 2009;11(3):213-22.
18. Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.
19. WHO. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. WHO. 2006.

5

20. Geijselaers SL, Sep SJ, Schram MT, van Boxtel MP, van Sloten TT, Op Het Roodt J, et al. Carotid circumferential wall stress is not associated with cognitive performance among individuals in late middle age: The Maastricht Study. Atherosclerosis. 2018;276:15-22.
21. van Sloten TT, Czernichow S, Houben AJ, Protogerou AD, Henry RM, Muris DM, et al. Association Between Arterial Stiffness and Skin Microvascular Function: The SUVIMAX2 Study and The Maastricht Study. Am J Hypertens. 2015;28(7):868-76.
22. Laurent S, Cockcroft J, Van Bortel L, Boutouyrie P, Giannattasio C, Hayoz D, et al. Expert consensus document on arterial stiffness: methodological issues and clinical applications. Eur Heart J. 2006;27(21):2588-605.
23. Willekes C, Hoeks AP, Bots ML, Brands PJ, Willigers JM, Reneman RS. Evaluation of off-line automated intima-media thickness detection of the common carotid artery based on M-line signal processing. Ultrasound Med Biol. 1999;25(1):57-64.
24. Reneman RS, Meinders JM, Hoeks AP. Non-invasive ultrasound in arterial wall dynamics in humans: what have we learned and what remains to be solved. Eur Heart J. 2005;26(10):960-6.
25. Henry RM, Kostense PJ, Dekker JM, Nijpels G, Heine RJ, Kamp O, et al. Carotid arterial remodeling: a maladaptive phenomenon in type 2 diabetes but not in impaired glucose metabolism: the Hoorn study. Stroke. 2004;35(3):671-6.
26. Trichopoulou A, Orfanos P, Norat T, Bueno-de-Mesquita B, Ocke MC, Peeters PH, et al. Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. BMJ. 2005;330(7498):991.
27. Inker LA, Schmid CH, Tighiouart H, Eckfeldt JH, Feldman HI, Greene T, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med. 2012;367(1):20-9.
28. Schiebinger L, Leopold SS, Miller VM. Editorial policies for sex and gender analysis. Lancet. 2016;388(10062):2841-2.
29. Nakagomi A, Sunami Y, Okada S, Ohno Y, Shoji T, Fujisawa T, et al. Association between 1-h post-load plasma glucose levels and arterial stiffness in normotensive subjects with normal glucose tolerance. Diab Vasc Dis Res. 2018;15(1):39-45.
30. Hulman A, Simmons RK, Vistisen D, Tabak AG, Dekker JM, Alssema M, et al. Heterogeneity in glucose response curves during an oral glucose tolerance test and associated cardiometabolic risk. Endocrine. 2017;55(2):427-34.
31. Sciacqua A, Maio R, Miceli S, Pascale A, Carullo G, Grillo N, et al. Association between one-hour post-load plasma glucose levels and vascular stiffness in essential hypertension. PLoS One. 2012;7(9):e44470.
32. Hulman A, Vistisen D, Glumer C, Bergman M, Witte DR, Faerch K. Glucose patterns during an oral glucose tolerance test and associations with future diabetes, cardiovascular disease and all-cause mortality rate. Diabetologia. 2018;61(1):101-7.
33. Temelkova-Kurktschiev TS, Koehler C, Henkel E, Leonhardt W, Fuecker K, Hanefeld M. Postchallenge plasma glucose and glycemic spikes are more strongly associated with atherosclerosis than fasting glucose or HbA1c level. Diabetes Care. 2000;23(12):1830-4.
34. Hu Y, Liu W, Huang R, Zhang X. Postchallenge plasma glucose excursions, carotid intima-media thickness, and risk factors for atherosclerosis in Chinese population with type 2 diabetes. Atherosclerosis. 2010;210(1):302-6.
35. Prenner SB, Chirinos JA. Arterial stiffness in diabetes mellitus. Atherosclerosis. 2015;238(2):370-9.
36. Maessen DE, Hanssen NM, Scheijen JL, van der Kallen CJ, van Greevenbroek MM, Stehouwer CD, et al. Post-Glucose Load Plasma alpha-Dicarbonyl Concentrations Are Increased in Individuals With Impaired Glucose Metabolism and Type 2 Diabetes: The CODAM Study. Diabetes Care. 2015;38(5):913-20.
37. Lee SB, Ahn CW, Lee BK, Kang S, Nam JS, You JH, et al. Association between triglyceride glucose index and arterial stiffness in Korean adults. Cardiovasc Diabetol. 2018;17(1):41.
38. Wang X, Zhao X, Zhou R, Gu Y, Zhu X, Tang Z, et al. Delay in glucose peak time during the oral glucose tolerance test as an indicator of insulin resistance and insulin secretion in type 2 diabetes patients. J Diabetes Investig. 2018;9(6):1288-95.

39.  van Eupen MG, Schram MT, van Sloten TT, Scheijen J, Sep SJ, van der Kallen CJ, et al. Skin Autofluorescence and Pentosidine Are Associated With Aortic Stiffening: The Maastricht Study. Hypertension. 2016;68(4):956-63.
40.  Baumann M, Richart T, Sollinger D, Pelisek J, Roos M, Kouznetsova T, et al. Association between carotid diameter and the advanced glycation end product N-epsilon-carboxymethyllysine (CML). Cardiovasc Diabetol. 2009;8:45.
41.  Groenwold RH, Klungel OH, Altman DG, van der Graaf Y, Hoes AW, Moons KG, et al. Adjustment for continuous confounders: an example of how to prevent residual confounding. CMAJ. 2013;185(5):401-6.
42.  Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016;31(4):337-50.
43.  Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009;20(4):488-95.
44.  McEniery CM, Yasmin, Hall IR, Qasem A, Wilkinson IB, Cockcroft JR, et al. Normal vascular aging: differential effects on wave reflection and aortic pulse wave velocity: the Anglo-Cardiff Collaborative Trial (ACCT). J Am Coll Cardiol. 2005;46(9):1753-60.
45.  Mooy JM, Grootenhuis PA, de Vries H, Kostense PJ, Popp-Snijders C, Bouter LM, et al. Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. Diabetologia. 1996;39(3):298-305.
46.  Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. BMJ. 2010;340:c2289.
47.  Muhammad IF, Borne Y, Ostling G, Kennback C, Gottsater M, Persson M, et al. Arterial Stiffness and Incidence of Diabetes: A Population-Based Cohort Study. Diabetes Care. 2017;40(12):1739-45.

**5**

# Supplemental materials

## *Supplementary methods*

*Statistical analysis*

We compared the groups stratified according to missing data reason with the total study population using independent T test in case of normally distributed continuous variables, Mann-Whitney U test in case of non-normally distributed continuous variables, and chi-square analysis in case of categorical variables.

## *Supplementary results*

**Table S5.1  Participant characteristics of the final study population and individuals excluded from the analyses due to missing data.**

| Characteristic | IGP study population (n=2758) | Missing OGTT data, OGTT contraindication (n=238) | p-value* | Missing OGTT data, missing samples (n=368) | p-value* | Missing all outcome data (n=46) | p-value* |
|---|---|---|---|---|---|---|---|
| Age, y | 59.8±8.2 | 62.0±8.0 | < 0.001 | 58.4±8.6 | 0.002 | 59.4±8.9 | NS |
| Women | 1,330 (48.2) | 65 (27.3) | < 0.001 | 238 (64.7) | < 0.001 | 21 (45.7) | NS |
| Body mass index, kg/m$^2$ | 26.8±4.3 | 31.1±5.3 | < 0.001 | 26.7±4.7 | < 0.001 | 27.4±5.0 | NS |
| Waist circumference, cm | | | | | | | |
| Men | 100.4±11.2 | 112.6±14.2 | < 0.001 | 99.6±11.4 | NS | 103.2±14.7 | NS |
| Women | 89.4±12.5 | 104.3±14.8 | < 0.001 | 90.1±13.1 | NS | 90.2±12.2 | NS |
| Office SBP, mmHg | 134.6±17.8 | 145.3±19.4 | < 0.001 | 132.4±18.3 | 0.031 | 133.1±17.0 | NS |
| Office DBP, mmHg | 76.5±9.9 | 75.1±9.6 | 0.033 | 74.9±9.4 | 0.005 | 75.6±9.4 | NS |
| Ambulatory 24-hour SBP, mmHg | 119.1±11.6 | 125.3±12.7 | < 0.001 | 116.4±10.9 | < 0.001 | 118.0±12.9 | NS |
| Ambulatory 24-hour DBP, mmHg | 74.1±7.2 | 72.9±7.0 | 0.018 | 73.1±6.8 | 0.005 | 73.6±7.6 | NS |
| Mean arterial pressure, mmHg | 96.7±10.4 | 97.8±9.1 | NS | 96.3±10.2 | NS | N/A (n=1) | N/A |
| Carotid pulse pressure, mmHg | 49.4±15.1 | 56.3±15.6 | < 0.001 | 48.5±15.4 | NS | N/A (n=2) | N/A |
| Mean heart rate, beats/minute | 62.3±9.2 | 68.4±10.1 | < 0.001 | 62.3±9.0 | NS | N/A (n=1) | N/A |
| Physical activity, hours/week | 13.0 [8.3-18.5] | 10.3 [5.8-15.9] | < 0.001 | 14.3 [8.5-21.1] | 0.052 | 13.8 [9.3-18.0] | NS |
| Mediterranean diet score, (range: 0-9) | 4.5±1.7 | 3.8±1.5 | < 0.001 | 4.5±1.6 | NS | 4.4±1.7 | NS |
| Smoking | | | | | | | |
| Never/former/current | 944/1,425/351 | 54/123/48 | < 0.001 | 146/163/51 | 0.037 | 16/18/10 | NS |
| Never/former/current, % | 34.7/52.4/12.9 | 24.0/54.7/21.3 | | 40.6/45.3/14.2 | | 36.4/40.9/22.7 | |
| Fasting plasma glucose, mmol/L | 5.8±1.2 | 9.4±3.1 | < 0.001 | 5.7±1.2 | 0.087 | 5.7±1.0 | NS |
| 2-hour post-load glucose, mmol/L | 7.9±4.3 | N/A | N/A | 7.7±3.9 | NS | 7.2±4.1 | NS |
| Glucose metabolism status | | | | | | | |
| NGM/prediabetes/type 2 diabetes | 1,662/448/648 | 0/0/238 | < 0.001 | 229/57/82 | < 0.001 | 33/6/7 | NS |
| NGM/prediabetes/type 2 diabetes , % | 60.3/16.2/23.5 | 0/0/100 | | 62.2/15.5/22.3 | | 71.7/13.0/15.2 | |
| Newly diagnosed type 2 diabetes | 109 (4.0) | 1 (0.4) | < 0.001 | 22 (6.0) | 0.081 | 1 (2.2) | NS |
| Incremental glucose peak (IGP), mmol/L | 4.1 [2.7-6.5] | N/A | N/A | N/A | N/A | 3.2 [2.4-5.6] | 0.072 |
| HbA$_{1c}$, % | 5.8±0.6 | 8.0±1.4 | < 0.001 | 5.7±0.6 | NS | 5.7±0.7 | NS |
| HbA$_{1c}$, mmol/mol | 39.5±7.0 | 63.7±14.9 | | 38.9±6.9 | | 38.4±7.2 | |
| Fasting plasma insulin, pmol/L | 61.0 [42.8-93.0] | 67.6 [39.4-116.0] | NS | 59.2 [40.7-83.3] | NS | 61.6 [41.8-114.0] | NS |
| HOMA2-IR | 1.4 [1.0-2.1] | 1.9 [1.2-2.8] | < 0.001 | 1.4 [0.9-1.9] | NS | 1.5 [1.0-2.6] | NS |
| Triglycerides, mmol/L | 1.2 [0.9-1.7] | 1.6 [1.1-2.3] | < 0.001 | 1.2 [0.9-1.7] | NS | 1.2 [1.0-1.7] | NS |

5

**Table S5.1** (continued)

| Characteristic | IGP study population (n=2758) | Missing OGTT data, OGTT contraindication (n=238) | p-value* | Missing OGTT data, missing samples (n=368) | p-value* | Missing all outcome data (n=46) | p-value* |
|---|---|---|---|---|---|---|---|
| Total-to-HDL cholesterol ratio | 3.5 [2.9-4.3] | 3.5 [2.7-4.4] | NS | 3.4 [2.8-4.3] | NS | 3.2 [2.7-4.0] | NS |
| Total cholesterol, mmol/L | 5.3±1.1 | 4.1±0.9 | < 0.001 | 5.4±1.1 | 0.020 | 5.1±1.0 | NS |
| LDL cholesterol, mmol/L | 3.1±1.0 | 2.1±0.7 | < 0.001 | 3.2±1.0 | NS | 3.0±0.9 | NS |
| HDL cholesterol, mmol/L | 1.5±0.5 | 1.2±0.4 | < 0.001 | 1.6±0.5 | 0.003 | 1.6±0.4 | NS |
| Lipid-modifying medication use | 922 (33.4) | 204 (85.7) | < 0.001 | 90 (24.5) | 0.001 | 14 (30.4) | NS |
| Antihypertensive medication use | 1,019 (36.9) | 195 (81.9) | < 0.001 | 122 (33.2) | NS | 19 (41.3) | NS |
| Diabetes medication use | 471 (17.1) | 235 (98.7) | < 0.001 | 54 (14.7) | NS | 5 (10.9) | NS |
| Insulin | 0 (0) | 216 (90.8) | < 0.001 | 0 (0) | N/A | 0 (0) | NS |
| Metformin | 444 (16.1) | 180 (75.6) | < 0.001 | 49 (13.3) | NS | 5 (10.9) | NS |
| Sulfonylureas | 152 (5.5) | 26 (10.9) | 0.002 | 17 (4.6) | NS | 3 (6.5) | NS |
| Thiazolidinediones | 9 (0.3) | 3 (1.3) | 0.064 | 1 (0.3) | NS | 0 (0) | NS |
| GLP-1 analogs | 6 (0.2) | 2 (0.9) | NS | 1 (0.3) | NS | 0 (0) | NS |
| DDP-4 inhibitors | 51 (1.9) | 4 (1.7) | NS | 4 (1.1) | NS | 0 (0) | NS |
| History of CVD | 400 (14.9) | 86 (38.7) | < 0.001 | 60 (16.7) | NS | 6 (13.6) | NS |
| eGFR, mL/min/1.73 m² | 88.4±14.2 | 83.6±20.0 | 0.013 | 86.4±17.6 | NS | 90.6±11.2 | NS |
| eGFR < 60 mL/min/1.73 m² | 92 (3.3) | 34 (14.6) | < 0.001 | 16 (4.6) | NS | 1 (2.2) | NS |
| (Micro)albuminuria | 200 (7.3) | 70 (30.6) | < 0.001 | 18 (5.0) | NS | 3 (7.3) | NS |
| Retinopathy | 16 (0.6) | 27 (12.7) | < 0.001 | 3 (0.9) | NS | 0 (0) | NS |
| Carotid-femoral pulse wave velocity (cf-PWV), m/s | 8.9±2.0 | 10.7±2.7 | < 0.001 | 8.7±2.1 | NS | N/A | N/A |
| Carotid distensibility coefficient (carDC), 10⁻³/kPa | 14.4±5.2 | 13.1±4.5 | < 0.001 | 14.1±5.0 | NS | N/A | N/A |
| Carotid intima-media thickness (cIMT), μm | 858.9±156.2 | 883.2±175.0 | 0.039 | 842.5±156.1 | 0.083 | N/A | N/A |
| Mean circumferential wall stress (CWS$_{mean}$), kPa | 45.9 [39.7-53.4] | 49.4 [41.5-55.9] | < 0.001 | 45.3 [38.9-52.8] | NS | N/A | N/A |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa | 22.8 [17.8-28.9] | 28.1 [21.5-34.7] | < 0.001 | 21.7 [17.5-27.3] | NS | N/A | N/A |
| Retinal arteriolar average dilatation (DVA), % | 2.7 [0.9-5.1] | 1.2 [0.05-2.9] | < 0.001 | 2.9 [1.3-4.9] | NS | N/A | N/A |
| Heat-induced skin hyperemia, % | 999.1 [588.9-1,512.9] | 823.8 [464.5-1,191.3] | < 0.001 | 1,147.0 [643.5-1,690.4] | NS | N/A | N/A |

Data are reported as mean ± SD, median [interquartile range], or number (percentage %) as appropriate. Data represent the study population of participants with complete oral glucose tolerance test data and results of at least one primary outcome. CVD, cardiovascular disease; SBP, systolic blood pressure; DBP, diastolic blood pressure; NGM, normal glucose metabolism; HbA$_{1c}$, glycated hemoglobin A1c; HDL, high-density lipoprotein; LDL, low-density lipoprotein; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4; eGFR, estimated glomerular filtration rate.* p-values >0.10 are presented as non-significant (NS) for the sake of clarity.

**Table S5.2**  Participant characteristics of the final study population and individuals excluded from the analyses due to missing data.

| Characteristic | IGP study population (n=2,758) | Missing data | Excluded due to missing OGTT and outcome values (n=652) | Missing data |
|---|---|---|---|---|
| Age, y | 59.8±8.2 | 0 | 59.8±8.6 | 0 |
| Women | 1,330 (48.2) | 0 | 324 (49.7) | 0 |
| Body mass index, kg/m2 | 26.8±4.3 | 2 | 28.3±5.4 | 1 |
| Waist circumference, cm | | 3 | | 1 |
| Men | 100.4±11.2 | | 106.7±14.6 | |
| Women | 89.4±12.5 | | 93.0±14.5 | |
| Office SBP, mmHg | 134.6±17.8 | 6 | 137.2±19.6 | 0 |
| Office DBP, mmHg | 76.5±9.9 | 2 | 75.0±9.5 | 0 |
| Ambulatory 24-hour SBP, mmHg | 119.1±11.6 | 323 | 119.7±12.4 | 93 |
| Ambulatory 24-hour DBP, mmHg | 74.1±7.2 | 323 | 73.0±6.9 | 93 |
| Mean arterial pressure, mmHg | 96.7±10.4 | 327 | 96.9±9.8 | 124 |
| Carotid pulse pressure, mmHg | 49.4±15.1 | 488 | 51.8±15.9 | 177 |
| Mean heart rate, beats/minute | 62.3±9.2 | 326 | 64.7±9.9 | 124 |
| Physical activity, hours/week | 13.0 [8.3-18.5] | 328 | 12.5 [7.5-18.8] | 120 |
| Mediterranean diet score, (range: 0-9) | 4.5±1.7 | 165 | 4.3±1.6 | 57 |
| Smoking | | 38 | | 23 |
| Never/former/current | 944/1,425/351 | | 216/304/109 | |
| Never/former/current, % | 34.7/52.4/12.9 | | 34.3/48.3/17.3 | |
| Fasting plasma glucose, mmol/L | 5.8±1.2 | 0 | 7.0±2.8 | 4 |
| 2-hour post-load glucose, mmol/L | 7.9±4.3 | 0 | 7.7±3.9 | 260 |
| Glucose metabolism status | | 0 | | 0 |
| NGM/prediabetes/type 2 diabetes | 1,662/448/648 | | 262/63/327 | |
| NGM/prediabetes/type 2 diabetes , % | 60.3/16.2/23.5 | | 40.2/9.7/50.2 | |
| Incremental glucose peak (IGP), mmol/L | 4.1 [2.7-6.5] | 0 | 3.2 [2.4-5.6] | 606 |
| Newly diagnosed type 2 diabetes | 109 (4.0) | 0 | 24 (3.7) | 0 |
| HbA1c, % | 5.8±0.6 | 1 | 6.5±1.4 | 12 |
| HbA1c, mmol/mol | 39.5±7.0 | | 48.1±16.0 | |
| Fasting plasma insulin, pmol/L | 61.0 [42.8-93.0] | 7 | 61.4 [40.5-98.2] | 41 |
| HOMA2-IR | 1.4 [1.0-2.1] | 105 | 1.5 [1.0-2.3] | 93 |
| Triglycerides, mmol/L | 1.2 [0.9-1.7] | 0 | 1.3 [0.9-1.9] | 4 |
| Total-to-HDL cholesterol ratio | 3.5 [2.9-4.3] | 0 | 3.4 [2.8-4.3] | 4 |
| Total cholesterol, mmol/L | 5.3±1.1 | 0 | 4.9±1.2 | 4 |
| LDL cholesterol, mmol/L | 3.1±1.0 | 0 | 2.8±1.0 | 4 |
| HDL cholesterol, mmol/L | 1.5±0.5 | 0 | 1.5±0.5 | 4 |
| Lipid-modifying medication use | 922 (33.4) | 4 | 308 (47.2) | 0 |
| Antihypertensive medication use | 1,019 (36.9) | 4 | 336 (51.5) | 0 |
| Diabetes medication use | 471 (17.1) | 4 | 294 (45.1) | 0 |
| Insulin | 0 (0) | | 216 (33.1) | |
| Metformin | 444 (16.1) | | 234 (35.9) | |
| Sulfonylureas | 152 (5.5) | | 46 (7.1) | |
| Thiazolidinediones | 9 (0.3) | | 4 (0.6) | |
| GLP-1 analogs | 6 (0.2) | | 3 (0.5) | |
| DDP-4 inhibitors | 51 (1.9) | | 8 (1.2) | |
| History of CVD | 400 (14.9) | 78 | 152 (24.3) | 27 |
| eGFR, mL/min/1.73 m2 | 88.4±14.2 | 6 | 86.8±17.3 | 27 |
| eGFR < 60 mL/min/1.73 m2 | 92 (3.3) | | 51 (8.2) | |

5

Table S5.2    (continued)

| Characteristic | IGP study population (n=2,758) | Missing data | Excluded due to missing OGTT and outcome values (n=652) | Missing data |
|---|---|---|---|---|
| (Micro)albuminuria | 200 (7.3) | 20 | 91 (14.4) | 22 |
| Retinopathy | 16 (0.6) | 266 | 30 (4.6) | 86 |
| Carotid-femoral pulse wave velocity (cf-PWV), m/s | 8.9±2.0 | 397 | 9.5±2.5 | 162 |
| Carotid distensibility coefficient (carDC), 10-3/kPa | 14.4±5.2 | 368 | 13.7±4.8 | 148 |
| Carotid intima-media thickness (cIMT), µm | 858.9±156.2 | 373 | 858.2±164.7 | 149 |
| Mean circumferential wall stress (CWSmean), kPa | 45.9 [39.7-53.4] | 377 | 46.4 [40.5-54.2] | 149 |
| Pulsatile circumferential wall stress (CWSpuls), kPa | 22.8 [17.8-28.9] | 536 | 23.8 [19.1-30.1] | 193 |
| Retinal arteriolar average dilatation (DVA), % | 2.7 [0.9-5.1] | 877 | 2.3 [0.7-4.1] | 272 |
| Heat-induced skin hyperemia, % | 999.1 [588.9-1,512.9] | 1388 | 954.1 [560.6-1,387.4] | 375 |

Data are reported as mean ± SD, median [interquartile range], or number (percentage %) as appropriate. Data represent the study population of participants with complete oral glucose tolerance test data and results of at least one primary outcome. CVD, cardiovascular disease; SBP, systolic blood pressure; DBP, diastolic blood pressure; NGM, normal glucose metabolism; HbA$_{1c}$, glycated hemoglobin A$_{1c}$; HDL, high-density lipoprotein; LDL, low-density lipoprotein; GLP-1, glucagon-like peptide-1; DPP-4, dipeptidase-4; eGFR, estimated glomerular filtration rate.

Table S5.3    Glucose peak time point during oral glucose tolerance test for the total study population and incremental glucose peak (IGP) tertiles.

| Peak time point during OGTT | Total study population (n=2,758) | first tertile (n=924) | second tertile (n=909) | third tertile (n=925) |
|---|---|---|---|---|
| Fasting plasma glucose (FPG) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| 15 minutes | 151 (5.5) | 143 (15.5) | 8 (0.9) | 0 (0) |
| 30 minutes | 682 (24.7) | 436 (47.2) | 225 (24.8) | 21 (2.3) |
| 45 minutes | 758 (27.5) | 217 (23.5) | 407 (44.8) | 124 (14.5) |
| 60 minutes | 574 (20.8) | 76 (8.2) | 199 (21.9) | 299 (32.3) |
| 90 minutes | 446 (16.2) | 33 (3.6) | 50 (5.5) | 363 (39.2) |
| 120 minutes | 147 (5.3) | 19 (2.1) | 20 (2.2) | 108 (11.7) |

Data are reported as number (percentage %). OGTT, oral glucose tolerance test

**Table S5.4** Multivariable-adjusted associations of incremental glucose peak (IGP) and arterial stiffness and arterial remodeling.

| Model | B (95%CI) | p-value |
|---|---|---|
| *Carotid femoral pulse wave velocity (cf-PWV), m/s (n=1,954)* | | |
| Crude | 0.172 (0.144; 0.201) | <0.001 |
| Model 2 | 0.110 (0.083; 0.136) | <0.001 |
| Model 3 | 0.099 (0.063; 0.134) | <0.001 |
| Model 4* | 0.067 (0.034; 0.100) | <0.001 |
| Model 5 | 0.054 (0.020; 0.089) | 0.002 |
| *Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=1,978)* | | |
| Crude | -0.315 (-0.389; -0.240) | <0.001 |
| Model 2 | -0.187 (-0.256; -0.118) | <0.001 |
| Model 3 | -0.093 (-0.184; -0.001) | 0.047 |
| Model 4* | -0.045 (-0.128; 0.039) | 0.294 |
| Model 5 | -0.026 (-0.112; 0.060) | 0.551 |
| Carotid intima-media thickness (cIMT), μm (n=1,973) | | |
| Crude | 4.157 (1.944; 6.370) | <0.001 |
| Model 2 | -0.094 (-2.274; 2.087) | 0.933 |
| Model 3 | -2.089 (-4.980; 0.801) | 0.157 |
| Model 4[†] | -3.159 (-6.047; -0.271) | 0.032 |
| Model 5 | -2.745 (-5.736; 0.245) | 0.072 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=1,870) | | |
| Crude | 0.441 (0.279; 0.603) | <0.001 |
| Model 2 | 0.325 (0.159; 0.491) | <0.001 |
| Model 3 | 0.404 (0.185; 0.623) | <0.001 |
| Model 4[‡] | 0.350 (0.137; 0.564) | 0.001 |
| Model 5 | 0.227 (0.008; 0.446) | 0.043 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=1,849) | | |
| Crude | 0.387 (0.257; 0.518) | <0.001 |
| Model 2 | 0.244 (0.110; 0.377) | <0.001 |
| Model 3 | 0.320 (0.144; 0.495) | <0.001 |
| Model 4* | 0.196 (0.040; 0.353) | 0.014 |
| Model 5 | 0.108 (-0.054; 0.270) | 0.192 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or carotid pulse pressure (‡) and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

5

**Table S5.5** Multivariable-adjusted associations of incremental glucose peak (IGP) and arterial stiffness and arterial remodeling, additionally adjusted for history of cardiovascular disease, retinopathy, estimated glomerular filtration rate, and urinary albumin excretion.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=1,722) | | |
| Crude | 0.168 (0.138; 0.199) | <0.001 |
| Model 2 | 0.105 (0.076; 0.134) | <0.001 |
| Model 3 | 0.096 (0.058; 0.134) | <0.001 |
| Model 4* | 0.065 (0.030; 0.100) | <0.001 |
| Model 5 | 0.056 (0.019; 0.093) | 0.003 |
| Model 6 | 0.054 (0.017; 0.091) | 0.004 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=1,744) | | |
| Crude | -0.321 (-0.400; -0.242) | <0.001 |
| Model 2 | -0.195 (-0.269; -0.121) | <0.001 |
| Model 3 | -0.109 (-0.207; -0.011) | 0.029 |
| Model 4* | -0.060 (-0.149; 0.030) | 0.189 |
| Model 5 | -0.035 (-0.128; 0.057) | 0.453 |
| Model 6 | -0.031 (-0.124; 0.062) | 0.511 |
| Carotid intima-media thickness (cIMT), μm (n=1,741) | | |
| Crude | 4.045 (1.729; 6.362) | 0.001 |
| Model 2 | -0.197 (-2.489; 2.094) | 0.866 |
| Model 3 | -2.491 (-5.537; 0.555) | 0.109 |
| Model 4[†] | -3.415 (-6.463; -0.367) | 0.028 |
| Model 5 | -2.920 (-6.088; 0.248) | 0.071 |
| Model 6 | -2.827 (-6.016; 0.362) | 0.082 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=1,633) | | |
| Crude | 0.408 (0.235; 0.582) | <0.001 |
| Model 2 | 0.281 (0.102; 0.459) | 0.002 |
| Model 3 | 0.420 (0.184; 0.656) | <0.001 |
| Model 4[‡] | 0.372 (0.143; 0.601) | 0.001 |
| Model 5 | 0.250 (0.013; 0.486) | 0.039 |
| Model 6 | 0.224 (-0.013; 0.462) | 0.064 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=1,633) | | |
| Crude | 0.333 (0.194; 0.471) | <0.001 |
| Model 2 | 0.188 (0.047; 0.329) | 0.009 |
| Model 3 | 0.280 (0.094; 0.466) | 0.003 |
| Model 4* | 0.152 (-0.014; 0.317) | 0.073 |
| Model 5 | 0.069 (-0.103; 0.242) | 0.431 |
| Model 6 | 0.066 (-0.108; 0.239) | 0.458 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for mean arterial pressure (*), office systolic blood pressure ([†]) or carotid pulse pressure ([‡]) and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels. Model 6: additionally adjusted for history of cardiovascular disease, retinopathy, estimated glomerular filtration rate, and urinary albumin excretion.

**Table S5.6** Multivariable-adjusted associations of incremental glucose peak (IGP) and microvascular function, additionally adjusted for history of cardiovascular disease, retinopathy, estimated glomerular filtration rate, and urinary albumin excretion.

| Model | B (95%CI) | p-value |
|---|---|---|
| Retinal arteriolar baseline diameter, MU (n=1,514) | | |
| Crude | -0.043 (-0.308; 0.223) | 0.753 |
| Model 2 | 0.076 (-0.202; 0.354) | 0.591 |
| Model 3 | -0.133 (-0.498; 0.232) | 0.474 |
| Model 4 | -0.079 (-0.446; 0.287) | 0.672 |
| Model 5 | -0.143 (-0.522; 0.235) | 0.458 |
| Model 6 | -0.149 (-0.527; 0.230) | 0.440 |
| Retinal arteriolar average dilatation, % (n=1,514) | | |
| Crude | -0.091 (-0.138; -0.044) | <0.001 |
| Model 2 | -0.077 (-0.125; -0.028) | 0.002 |
| Model 3 | -0.041 (-0.105; 0.022) | 0.203 |
| Model 4 | -0.045 (-0.109; 0.019) | 0.169 |
| Model 5 | -0.023 (-0.090; 0.043) | 0.497 |
| Model 6 | -0.018 (-0.084; 0.049) | 0.607 |
| Skin baseline blood flow, PU (n=1,013) | | |
| Crude | 0.010 (-0.123; 0.143) | 0.880 |
| Model 2 | -0.034 (-0.172; 0.105) | 0.634 |
| Model 3 | 0.028 (-0.156; 0.211) | 0.767 |
| Model 4 | 0.052 (-0.134; 0.237) | 0.582 |
| Model 5 | 0.077 (-0.117; 0.270) | 0.437 |
| Model 6 | 0.042 (-0.153; 0.237) | 0.674 |
| Heat-induced skin hyperemia, % (n=1,013) | | |
| Crude | -26.408 (-41.668; -11.147) | 0.001 |
| Model 2 | -8.673 (-24.293; 6.946) | 0.276 |
| Model 3 | -1.164 (-21.867; 19.540) | 0.912 |
| Model 4 | -3.704 (-24.615; 17.208) | 0.728 |
| Model 5 | -0.211 (-22.086; 21.664) | 0.985 |
| Model 6 | 0.194 (-21.935; 22.322) | 0.986 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for office systolic blood pressure. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels. Model 6: additionally adjusted for history of cardiovascular disease, retinopathy, estimated glomerular filtration rate, and urinary albumin excretion.

5

**Table S5.7** Multivariable-adjusted associations of incremental glucose peak (IGP) and arterial stiffness and arterial remodeling, additionally adjusted for fasting plasma insulin (6.1) or HOMA2-IR (6.2).

| Model | B (95%CI) | p-value |
|---|---|---|
| **Carotid femoral pulse wave velocity (cf-PWV), m/s (n=1,890)** | | |
| Crude | 0.174 (0.144; 0.203) | <0.001 |
| Model 2 | 0.109 (0.081; 0.136) | <0.001 |
| Model 3 | 0.101 (0.065; 0.138) | 0.001 |
| Model 4* | 0.070 (0.036; 0.103) | <0.001 |
| Model 5 | 0.056 (0.021; 0.091) | 0.002 |
| Model 6.1 | 0.056 (0.021; 0.091) | 0.002 |
| Model 6.2 | 0.056 (0.021; 0.091) | 0.002 |
| **Carotid distensibility coefficient (carDC), 10-3/kPa (n=1,908)** | | |
| Crude | -0.316 (-0.392; -0.240) | < 0.001 |
| Model 2 | -0.185 (-0.256; -0.114) | < 0.001 |
| Model 3 | -0.099 (-0.192; -0.005) | 0.039 |
| Model 4* | -0.050 (-0.135; 0.036) | 0.254 |
| Model 5 | -0.030 (-0.118; 0.059) | 0.509 |
| Model 6.1 | -0.030 (-0.118; 0.059) | 0.509 |
| Model 6.2 | -0.031 (-0.120; 0.058) | 0.491 |
| **Carotid intima-media thickness (cIMT), μm (n=1,903)** | | |
| Crude | 3.975 (1.710; 6.239) | 0.001 |
| Model 2 | -0.306 (-2.542; 1.929) | 0.788 |
| Model 3 | -2.105 (-5.064; 0.854) | 0.163 |
| Model 4† | -3.215 (-6.172; -0.258) | 0.033 |
| Model 5 | -2.791 (-5.859; 0.277) | 0.075 |
| Model 6.1 | -2.797 (-5.866; 0.272) | 0.074 |
| Model 6.2 | -2.782 (-5.851; 0.287) | 0.076 |
| **Mean circumferential wall stress (CWSmean), kPa (n=1,784)** | | |
| Crude | 0.445 (0.278; 0.612) | < 0.001 |
| Model 2 | 0.329 (0.157; 0.501) | < 0.001 |
| Model 3 | 0.406 (0.180; 0.633) | < 0.001 |
| Model 4‡ | 0.345 (0.124; 0.565) | 0.002 |
| Model 5 | 0.201 (-0.025; 0.428) | 0.082 |
| Model 6.1 | 0.202 (-0.025; 0.428) | 0.082 |
| Model 6.2 | 0.203 (-0.024; 0.430) | 0.079 |
| **Pulsatile circumferential wall stress (CWSpuls), kPa (n=1,784)** | | |
| Crude | 0.394 (0.261; 0.527) | < 0.001 |
| Model 2 | 0.253 (0.117; 0.389) | < 0.001 |
| Model 3 | 0.326 (0.147; 0.505) | < 0.001 |
| Model 4* | 0.202 (0.043; 0.362) | 0.013 |
| Model 5 | 0.105 (-0.060; 0.271) | 0.213 |
| Model 6.1 | 0.105 (-0.060; 0.271) | 0.212 |
| Model 6.2 | 0.106 (-0.059; 0.272) | 0.209 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for HbA$_{1c}$. Model 4: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or carotid pulse pressure (‡)and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels. Model 6 additionally adjusted for fasting plasma insulin (6.1) or HOMA2-IR (6.2).

Table S5.8   Multivariable-adjusted associations of incremental glucose peak (IGP) and arterial stiffness, arterial remodeling and microvascular function after adjustment for glucose metabolism status or fasting plasma glucose instead of HbA$_{1c}$.

| Model | B (95%CI) | p-value | Model | B (95%CI) | p-value |
|---|---|---|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=1,954) | | | | | |
| Crude | 0.172 (0.144; 0.201) | <0.001 | Crude | | |
| Model 2 | 0.110 (0.083; 0.136) | <0.001 | Model 2 | | |
| Model 3a | 0.074 (0.032; 0.116) | 0.001 | Model 3b | 0.079 (0.043; 0.114) | <0.001 |
| Model 4a* | 0.040 (0.001; 0.079) | 0.045 | Model 4b* | 0.061 (0.028; 0.094) | <0.001 |
| Model 5a | 0.031 (-0.008; 0.071) | 0.120 | Model 5b | 0.048 (0.014; 0.082) | 0.006 |
| Carotid distensibility coefficient (carDC), 10-3/kPa (n=1,978) | | | | | |
| Crude | -0.315 (-0.389; -0.240) | < 0.001 | Crude | | |
| Model 2 | -0.187 (-0.256; -0.118) | < 0.001 | Model 2 | | |
| Model 3a | -0.162 (-0.270; -0.053) | 0.003 | Model 3b | -0.090 (-0.182; 0.002) | 0.054 |
| Model 4a* | -0.109 (-0.208; -0.011) | 0.030 | Model 4b* | -0.078 (-0.162; 0.005) | 0.067 |
| Model 5a | -0.095 (-0.194; 0.005) | 0.062 | Model 5b | -0.061 (-0.148; 0.026) | 0.166 |
| Carotid intima-media thickness (cIMT), μm (n=1,973) | | | | | |
| Crude | 4.157 (1.944; 6.370) | <0.001 | Crude | | |
| Model 2 | -0.094 (-2.274; 2.087) | 0.933 | Model 2 | | |
| Model 3a | -3.582 (-6.991; -0.172) | 0.039 | Model 3b | -2.446 (-5.347; 0.454) | 0.098 |
| Model 4a† | -4.181 (-7.568; -0.793) | 0.016 | Model 4b† | -3.067 (-5.952; -0.182) | 0.037 |
| Model 5a | -4.282 (-7.706; -0.857) | 0.014 | Model 5b | -3.004 (-6.005; -0.002) | 0.050 |
| Mean circumferential wall stress (CWSmean), kPa (n=1,870) | | | | | |
| Crude | 0.441 (0.279; 0.603) | <0.001 | Crude | | |
| Model 2 | 0.325 (0.159; 0.491) | <0.001 | Model 2 | | |
| Model 3a | 0.429 (0.169; 0.689) | 0.001 | Model 3b | 0.258 (0.038; 0.479) | 0.022 |
| Model 4a‡ | 0.380 (0.127; 0.633) | 0.003 | Model 4b‡ | 0.219 (0.004; 0.434) | 0.046 |
| Model 5a | 0.343 (0.090; 0.596) | 0.008 | Model 5b | 0.156 (-0.065; 0.377) | 0.167 |
| Pulsatile circumferential wall stress (CWSpuls), kPa (n=1,849) | | | | | |
| Crude | 0.387 (0.257; 0.518) | <0.001 | Crude | | |
| Model 2 | 0.244 (0.110; 0.377) | <0.001 | Model 2 | | |
| Model 3a | 0.265 (0.057; 0.472) | 0.013 | Model 3b | 0.214 (0.037; 0.391) | 0.018 |
| Model 4a* | 0.126 (-0.059; 0.312) | 0.182 | Model 4b* | 0.159 (0.001; 0.316) | 0.048 |
| Model 5a | 0.076 (-0.111; 0.264) | 0.423 | Model 5b | 0.067 (-0.097; 0.230) | 0.425 |
| Retinal arteriolar average dilatation (n=1,591) | | | | | |
| Crude | -0.088 (-0.134; -0.043) | <0.001 | Crude | | |
| Model 2 | -0.073 (-0.121; -0.026) | 0.002 | Model 2 | | |
| Model 3a | 0.019 (-0.055; 0.093) | 0.619 | Model 3b | -0.026 (-0.089; 0.037) | 0.424 |
| Model 4a† | 0.015 (-0.059; 0.090) | 0.689 | Model 4b† | -0.029 (-0.092; 0.035) | 0.375 |
| Model 5a | 0.026 (-0.049; 0.101) | 0.501 | Model 5b | -0.006 (-0.072; 0.060) | 0.854 |
| Heat-induced skin hyperemia (n=1,134) | | | | | |
| Crude | -28.109 (-42.778; -13.440) | <0.001 | Crude | | |
| Model 2 | -12.503 (-27.509; 2.504) | 0.102 | Model 2 | | |
| Model 3a | 1.347 (-22.597; 25.291) | 0.912 | Model 3b | -10.477 (-30.879; 9.926) | 0.314 |
| Model 4a† | -0.043 (-24.042; 23.956) | 0.997 | Model 4b† | -11.996 (-32.497; 8.505) | 0.251 |
| Model 5a | 4.340 (-19.989; 28.669) | 0.726 | Model 5b | -6.214 (-27.437; 15.009) | 0.566 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for glucose metabolism status (a, left column) or fasting plasma glucose (FPG) (b, right column). Model 4: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or carotid pulse pressure (‡) and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

5

**Table S5.9** Multivariable-adjusted associations of absolute glucose peak (AGP) and arterial stiffness, arterial remodeling and microvascular function.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=1,954) | | |
| Crude | 0.144 (0.121; 0.166) | <0.001 |
| Model 2 | 0.090 (0.069; 0.112) | <0.001 |
| Model 3 | 0.092 (0.061; 0.123) | <0.001 |
| Model 4* | 0.059 (0.031; 0.088) | <0.001 |
| Model 5 | 0.049 (0.019; 0.079) | 0.001 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=1,978) | | |
| Crude | -0.268 (-0.326; -0.210) | <0.001 |
| Model 2 | -0.163 (-0.217; -0.108) | <0.001 |
| Model 3 | -0.097 (-0.176; -0.017) | 0.017 |
| Model 4* | -0.042 (-0.115; 0.030) | 0.250 |
| Model 5 | -0.020 (-0.096; 0.055) | 0.598 |
| Carotid intima-media thickness (cIMT), μm (n=1,973) | | |
| Crude | 4.013 (2.290; 5.736) | <0.001 |
| Model 2 | 0.380 (-1.338; 2.098) | 0.664 |
| Model 3 | -1.217 (-3.719; 1.285) | 0.340 |
| Model 4[†] | -2.296 (-4.803; 0.212) | 0.073 |
| Model 5 | -1.778 (-4.389; 0.833) | 0.182 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=1,870) | | |
| Crude | 0.363 (0.237; 0.490) | <0.001 |
| Model 2 | 0.263 (0.132; 0.394) | <0.001 |
| Model 3 | 0.382 (0.192; 0.572) | 0.001 |
| Model 4[‡] | 0.333 (0.148; 0.517) | <0.001 |
| Model 5 | 0.200 (0.008; 0.391) | 0.041 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=1,849) | | |
| Crude | 0.315 (0.213; 0.417) | <0.001 |
| Model 2 | 0.193 (0.088; 0.299) | <0.001 |
| Model 3 | 0.296 (0.144; 0.447) | <0.001 |
| Model 4* | 0.167 (0.031; 0.302) | 0.016 |
| Model 5 | 0.087 (-0.054; 0.228) | 0.228 |
| Retinal arteriolar average dilatation (n=1,591) | | |
| Crude | -0.076 (-0.111; -0.041) | <0.001 |
| Model 2 | -0.065 (-0.102; -0.028) | 0.001 |
| Model 3 | -0.043 (-0.097; 0.011) | 0.116 |
| Model 4[†] | -0.048 (-0.102; 0.007) | 0.085 |
| Model 5 | -0.031 (-0.087; 0.026) | 0.284 |
| Heat-induced skin hyperemia (n=1,134) | | |
| Crude | -23.140(-34.467; -11.813) | <0.001 |
| Model 2 | -9.906 (-21.625; 1.813) | 0.097 |
| Model 3 | -1.556 (-18.602; 15.490) | 0.858 |
| Model 4[†] | -3.408 (-20.647; 13.831) | 0.698 |
| Model 5 | -0.757 (-18.811; 17.297) | 0.934 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of absolute glucose peak (AGP). Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for mean arterial blood pressure and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

**Table S5.10** Multivariable-adjusted associations of percentage increase from baseline (IGP$_{percentage}$) and arterial stiffness, arterial remodeling and microvascular function.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=1,954) | | |
| Crude | 0.011 (0.009; 0.013) | <0.001 |
| Model 2 | 0.007 (0.005; 0.009) | <0.001 |
| Model 3 | 0.005 (0.003; 0.007) | <0.001 |
| Model 4* | 0.004 (0.001; 0.006) | 0.001 |
| Model 5 | 0.003 (0.001; 0.005) | 0.010 |
| Carotid distensibility coefficient (carDC), 10$^{-3}$/kPa (n=1,978) | | |
| Crude | -0.020 (-0.025; -0.014) | <0.001 |
| Model 2 | -0.011 (-0.016; -0.006) | <0.001 |
| Model 3 | -0.005 (-0.011; 0.000) | 0.071 |
| Model 4* | -0.003 (-0.008; 0.002) | 0.274 |
| Model 5 | -0.002 (-0.007; 0.003) | 0.498 |
| Carotid intima-media thickness (cIMT), μm (n=1,973) | | |
| Crude | 0.204 (0.040; 0.368) | 0.015 |
| Model 2 | -0.062 (-0.221; 0.097) | 0.445 |
| Model 3 | -0.154 (-0.335; 0.026) | 0.094 |
| Model 4† | -0.203 (-0.383; -0.023) | 0.027 |
| Model 5 | -0.193 (-0.378; -0.008) | 0.041 |
| Mean circumferential wall stress (CWS$_{mean}$), kPa (n=1,870) | | |
| Crude | 0.029 (0.017; 0.041) | <0.001 |
| Model 2 | 0.022 (0.010; 0.034) | <0.001 |
| Model 3 | 0.022 (0.008; 0.036) | 0.002 |
| Model 4‡ | 0.019 (0.005; 0.032) | 0.006 |
| Model 5 | 0.013 (-0.001; 0.026) | 0.069 |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa (n=1,849) | | |
| Crude | 0.027 (0.017; 0.036) | <0.001 |
| Model 2 | 0.017 (0.008; 0.027) | <0.001 |
| Model 3 | 0.018 (0.007; 0.029) | 0.001 |
| Model 4* | 0.012 (0.002; 0.021) | 0.019 |
| Model 5 | 0.007 (-0.003; 0.017) | 0.173 |
| Retinal arteriolar average dilatation (n=1,591) | | |
| Crude | -0.005 (-0.008; -0.001) | 0.005 |
| Model 2 | -0.004 (-0.007; 0.000) | 0.040 |
| Model 3 | -0.001 (-0.005; 0.003) | 0.536 |
| Model 4† | -0.001 (-0.005; 0.003) | 0.476 |
| Model 5 | 0.000 (-0.004; 0.004) | 0.910 |
| Heat-induced skin hyperemia (n=1,134) | | |
| Crude | -1.852 (-2.951; -0.754) | 0.001 |
| Model 2 | -0.896 (-2.002; 0.209) | 0.112 |
| Model 3 | -0.440 (-1.686; 0.805) | 0.488 |
| Model 4† | -0.544 (-1.797; 0.709) | 0.395 |
| Model 5 | -0.289 (-1.577; 1.000) | 0.661 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (%) increase of percentage increase from baseline (IGPpercentage). Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for HbA$_{1c}$. Model 4: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or carotid pulse pressure (‡) and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

5

Table S5.11. Multivariable-adjusted associations of incremental glucose peak (IGP) and arterial stiffness after adjustment for alternative blood pressure measurements.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid-femoral pulse wave velocity (cf-PWV), m/s (n=1,950) | | |
| Crude | 0.171 (0.142; 0.199) | <0.001 |
| Model 2 | 0.109 (0.082; 0.136) | <0.001 |
| Model 3 | 0.097 (0.062; 0.133) | <0.001 |
| Model 4* | 0.048 (0.014; 0.082) | 0.005 |
| Model 5 | 0.041 (0.006; 0.076) | 0.021 |
| Carotid-femoral pulse wave velocity (cf-PWV), m/s (n=1,751) | | |
| Crude | 0.170 (0.140; 0.200) | <0.001 |
| Model 2 | 0.110 (0.081; 0.138) | <0.001 |
| Model 3 | 0.100 (0.062; 0.138) | <0.001 |
| Model 4† | 0.074 (0.039; 0.110) | <0.001 |
| Model 5 | 0.061 (0.024; 0.098) | 0.001 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=1,974) | | |
| Crude | -0.313 (-0.388; -0.239) | <0.001 |
| Model 2 | -0.186 (-0.255; -0.117) | <0.001 |
| Model 3 | -0.092 (-0.184; -0.001) | 0.048 |
| Model 4* | -0.029 (-0.119; 0.060) | 0.521 |
| Model 5 | -0.021 (-0.113; 0.070) | 0.648 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=1,769) | | |
| Crude | -0.274 (-0.353; -0.195) | <0.001 |
| Model 2 | -0.153 (-0.227; -0.079) | <0.001 |
| Model 3 | -0.061 (-0.158; 0.035) | 0.214 |
| Model 4† | -0.054 (-0.148; 0.041) | 0.266 |
| Model 5 | -0.030 (-0.128; 0.067) | 0.539 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for office systolic blood pressure (*) or ambulatory 24-h systolic blood pressure (†) and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

**Table S5.12** Multivariable-adjusted associations of incremental glucose peak (IGP) and arterial remodeling after adjustment for alternative blood pressure measurements.

| Model | B (95%CI) | p-value |
|---|:---:|:---:|
| Carotid intima-media thickness (cIMT), μm (n=1,769) | | |
| Crude | 3.315 (0.947; 5.683) | 0.006 |
| Model 2 | -1.085 (-3.409; 1.239) | 0.360 |
| Model 3 | -3.088 (-6.147; -0.029) | 0.048 |
| Model 4* | -3.287 (-6.310; -0.265) | 0.033 |
| Model 5 | -3.126 (-6.269; 0.017) | 0.051 |
| Carotid intima-media thickness (cIMT), μm (n=1,974) | | |
| Crude | 4.137 (1.919; 6.355) | <0.001 |
| Model 2 | -0.124 (-2.309; 2.061) | 0.912 |
| Model 3 | -2.136 (-5.029; 0.758) | 0.148 |
| Model 4† | -2.455 (-5.345; 0.434) | 0.096 |
| Model 5 | -2.193 (-5.198; 0.813) | 0.153 |
| Mean circumferential wall stress (CWSmean), kPa (n=1,765) | | |
| Crude | 0.443 (0.277; 0.609) | <0.001 |
| Model 2 | 0.33 (0.16; 0.501) | <0.001 |
| Model 3 | 0.404 (0.180; 0.628) | <0.001 |
| Model 4* | 0.383 (0.165; 0.602) | 0.001 |
| Model 5 | 0.268 (0.043; 0.492) | 0.019 |
| Mean circumferential wall stress (CWSmean), kPa (n=1,969) | | |
| Crude | 0.408 (0.251; 0.564) | <0.001 |
| Model 2 | 0.288 (0.128; 0.449) | <0.001 |
| Model 3 | 0.385 (0.172; 0.598) | <0.001 |
| Model 4‡ | 0.248 (0.040; 0.457) | 0.020 |
| Model 5 | 0.162 (-0.053; 0.377) | 0.139 |
| Pulsatile circumferential wall stress (CWSpuls), kPa (n=1,767) | | |
| Crude | 0.372 (0.235; 0.509) | <0.001 |
| Model 2 | 0.227 (0.088; 0.366) | 0.001 |
| Model 3 | 0.283 (0.101; 0.464) | 0.002 |
| Model 4* | 0.250 (0.077; 0.423) | 0.005 |
| Model 5 | 0.163 (-0.016; 0.342) | 0.074 |
| Pulsatile circumferential wall stress (CWSpuls), kPa (n=1,848) | | |
| Crude | 0.398 (0.268; 0.529) | <0.001 |
| Model 2 | 0.256 (0.123; 0.389) | <0.001 |
| Model 3 | 0.333 (0.158; 0.509) | <0.001 |
| Model 4‡ | 0.155 (-0.011; 0.320) | 0.067 |
| Model 5 | 0.099 (-0.073; 0.270) | 0.259 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for ambulatory 24-hour systolic blood pressure (*), mean arterial pressure (†), or office systolic blood pressure (‡). Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

5

**Table S5.13** Multivariable-adjusted associations of incremental glucose peak (IGP) and microvascular function after adjustment for alternative blood pressure measurements.

| Model | B (95%CI) | p-value |
|---|---|---|
| Retinal arteriolar average dilatation (n=1,422) | | |
| Crude | -0.091 (-0.14; -0.043) | <0.001 |
| Model 2 | -0.075 (-0.125; -0.025) | 0.003 |
| Model 3 | -0.040 (-0.106; 0.026) | 0.238 |
| Model 4 | -0.040 (-0.106; 0.026) | 0.231 |
| Model 5 | -0.023 (-0.092; 0.046) | 0.512 |
| Heat-induced skin hyperemia (n=1,000) | | |
| Crude | -27.164 (-42.764; -11.563) | 0.001 |
| Model 2 | -13.087 (-28.994; 2.819) | 0.107 |
| Model 3 | -5.924 (-27.064; 15.216) | 0.583 |
| Model 4 | -5.654 (-26.806; 15.499) | 0.600 |
| Model 5 | -0.237 (-22.360; 21.886) | 0.983 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for HbA$_{1c}$. Model 4: additionally adjusted for ambulatory 24-hour systolic blood pressure. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

**Table S5.14** Multivariable-adjusted associations of time to glucose peak and arterial stiffness, arterial remodeling and microvascular function.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=1,954) | | |
| Crude | 0.018 (0.015; 0.022) | <0.001 |
| Model 2 | 0.011 (0.008; 0.008) | <0.001 |
| Model 3 | 0.009 (0.005; 0.005) | <0.001 |
| Model 4* | 0.006 (0.002; 0.009) | 0.001 |
| Model 5 | 0.005 (0.001; 0.008) | 0.007 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=1,978) | | |
| Crude | -0.030 (-0.039; -0.022) | <0.001 |
| Model 2 | -0.014 (-0.022; -0.006) | <0.001 |
| Model 3 | -0.004 (-0.013; 0.005) | 0.425 |
| Model 4* | 0.004 (-0.005; 0.012) | 0.402 |
| Model 5 | 0.006 (-0.002; 0.015) | 0.131 |
| Carotid intima-media thickness (cIMT), μm (n=1,973) | | |
| Crude | 0.501 (0.246; 0.736) | <0.001 |
| Model 2 | 0.031 (-0.219; 0.282) | 0.807 |
| Model 3 | -0.081 (-0.368; 0.206) | 0.579 |
| Model 4† | -0.189 (-0.475; 0.098) | 0.196 |
| Model 5 | -0.140 (-0.433; 0.152) | 0.348 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=1,869) | | |
| Crude | 0.048 (0.029; 0.067) | <0.001 |
| Model 2 | 0.036 (0.017; 0.055) | <0.001 |
| Model 3 | 0.037 (0.015; 0.059) | 0.001 |
| Model 4‡ | 0.029 (0.008; 0.051) | 0.007 |
| Model 5 | 0.015 (-0.006; 0.037) | 0.162 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=1,849) | | |
| Crude | 0.043 (0.028; 0.058) | <0.001 |
| Model 2 | 0.027 (0.012; 0.042) | 0.001 |
| Model 3 | 0.029 (0.011; 0.046) | 0.001 |
| Model 4* | 0.011 (-0.004; 0.027) | 0.155 |
| Model 5 | 0.005 (-0.011; 0.021) | 0.546 |
| Retinal arteriolar average dilatation (n=1,591) | | |
| Crude | -0.011 (-0.016; -0.005) | <0.001 |
| Model 2 | -0.009 (-0.014; -0.028) | 0.001 |
| Model 3 | -0.007 (-0.012; -0.001) | 0.029 |
| Model 4† | -0.007 (-0.013; -0.001) | 0.022 |
| Model 5 | -0.006 (-0.012; 0.001) | 0.074 |
| Heat-induced skin hyperemia (n=1,134) | | |
| Crude | -3.181 (-4.867; -1.495) | <0.001 |
| Model 2 | -1.521 (-3.236; 0.194) | 0.082 |
| Model 3 | -0.814 (-2.777; 1.149) | 0.416 |
| Model 4† | -1.046 (-3.032; 0.940) | 0.302 |
| Model 5 | -1.044 (-3.080; 0.993) | 0.315 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (minute) increase of time to glucose peak. Model 1: crude. Model 2: additionally adjusted for age and sex. Model 3: additionally adjusted for $HbA_{1c}$. Model 4: additionally adjusted for mean arterial blood pressure and heart rate in case of cf-PWV. Model 5: additionally adjusted for body mass index, smoking status, physical activity, Mediterranean diet score, use of antihypertensive and lipid-modifying drugs, fasting triglycerides, and total-to-HDL cholesterol levels.

5

Table S5.15  Fully adjusted associations of incremental glucose peak (IGP) and carotid-femoral pulse wave velocity (cf-PWV) stratified according to age tertiles.

| Age tertiles | Model 5: B (95%CI) | p-value |
|---|---|---|
| Carotid-femoral pulse wave velocity (cf-PWV), m/s | | |
| 1st (n=627) | 0.032 (-0.021; 0.085) | 0.234 |
| 2nd (n=707) | 0.034 (-0.020; 0.088) | 0.219 |
| 3rd (n=620) | 0.101 (0.032; 0.169) | 0.004 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 5: fully adjusted model.


Table S5.16  Fully adjusted associations of incremental glucose peak (IGP) and carotid distensibility coefficient (carDC) stratified according to age tertiles.

| Age tertiles | Model 5: B (95%CI) | p-value |
|---|---|---|
| Carotid destensibility coefficient (carDC), 10-3/kPa | | |
| 1st (n=634) | -0.120 (-0.306; 0.066) | 0.207 |
| 2nd (n=712) | -0.070 (-0.205; 0.066) | 0.313 |
| 3rd (n=632) | 0.071 (-0.067; 0.208) | 0.313 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 5: fully adjusted model.


Table S5.17  Fully adjusted associations of incremental glucose peak (IGP) and pulsatile circumferential wall stress (CWS$_{puls}$) stratified according to age tertiles.

| Age tertiles | Model 5: B (95%CI) | p-value |
|---|---|---|
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa | | |
| 1st (n=525) | -0.135 (-0.405; 0.135) | 0.326 |
| 2nd (n=594) | 0.170 (-0.112; 0.451) | 0.237 |
| 3rd (n=514) | 0.270 (-0.027; 0.566) | 0.075 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of IGP. Model 5: fully adjusted model.

# Chapter **6**

Greater daily glucose variability and lower time in range assessed with continuous glucose monitoring are associated with greater aortic stiffness – The Maastricht Study

Yuri D. Foreman, William P.T.M. van Doorn, Nicolaas C. Schaper, Marleen M.J. van Greevenbroek, Carla J.H. van der Kallen, Ronald M.A. Henry, Annemarie Koster, Simone J.P.M. Eussen, Anke Wesselius, Koen D. Reesink, Miranda T. Schram, Pieter C. Dagnelie, Abraham A. Kroon, Martijn C.G.J. Brouwers, and Coen D.A. Stehouwer

# Abstract

### Aims

Cardiovascular disease (CVD) is the main cause of morbidity and mortality in individuals with diabetes. It is currently unclear whether daily glucose variability contributes to CVD. We therefore investigated whether glucose variability is associated with arterial measures that are considered important in the pathogenesis of CVD.

### Methods

We included participants of The Maastricht Study, an observational population-based cohort, who underwent at least 48h of continuous glucose monitoring (CGM) (n=853; age: 59.9±8.6y), 49% women, 23% type 2 diabetes). We studied the cross-sectional associations of two glucose variability indices (standard deviation [$SD_{CGM}$]; coefficient of variation [$CV_{CGM}$]) and time in range ($TIR_{CGM}$) with carotid-femoral pulse wave velocity (cf-PWV), carotid distensibility coefficient, carotid intima-media thickness, ankle-brachial index, and circumferential wall stress via multiple linear regression.

### Results

Higher $SD_{CGM}$ was associated with higher cf-PWV after adjustment for demographics, cardiovascular risk factors, lifestyle factors, and medication use (regression coefficient [B] per 1mmol/L $SD_{CGM}$ [and corresponding 95% confidence interval; 95%CI]: 0.413m/s [0.147;0.679], p=0.002). In the model additionally adjusted for mean sensor glucose ($MSG_{CGM}$), $SD_{CGM}$ and $MSG_{CGM}$ contributed similarly to cf-PWV (respective standardized regression coefficients [st.β and 95%CI] of 0.065 [-0.018; 0.167], p=0.160; and 0.059 [-0.043; 0.164], p=0.272). In the fully adjusted models, both higher $CV_{CGM}$ (B [95%CI] per 10% $CV_{CGM}$: 0.303m/s [0.046;0.559], p=0.021) and lower $TIR_{CGM}$ (B [95%CI] per 10% $TIR_{CGM}$: -0.145m/s [-0.252;-0.038] p=0.008) were statistically significantly associated with higher cf-PWV. Such consistent associations were not observed for the other arterial measures.

### Conclusions

Our findings show that greater daily glucose variability and lower $TIR_{CGM}$ are associated with greater aortic stiffness (cf-PWV), but not with other arterial measures. If corroborated in prospective studies, these results support the development of therapeutic agents that target both daily glucose variability and $TIR_{CGM}$ in order to prevent CVD.

# Introduction

Cardiovascular disease (CVD) is the main cause of morbidity and mortality in individuals with type 2 diabetes[1]. Additionally, individuals with prediabetes are already at an elevated risk of CVD[2]. Hyperglycemia contributes to this CVD risk, in part, by its adverse effects on arterial stiffness[3-5], atherosclerosis[1,6], and large-artery endothelial function[5,7]. Accordingly, both achieving and maintaining normoglycemia are important for the reduction of CVD risk[1]. However, current treatment modalities have not been able to fully normalize the elevated CVD risk of individuals with type 2 diabetes[1]. A better understanding of the involved pathophysiologic processes could yield new therapeutic targets to further reduce CVD risk.

Glucose variability (GV) is thought to contribute to the development of CVD, irrespective of mean glucose values. Notably, two types of GV need to be distinguished, as they are measured differently and represent different underlying etiologic concepts[8,9]. Short-term (or daily) GV reflects actual glucose fluctuations over the day[9,10]. By contrast, long-term (or visit-to-visit) GV reflects variance in classic glycemic indices (e.g., HbA$_{1c}$) that have been periodically measured over weeks, months, or years[8, 9]. While long-term GV may assess daily glucose fluctuations to some extent, it is viewed to largely represent difficult to measure factors that affect glycemic control (e.g., therapy adherence, multimorbidity, infections)[8]. Whereas multiple studies have shown that long-term GV is independently associated with CVD and all-cause mortality[8,11-15], the association between daily GV and CVD has only been sparsely investigated[16].

In general, the study of incident CVD requires both a substantial sample size and an ample follow-up period. Large-scale measurement of daily GV with the gold standard method (i.e., continuous glucose monitoring; CGM)[17] has been challenging up until recently, in part, due to costliness and relative invasiveness[18]. Therefore, most studies on this topic have cross-sectionally associated daily GV with measures that reflect the aforementioned processes leading up to CVD[19-25]. Importantly, however, these studies either did not adjust for certain important potential confounders[20-23] or assessed daily GV with less precise methods than CGM[24,25].

Hence, we investigated whether daily GV is associated with arterial measures that are considered important in the pathogenesis of CVD in a population-based cohort study. We studied whether the associations were independent of key demographics, cardiovascular risk factors, and lifestyle factors, and assessed to what extent the associations were explained by mean glycaemia. Based on previous work[25], we hypothesized that CGM-derived indices of GV would be most strongly associated with carotid-femoral pulse wave velocity (cf-PWV), which is the gold standard measure of aortic stiffness because of its independent association with incident CVD, cardiovascular mortality, and all-cause mortality[26-28]. In secondary analyses, we assessed the associations of coefficient of variation (CV$_{CGM}$), an index that is intrinsically adjusted for mean glycaemia, and time in range (TIR$_{CGM}$), an emerging

6

glycemic index that is partly determined by GV[29], with the same arterial outcome variables.

# Research design and methods

## *Study population and design*

We used data from The Maastricht Study, an observational, prospective, population-based cohort study. The rationale and methodology have been described previously[30]. In brief, The Maastricht Study focuses on the etiology, pathophysiology, complications, and comorbidities of type 2 diabetes, and is characterized by an extensive phenotyping approach. All individuals aged between 40 and 75 years and living in the southern part of the Netherlands were eligible for participation. Participants were recruited through mass media campaigns and from the municipal registries and the regional Diabetes Patient Registry via mailings. For reasons of efficiency, recruitment was stratified according to known type 2 diabetes status, with an oversampling of individuals with type 2 diabetes. In general, the examinations of each participant were performed within a time window of three months. The Maastricht Study has been approved by the institutional medical ethical committee (NL31329.068.10) and the Minister of Health, Welfare and Sports of the Netherlands (Permit 131088-105234-PG). All participants gave written informed consent.

## *Continuous glucose monitoring*

The rationale and methodology of CGM (iPro2 and Enlite Glucose Sensor; Medtronic, Tolochenaz, Switzerland) have been described previously[31]. From 19 September 2016 until 13 September 2018, all participants were invited to undergo CGM as part of their regular work-up at The Maastricht Study. To accelerate the inclusion process and to ensure inclusion of a sufficient number of participants with prediabetes and type 2 diabetes, we re-invited a selected group of participants who had recently visited The Maastricht Study to undergo CGM as a separate research visit (further referred to as 'catch-up visit'). The CGM device was worn on the lower abdomen and recorded subcutaneous interstitial glucose values (range: 2.2 - 22.2 mmol/L) every five minutes for a seven-day period. Participants were asked to perform self-measurements of blood glucose four times daily (Contour Next; Ascensia Diabetes Care, Mijdrecht, the Netherlands) for retrospective CGM calibration. Participants were blinded to the CGM recording, but not to the self-measured values. Diabetes medication use was allowed, and no dietary or physical activity instructions were given.

The first 24 hours of CGM were excluded, because of insufficient calibration. Next, we excluded individuals of whom less than 24 hours of recording (less than one data day) remained. Then, we calculated, per participant, mean sensor glucose ($MSG_{CGM}$), $SD_{CGM}$, $CV_{CGM}$ (i.e., $SD_{CGM}/MSG_{CGM}$ * 100%), and $TIR_{CGM}$ (i.e., % of time between

3.9 - 10.0 mmol/L) using the total recording period. Based on international consensus, we used $SD_{CGM}$ and $CV_{CGM}$ as indices of GV[17].

## Arterial measurements

The rationale and methodology of the arterial measurements have been described previously[25,32,33]. We assessed cf-PWV by use of applanation tonometry (SphygmoCor, Atcor Medical, Sydney, Australia)[26], and used the median of at least three consecutive cf-PWV recordings in our analyses. Because of its established clinical relevance[26-28], cf-PWV was our main outcome measure of interest.

In addition, we measured the left common carotid artery with the use of an ultrasound scanner equipped with a 7.5-MHz linear probe (MyLab 70, Esaote Europe B.V., Maastricht, the Netherlands) to assess local carotid distension, intima-media thickness (cIMT), and interadventitial diameter (IAD)[34]. We quantified local arterial stiffness by calculating the carotid distensibility coefficient (carDC) according to the following formula: $carDC = (2*\Delta D*IAD + \Delta D^2)/(braPP*IAD^2)$, where IAD interadventitial arterial diameter, $\Delta D$ distension, and braPP brachial pulse pressure[35]. We defined cIMT as the distance between the lumen-intima and media-adventitia interfaces of the far (posterior) wall[34], and IAD as the distance between the media-adventitia interfaces of the near and far wall. The median carDC, cIMT and IAD of three consecutive measurements were used.

We calculated carotid lumen diameter (LD) according to the following formula[36]: $LD = IAD - (2*cIMT)$. In parallel with the vascular measurements, we also determined mean heart rate and mean arterial pressure (MAP) every five minutes with an oscillometric device (Accutorr Plus, Datascope Inc., Montvale, NJ, USA). We calculated mean ($CWS_{mean}$) and pulsatile ($CWS_{puls}$) carotid circumferential wall stress by use of the Lamé equation as $CWS_{mean} = (MAP*(LD/2))/cIMT$ and $CWS_{puls} = (braPP*(LD/2))/cIMT$[32].

Last, the Omron VP2000 (Omron, Kyoto, Japan) was used to automatically determine the ankle-brachial index (ABI) based on simultaneous blood pressure measurements at both ankles and upper arms. The left and right ABI were calculated by dividing the systolic blood pressure (BP) measured at the ankle by the highest systolic BP measured at either upper arm. We used the lowest ABI in our analyses, and excluded individuals with an ABI above 1.4[37].

## Measurement of covariates

As described previously[30], we categorized glucose metabolism status (GMS) based on both a standardized 2-hour 75 gram oral glucose tolerance test (OGTT) and the participant's medication use as either normal glucose metabolism (NGM), prediabetes, or type 2 diabetes[38]. Participants who used insulin or had a fasting plasma glucose value above 11.0 mmol/L did not undergo the OGTT. In addition, we assessed educational level (low, intermediate, high), moderate-to-vigorous physical activity,

6

smoking status (never, former, current), alcohol use (none, low, high), and history of CVD by questionnaire; calculated Dutch Healthy Diet index sum score, a measure of adherence to the Dutch dietary guidelines 2015[39], based on a food frequency questionnaire[40]; assessed lipid-modifying, antihypertensive, and glucose-lowering medication use as part of a medication interview; measured weight, height, and waist circumference during a physical examination; calculated body mass index (BMI); measured office and 24-hour ambulatory BP; measured $HbA_{1c}$ and lipid profile in fasting venous blood samples; measured albumin excretion in two 24-hour urine collections; and calculated the estimated glomerular filtration rate (eGFR) based on serum creatinine only, as cystatin C was not presently available in this subpopulation.

## *Statistical analysis*

Normally distributed data are presented as mean and standard deviation (SD), non-normally distributed data as median and interquartile range, and categorical data as n (%). We used multiple linear regression with a complete-case approach to study the associations of daily GV with arterial measures. The crude analyses only included $SD_{CGM}$ as determinant. Model 1 was adjusted for demographics: age, sex, and education level. Model 2 was additionally adjusted for cardiovascular risk and lifestyle factors: MAP (in case of cf-PWV, carDC, and $CWS_{puls}$), office systolic BP (in case of cIMT and ABI), or braPP (in case of $CWS_{mean}$), mean heart rate (in case of cf-PWV and ABI only), BMI, total-to-high-density lipoprotein cholesterol levels, smoking status, alcohol use, and antihypertensive and lipid-modifying drug use. In order to study its contribution relative to $SD_{CGM}$, the associations were further adjusted for $MSG_{CGM}$ in an additional model (i.e., model 2 + $MSG_{CGM}$). The main regression results are presented as: regression coefficient (B) (corresponding 95% confidence interval [95%CI]), p-value.

We presumed the reliability of our model 2 + $MSG_{CGM}$ results to be negatively impacted by multicollinearity, due to the strong correlation between $SD_{CGM}$ and $MSG_{CGM}$ (rho=0.69)[41]. Hence, we additionally performed ridge regression, a L2-regularized form of linear regression (formula is provided in the Supplemental Methods), which is a valid statistical method to counter a degree of model instability caused by multicollinearity[42]. Ridge regression estimates are computed according to the combination of the residual sum of squares, characteristic of regular linear regression, and predefined penalization of the coefficients. As such, it slightly biases the regression coefficients and can strongly reduce inflated variances that arise when high levels of multicollinearity are present. We pragmatically chose the level of penalization based on the lambda (λ) required to reduce the variance inflation factor (VIF) of model 2 + $MSG_{CGM}$ back to the VIF of model 2 (or halfway back). The ridge regression results are presented as: standardized regression coefficient (st.β) (95%CI), p-value. The median st.βs (95%CIs) were estimated with use of resampling (1,000 bootstrap).

In secondary analyses, we replaced the main determinant $SD_{CGM}$ with $CV_{CGM}$ and $TIR_{CGM}$. For reasons of clarity, the regression coefficients of both indices are presented per 10% difference instead of per 1%. In order to further explore the clinical applicability of our results in the context of the International Consensus on $TIR_{CGM}$[43], we repeated the analyses with $TIR_{CGM} \geq 70\%$ (yes/no) as main determinant. In addition, we investigated whether the associations were modified by sex[44], age[25], or (type 2) diabetes status by adding interaction terms (e.g., $SD_{CGM}$*sex) to model 2.

To test the robustness of our main findings, we performed several sensitivity analyses by 1) replacing $MSG_{CGM}$ with GMS, $HbA_{1c}$ or fasting plasma glucose; 2) adding physical activity and diet as a separate model, as for these confounders relatively many missing values were observed (**Supplemental Table S6.1**); 3) adding specific variables (eGFR, urinary albumin excretion, history of CVD) as a separate model, since they may introduce overadjustment bias[45]; 4) substituting office systolic BP with ambulatory systolic BP; and 5) excluding individuals with type 1 diabetes, individuals with CGM data gaps, individuals who underwent CGM as part of a "catch-up visit", or individuals with a suboptimal CGM recording period (i.e., less than two data days)[31]. Last, we also repeated the primary analyses with $MSG_{CGM}$ as main determinant.

We considered a p-value of <0.05 statistically significant. Statistical analyses were performed using the Statistical Package for Social Sciences (version 25.0; IBM, Chicago, Illinois, USA) and the R programming language (version 3.6.1; R Foundation for Statistical Computing, Vienna, Austria) with package glmnet (version 4.0.2).

## Results

### Study population characteristics

The total CGM study population comprised 853 individuals. Because outcome and covariate data could not be obtained in all individuals (**Supplemental Figure S6.1**, **Supplemental Table S6.1**), the number of participants who were included in the different regression analyses varied (n=643–816). **Table 6.1** shows the participant characteristics of the largest sample size (i.e., ABI study population) stratified according to tertiles of $SD_{CGM}$. With higher GV, participants were older, more often male, and were generally characterized by a more unfavorable cardiometabolic profile (i.e., higher $HbA_{1c}$, BP, and BMI values and more often current smoker). GMS did not fully correspond with daily GV. Namely, 31 (17%) of the 185 individuals with type 2 diabetes were not in the highest tertile of $SD_{CGM}$, participants with prediabetes were evenly distributed between the tertiles, and 58 (13%) of the 454 individuals with NGM were not in the lowest or middle tertiles. **Supplemental Table S6.2** and **Supplemental Figures S6.2-4** additionally show that the different GMS categories have substantially overlapping $SD_{CGM}$ values.

6

Table 6.1    Participant characteristics of ankle-brachial index study population (n=816) stratified according to tertiles of $SD_{CGM}$.

| Characteristic | First $SD_{CGM}$ tertile: 0.32 – 0.72 mmol/L (n=276) | Second $SD_{CGM}$ tertile: 0.73 – 1.00 mmol/L (n=267) | Third $SD_{CGM}$ tertile: 1.01 – 4.81 mmol/L (n=273) |
|---|---|---|---|
| Demographics | | | |
| Age, y | 57.8 ± 8.9 | 59.3 ± 8.7 | 62.1 ± 7.7 |
| Women, n | 147 (53.3) | 126 (47.2) | 125 (45.8) |
| Educational level (low/medium/high), n | 63/76/137 | 86/80/101 | 107/71/95 |
| Educational level (low/medium/high), % | 22.8/27.5/49.6 | 32.2/30.0/37.8 | 39.2/26.0/34.8 |
| Glycemic parameters | | | |
| Glucose metabolism status | | | |
| NGM/PreD/T2D/T1D, n | 230/40/6/0 | 166/76/25/0 | 58/59/154/2 |
| NGM/PreD/T2D/T1D, % | 83.3/14.5/2.2/0 | 62.2/28.5/9.4/0 | 21.2/21.6/56.4/0.7 |
| Newly diagnosed T2D, n | 6 (2.2) | 18 (6.7) | 44 (16.1) |
| Fasting plasma glucose, mmol/L | 5.1 [4.9-5.5] | 5.4 [5.0-5.9] | 6.5 [5.4-7.6] |
| 2-hour post-load glucose, mmol/L | 5.5 [4.7-6.9] | 6.4 [5.2-8.0] | 10.3 [7.2-14.5] |
| $MSG_{CGM}$, mmol/L | 5.7 [5.4-6.0] | 6.0 [5.7-6.3] | 7.1 [6.4-8.1] |
| $SD_{CGM}$, mmol/L | 0.63 [0.55-0.68] | 0.84 [0.77-0.93] | 1.40 [1.17-1.86] |
| $CV_{CGM}$, % | 10.8 [9.9-11.7] | 14.0 [13.0-15.3] | 19.9 [17.5-23.9] |
| $TIR_{CGM}$, % | 100.0 [100.0-100.0] | 100.0 [99.5-100.0] | 94.6 [82.1-98.4] |
| $HbA_{1c}$, % | 5.4 [5.2-5.5] | 5.5 [5.4-5.7] | 6.0 [5.6-6.8] |
| $HbA_{1c}$, mmol/mol | 35.0 [33.0-37.0] | 37.0 [35.0-39.0] | 42.0 [38.0-51.0] |
| Diabetes medication use, n | 0 (0) | 6 (2.2) | 96 (35.2) |
| Insulin | 0 (0) | 1 (0.4) | 19 (7.0) |
| Metformin | 0 (0) | 6 (2.2) | 91 (33.3) |
| Sulfonylureas | 0 (0) | 0 (0) | 21 (7.7) |
| GLP-1 analogs | 0 (0) | 0 (0) | 4 (1.5) |
| DDP-4 inhibitors | 0 (0) | 0 (0) | 1 (0.4) |
| SGLT-2 inhibitors | 0 (0) | 0 (0) | 1 (0.4) |
| Lifestyle factors | | | |
| BMI, kg/m$^2$ | 26.1 ± 3.7 | 26.7 ± 3.9 | 28.3 ± 4.8 |
| Waist circumference (men), cm | 98.8 ± 9.9 | 100.7 ± 10.6 | 106.3 ± 12.4 |
| Waist circumference (women), cm | 87.2 ± 10.7 | 90.4 ± 11.5 | 94.2 ± 12.8 |
| Physical activity, hours/week | 12.5 [7.8-18.5] | 12.5 [7.5-19.6] | 11.5 [6.8-17.9] |
| Dutch healthy diet index, (range: 0-150) | 85.4 ± 17.3 | 84.5 ± 16.2 | 81.3 ± 14.6 |
| Alcohol use (none/low/high), n | 38/179/59 | 36/180/51 | 69/164/40 |
| Alcohol use (none/low/high), % | 13.8/64.9/21.4 | 13.5/67.4/19.1 | 25.3/60.1/14.7 |
| Smoking (never/former/current), n | 122/126/28 | 100/135/32 | 95/136/42 |
| Smoking (never/former/current), % | 44.2/45.7/10.1 | 37.5/50.6/12.0 | 34.8/49.8/15.4 |
| Cardiovascular risk factors | | | |
| History of CVD | 41 (14.9) | 28 (10.6) | 53 (19.4) |
| Office systolic BP, mmHg | 129.0 ± 17.5 | 133.3 ± 17.9 | 137.0 ± 17.9 |
| Office diastolic BP, mmHg | 73.7 ± 9.8 | 75.4 ± 10.4 | 75.9 ± 10.2 |
| Mean arterial pressure, mmHg | 95.5 ± 10.9 | 96.8 ± 10.7 | 98.6 ± 10.7 |
| Mean heart rate, beats/minute | 59.2 ± 8.1 | 60.3 ± 8.6 | 63.3 ± 8.9 |
| Antihypertensive medication use, n | 58 (21.0) | 84 (31.5) | 142 (52.0) |
| Total-to-HDL cholesterol ratio | 3.3 [2.8-4.3] | 3.6 [2.9-4.3] | 3.6 [2.8-4.3] |
| Triglycerides, mmol/L | 1.2 [0.9-1.5] | 1.3 [0.9-1.7] | 1.4 [1.0-1.9] |
| Lipid-modifying medication use, n | 31 (11.2) | 42 (15.7) | 128 (46.9) |
| eGFR, mL/min/1.73 m$^2$ | 81.8 ± 13.0 | 79.8 ± 13.8 | 80.0 ± 10.2 |
| Albuminuria, n | 7 (2.5) | 23 (8.6) | 33 (12.2) |

**Table 6.1**     (continued)

| Characteristic | First SD$_{CGM}$ tertile: 0.32 – 0.72 mmol/L (n=276) | Second SD$_{CGM}$ tertile: 0.73 – 1.00 mmol/L (n=267) | Third SD$_{CGM}$ tertile: 1.01 – 4.81 mmol/L (n=273) |
|---|---|---|---|
| Outcome measures | | | |
| Carotid-femoral pulse wave velocity (cf-PWV), m/s | 8.3 ± 1.8 | 8.5 ± 1.9 | 9.5 ± 2.5 |
| Carotid distensibility coefficient (carDC), 10$^{-3}$/kPa | 16.3 ± 5.8 | 16.5 ± 5.9 | 14.9 ± 6.1 |
| Carotid intima-media thickness (cIMT), μm | 865.6 ± 144.0 | 899.2 ± 152.3 | 906.7 ± 160.2 |
| Ankle-brachial index (ABI) | 1.14 ± 0.10 | 1.14 ± 0.10 | 1.13 ± 0.12 |
| Ankle-brachial index (ABI) < 0.9, n | 6 (2.2) | 8 (3.0) | 10 (3.7) |
| Mean circumferential wall stress (CWS$_{mean}$), kPa | 43.8 [38.1-49.5] | 44.0 [37.7-49.7] | 44.3 [37.9-52.1] |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa | 21.7 [18.6-26.1] | 22.5 [18.7-26.5] | [19.7-29.1] |

Data are reported as mean ± SD, median [interquartile range], or number (percentage %) as appropriate. Data represent the study population of participants with complete data on determinant, outcome (i.e., ABI) and confounders. NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; T1D, type 1 diabetes, MSG$_{CGM}$, mean sensor glucose; SD$_{CGM}$, standard deviation; CV$_{CGM}$, coefficient of variation; TIR$_{CGM}$, time in range; HbA1c glycated hemoglobin A1c; GLP-1 glucagon-like peptide-1; DPP-4 dipeptidase-4; SGLT2, sodium/glucose cotransporter 2; BMI, body mass index; CVD, cardiovascular disease; BP, blood pressure; HDL high-density lipoprotein; eGFR, estimated glomerular filtration rate.

6

## *Daily glucose variability and arterial stiffness*

**Figure 6.1** and **Supplemental Table S6.3** show the associations of SD$_{CGM}$ with cf-PWV and carDC estimated by use of multiple linear regression. Higher SD$_{CGM}$ was statistically significantly associated with higher cf-PWV after adjustment for demographics, cardiovascular risk factors, and lifestyle factors (model 2, B: 0.413 m/s [0.147; 0.679], p=0.003). Although, numerically, the regression estimate was attenuated by a third after additional adjustment for MSG$_{CGM}$ (model 2 + MSG$_{CGM}$, B: 0.270 m/s [-0.125; 0.666], p=0.180), the coefficients were not statistically significantly different.

   **Table 6.2** shows the fully adjusted st.βs of SD$_{CGM}$ and MSG$_{CGM}$, as estimated with ridge regression, to allow better comparison of the strength of association of both indices with cf-PWV. The coefficients were comparable and both not statistically significant (st.β: 0.065 [-0.018; 0.167], p=0.160 for SD$_{CGM}$; and st.β: 0.059 [-0.043; 0.164], p=0.272 for MSG$_{CGM}$).
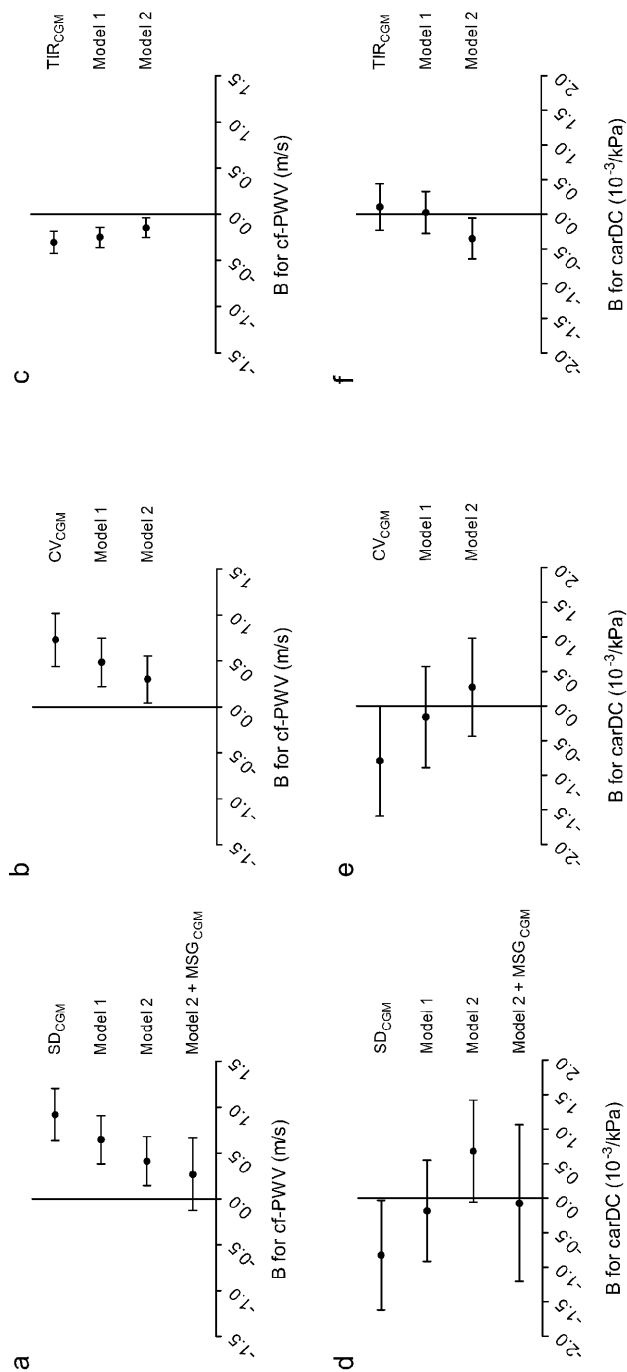
**Figure 6.1** Multivariable-adjusted associations of $SD_{CGM}$, $CV_{CGM}$, and $TIR_{CGM}$ with measures of arterial stiffness. Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 mmol/L increase in $SD_{CGM}$ or 10% increase in $CV_{CGM}$ or $TIR_{CGM}$. The panels depict the associations with carotid-femoral pulse wave velocity (cf-PWV; panels a-c) and carotid distensibility coefficient (carDC; panels d-f). Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure, mean heart rate (in case of cf-PWV only), BMI, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

**Table 6.2** Standardized regression coefficients of standard deviation and mean sensor glucose in the fully adjusted models with arterial outcome variables.

| Ridge regression penalization (λ) | SD$_{CGM}$ (st.β, 95%CI) | p-value | MSG$_{CGM}$ (st.β, 95%CI) | p-value |
|---|---|---|---|---|
| Carotid femoral pulse wave velocity* (cf-PWV), SD (n=643) | | | | |
| λ=0.11 | 0.065 (-0.018; 0.167) | 0.160 | 0.059 (-0.043; 0.164) | 0.272 |
| Carotid distensibility coefficient* (carDC), SD (n=725) | | | | |
| λ=0.12 | -0.003 (-0.097; 0.092) | 0.952 | 0.088 (-0.014; 0.184) | 0.102 |
| Carotid intima-media thickness† (cIMT), SD (n=726) | | | | |
| λ=0.12 | -0.007 (-0.123; 0.111) | 0.916 | 0.078 (-0.038; 0.207) | 0.198 |
| Ankle-brachial index† (ABI), SD (n=816) | | | | |
| λ=0.11 | -0.033 (-0.071; 0.002) | 0.060 | -0.008 (-0.032; 0.017) | 0.548 |
| Mean circumferential wall stress‡ (CWS$_{mean}$), SD (n=725) | | | | |
| λ=0.12 | -0.059 (-0.169; 0.066) | 0.318 | 0.082 (-0.044; 0.204) | 0.180 |
| Pulsatile circumferential wall stress* (CWS$_{puls}$), SD (n=725) | | | | |
| λ=0.12 | -0.045 (-0.145; 0.053) | 0.374 | 0.042 (-0.055; 0.138) | 0.410 |

Standardized regression coefficients (st.β) indicate the median difference (95% confidence interval) associated with 1 SD higher SD$_{CGM}$ or MSG$_{CGM}$. In the cf-PWV study population, 1 SD corresponds to 0.57 mmol/L for SD$_{CGM}$, 1.3 mmol/L for MSG$_{CGM}$, and 2.2 m/s for cf-PWV. In the carDC, cIMT, and CWS study populations, 1 SD corresponds to 0.57 mmol/L for SD$_{CGM}$, 1.3 mmol/L for MSG$_{CGM}$, 6.0 $10^{-3}$/kPa for carDC, 152.7 μm for cIMT, 10.2 kPa for CWS$_{mean}$, and 6.6 kPa for CWS$_{puls}$. In the ABI study population, 1 SD corresponds to 0.56 mmol/L for SD$_{CGM}$, 1.3 mmol/L for MSG$_{CGM}$, and 0.11 for ABI. The associations were adjusted for age, sex, educational level, mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, use of antihypertensive and lipid-modifying drugs, and the other CGM-assessed index. All coefficients were estimated by use of ridge regression. Point estimates and 95% confidence intervals were calculated by use of 1,000 bootstrap estimates. SD$_{CGM}$, standard deviation; st.β, standardized regression coefficient; MSG$_{CGM}$, mean sensor glucose.

In the analysis with CV$_{CGM}$ as determinant, the association with cf-PWV was statistically significant after full adjustment (model 2, B per 10% CV$_{CGM}$: 0.303 m/s [0.046; 0.559], p=0.021; **Supplemental Table S6.4**). In line with the main results, higher TIR$_{CGM}$ was independently associated with lower cf-PWV (model 2, B per 10% TIR$_{CGM}$: -0.145 m/s [-0.252; -0.038] p=0.008; **Figure 6.1**, **Supplemental Table S6.5**). Correspondingly, TIR$_{CGM}$≥70% was independently associated with lower cf-PWV (model 2, B: -1.098 m/s (-1.745; -0.451), p=0.001; **Supplemental Table 6.6**).

SD$_{CGM}$ was not associated with carDC after adjustment for demographics, cardiovascular risk factors, lifestyle factors, and MSG$_{CGM}$ (model 2 + MSG$_{CGM}$, B: -0.071 $10^{-3}$/kPa [-1.204; 1.063], p=0.903). CV$_{CGM}$ and TIR$_{CGM}$≥70% were also not associated with carDC (**Supplemental Table S6.4** and **S6.6**). Inconsistently, TIR$_{CGM}$ was independently associated with carDC (model 2, B per 10% TIR$_{CGM}$: -0.350 $10^{-3}$/kPa [-0.646; -0.055] p=0.020; **Supplemental Table S6.5**).

## Daily glucose variability and arterial structure

**Figure 6.2** and **Supplemental Table S6.3** show the associations of SD$_{CGM}$ with cIMT and ABI. SD$_{CGM}$ and cIMT were not associated after adjustment for all potential

confounders and $MSG_{CGM}$ (model 2 + $MSG_{CGM}$, B: -1.648 µm [-33.984; 30.688], p=0.920). While $CV_{CGM}$ and $TIR_{CGM}$ were not independently associated with cIMT (Supplemental Table 4 and 5), $TIR_{CGM} \geq 70\%$ was (model 2: B: -63.722 [-115.422; -12.023], p=0.016; **Supplemental Table S6.6**).

Higher $SD_{CGM}$ was statistically significantly associated with lower ABI after adjustment for demographics, but not after further adjustment for cardiovascular risk and lifestyle factors (model 2, B: -0.011 [-0.026; 0.003], p=0.126). Adjustment for $MSG_{CGM}$ numerically altered the regression coefficient, but did not affect statistical significance (model 2 + $MSG_{CGM}$, B: -0.017 [-0.039; 0.005], p=0.121). Although $CV_{CGM}$ and $TIR_{CGM}$ were not independently associated with ABI (**Supplemental Tables S6.4 and S6.5**), $TIR_{CGM} \geq 70\%$ was (model 2, B: 0.041 [0.004; 0.077], p=0.030; **Supplemental Table S6.6**).

### Daily glucose variability and circumferential wall stress

After full adjustment, $SD_{CGM}$ was not associated with $CWS_{mean}$ (model 2, B: 0.077 kPa [-1.313; 1.467], p=0.913; **Supplemental Table S6.3**) or $CWS_{puls}$ (model 2, B: -0.202 kPa [-1.019; 0.614], p=0.627; **Supplemental Table S6.3**). Further adjustment for $MSG_{CGM}$ did not materially alter the results. $CV_{CGM}$ and $TIR_{CGM}$ were not independently associated with $CWS_{mean}$ and $CWS_{puls}$ (**Supplemental Tables S6.4 and S6.5**).

### Interaction analyses

**Supplemental Table S6.7** shows all $P_{interaction}$ values for the associations between $SD_{CGM}$ and the arterial outcome measures. A statistically significant $P_{interaction}$ for age was only observed for the association between $SD_{CGM}$ and cIMT (p=0.044). The association between $SD_{CGM}$ and cIMT was stronger in women (**Supplemental Table S6.8**). Age and (type 2) diabetes status did not modify any of the associations under study (**Supplemental Tables S6.7 and S6.9**).
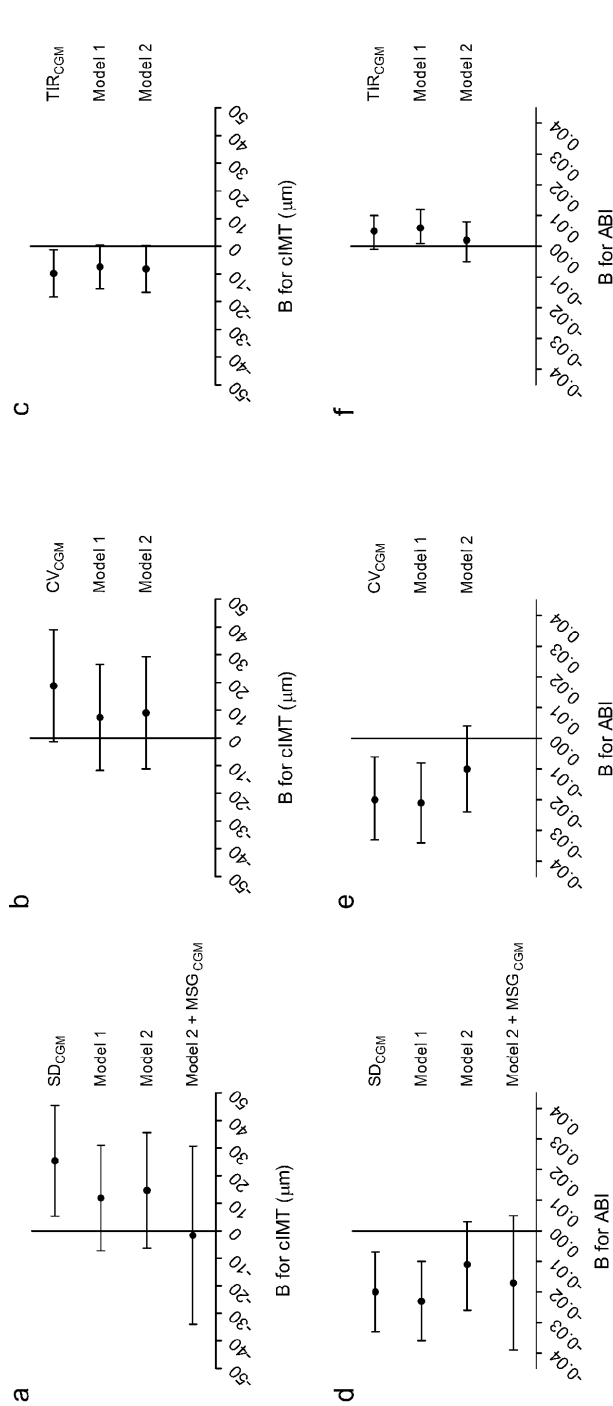
**Figure 6.2** Multivariable-adjusted associations of $SD_{CGM}$, $CV_{CGM}$, and $TIR_{CGM}$ with measures of arterial structure. Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 mmol/L increase in $SD_{CGM}$ or 10% increase in $CV_{CGM}$ or $TIR_{CGM}$. The panels depict the associations with carotid intima-media thickness (cIMT; panels a-c) and ankle-brachial index (ABI; panels d-f). Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for office systolic blood pressure, mean heart rate (in case of ABI only), BMI, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

6

### *Additional analyses*

In general, replacement of $MSG_{CGM}$ with GMS, $HbA_{1c}$ or fasting plasma glucose (**Supplemental Table S6.10**); additional adjustment for physical activity and diet (**Supplemental Table S6.11**) or for eGFR, urinary albumin excretion, and history of CVD (**Supplemental Table S6.12**); replacement of office systolic BP with ambulatory systolic BP (**Supplemental Table S6.13**); or exclusion of individuals with type 1 diabetes (**Supplemental Table S6.14**) did not materially alter the main results. The associations of $SD_{CGM}$ with arterial measures were, in general, slightly stronger after exclusion of individuals with CGM data gaps or with a suboptimal CGM recording period (**Supplemental Tables S6.15** and **S6.16**). Exclusion of individuals who underwent CGM as part of a 'catch-up visit' substantially altered the associations between $SD_{CGM}$ and the arterial measures (**Supplemental Table S6.17**). **Supplemental Table S6.18** provides the associations of $MSG_{CGM}$ with the arterial measures. Last, **Supplemental Table S6.19** shows the effects of different degrees of ridge regression penalization on the studied associations. In case of ABI, slight regularization ($\lambda=0.11$) reversed the st.$\beta$ of $MSG_{CGM}$.

## Discussion

In the present study, we investigated the cross-sectional associations of daily GV with several arterial outcome variables in a relatively large population of individuals who underwent more than 48h of CGM. Our study has two main findings. First, greater GV was linearly associated with higher cf-PWV, the gold-standard measure to assess aortic stiffness, irrespective of demographics, cardiovascular risk factors, and lifestyle factors. The observed association between $SD_{CGM}$ and cf-PWV was corroborated by our $CV_{CGM}$ and $TIR_{CGM}$ results. Notably, $SD_{CGM}$ and $MSG_{CGM}$ contributed to a similar extent to the association with cf-PWV, which suggests an equivalent pathophysiological relevance to aortic stiffness. Second, we established no consistent independent associations between indices of daily GV and the other investigated arterial measures.

Our main analyses were performed in a study population that comprises the complete spectrum of daily GV (i.e., individuals with NGM, prediabetes, type 2 diabetes, and type 1 diabetes). This approach is justified by the substantial overlap in CGM-derived indices between GMS groups, which can be appreciated from **Supplemental Table S6.2**, **Supplemental Figures S6.2-6.4**, and a recent publication on this cohort[31], and has several advantages over subgroup analyses, such as more statistical power[46] and less range restriction[47]. In addition, because no effect modification by (type 2) diabetes status was observed (**Supplemental Table S6.7**), stratification was not indicated. Further, the linearity of the observed associations between daily GV and arterial measures is consistent with work on the "ticking clock hypothesis", which postulates that hyperglycemia-induced damage is a continuous

process that starts in prediabetes, progresses with the onset of type 2 diabetes, and continues during type 2 diabetes[48, 49].

Few studies have investigated the association of CGM-measured GV with arterial measures[20-22] in concert with sufficient adjustment for potential confounders[19]. Lu et al. did not establish an association of GV with cIMT[19], which is in line with our cIMT results. Recently, we observed that the incremental glucose peak, an OGTT-based proxy of daily GV[31], was statistically significantly associated with higher cf-PWV and $CWS_{mean}$, but not with carDC, cIMT and $CWS_{puls}$[25]. Notably, our current findings are corroborated by this larger study, as the directions of the regression coefficients generally correspond, and in both instances the strongest association was found with cf-PWV. We presume that discrepancies in statistical significance are largely attributable to the almost threefold sample size differences of our previous (n=1,849-1,978) and current study populations (n=643–816). Although Lu et al. previously reported on the relation between $TIR_{CGM}$ and cIMT[19], we are the first to establish a statistically significant association of $TIR_{CGM}$ with cf-PWV.

We present—as the primary analysis—$MSG_{CGM}$-adjusted associations with $SD_{CGM}$, and —as secondary analyses— associations with the intrinsically $MSG_{CGM}$-adjusted index CV, and with $TIR_{CGM}$, which inversely reflects both mean glycaemia and GV[29]. Due to their strong correlation, it is both necessary and complex to disentangle the effects of glucose fluctuations (i.e., $SD_{CGM}$) and mean glucose (i.e., $MSG_{CGM}$)[18]. The strong correlation between $SD_{CGM}$ and $MSG_{CGM}$ (rho = 0.69), the substantial increase (121-139%) in VIF from model 2 to model 2 + $MSG_{CGM}$ (**Supplemental Table S6.3**), and the opposite directions of the regression coefficients of $SD_{CGM}$ and $MSG_{CGM}$ (e.g., ABI) all indicate the presence of multicollinearity[41]. Previous studies on other potential consequences of GV encountered similar contrariety[50,51], but insufficiently addressed this. We employed ridge regression to partially counter the potential adverse effects of multicollinearity, thereby allowing for better comparison of $SD_{CGM}$ and $MSG_{CGM}$ (**Table 6.2**). Notably in case of ABI, slight regularization (λ=0.11) reversed the st.β of $MSG_{CGM}$ (**Supplemental Table S6.19**). Interestingly, the relative contributions of $SD_{CGM}$ and $MSG_{CGM}$ differed per measure. In case of cf-PWV, the estimates were similar, which is corroborated by its independent association with $CV_{CGM}$ and $TIR_{CGM}$.

The biological mechanisms that mediate the relationship between GV and aortic stiffness require further elucidation. Several studies observed that greater GV augments inflammation and oxidative stress[52,53]. This could promote the formation of advanced glycation end-products (AGEs)[54], which have been suggested to induce arterial stiffening by accumulating in the arterial wall and forming cross-links between elastin and collagen[3-5]. An association of tissue and circulating AGEs has, thus far, only been reported with cf-PWV[55,56], which might explain our contrasting findings for the structurally different aorta (i.e., cf-PWV) and carotid artery (i.e., carDC, cIMT). In addition, cultured human fibroblasts synthesized more collagen during intermittently high glucose concentrations than during stable hyperglycaemia[57]. Higher GV could,

6

thus, lead to higher aortic stiffness by altering the elastin-collagen ratio. Additionally, large-artery endothelial dysfunction may, in part, explain the association between daily GV and cf-PWV[5,58]. Further, not only higher glucose peaks, but also more pronounced glucose nadirs could contribute to CVD development[59]. Recurrent hypoglycaemia has, for example, been shown to negatively affect certain preclinical vascular measures in individuals with type 1 diabetes[60].

Aortic stiffness, assessed via cf-PWV, is an independent determinant of CVD, cardiovascular mortality, and all-cause mortality[26-28]. We found that cf-PWV was 0.27-0.41 m/s higher per $SD_{CGM}$ unit (mmol/L) increase in the final regression models (i.e., model 2, model 2 + $MSG_{CGM}$), which corresponds with three to four years of vascular aging[61]. Hence, the 0.8 mmol/L $SD_{CGM}$ difference between the first and third $SD_{CGM}$ tertile (**Table 6.1**) can be translated to a 2- or 3-year vascular aging difference, which closely matches our recent findings on the OGTT-based incremental glucose peak[25]. Moreover, per 10% higher $TIR_{CGM}$ cf-PWV was 0.15 m/s lower, which equals minus 18 months of vascular aging[61]. After full adjustment, a $TIR_{CGM} \geq 70\%$ corresponded with 1.10 m/s lower cf-PWV, an 11-year vascular aging difference[61]. This statistically significant association remained even after further adjustment for $HbA_{1c}$ (**Supplemental Table S6.6**), which strengthens the recommendations from the International Consensus on $TIR_{CGM}$[43]. Prospective studies should further explore the observed association with aortic stiffness. If confirmative, it would be justified to study whether interventions that specifically target CGM-measured GV or $TIR_{CGM}$ (e.g., closed-loop insulin delivery systems) can improve CVD risk or incidence[16,62].

This study has certain strengths and limitations. Strengths include 1) the use of the gold-standard methods for daily GV quantification[17]; 2) the use of several, state-of-the-art arterial outcome measures; 3) the extensive participant characterization, which enabled adjustment for a broad array of possible confounders; 4) the additional use of ridge regression, which allowed us to partly address multicollinearity between $SD_{CGM}$ and $MSG_{CGM}$; and 5) the robustness of the results, i.e., the overall consistency of several sensitivity analyses, in particular for cf-PWV.

Our study has specific limitations. First, a relatively large number of individuals were excluded due to missing outcome data (**Supplemental Figure S6.1**). Although the study populations were generally comparable (**Supplemental Table S6.1**), the smaller sample size of the cf-PWV study population negatively impacted statistical power. Second, most of the individuals with diabetes had relatively well-controlled glycemic indices[31]. The consequent range restriction in the upper $SD_{CGM}$ and lower $TIR_{CGM}$ spectrum may have biased the regression estimates towards null[47]. Third, the strength of the associations may have been additionally underestimated due to individuals who underwent CGM as a catch-up visit (n=249; 29.2%)[63], as for these there was a median time of 2.1 years between CGM and the other measurements[31]. While the associations were also investigated in newly recruited individuals only (**Supplemental Table S6.17**), their applicability is substantially hampered by the smaller sample size and different

GMS distribution (i.e., lower number of individuals with prediabetes and type 2 diabetes) of the study populations. Fourth, due to the cross-sectional design of our study, we are unable to rule out reverse causality. As, for example, greater arterial stiffness has been associated with incident diabetes[64], it could increase GV. Fifth, it could be argued that adjustment for multiple testing would be required in our study[65]. However, we regarded the consequently higher chance of type 2 error undesirable[65, 66], especially in the context of a CGM-based study, which commonly has a relatively small size because of the costliness and relative invasiveness of CGM[18]. Further, it would be overly strict to enforce adjustment based on the determinants used, since $SD_{CGM}$, $CV_{CGM}$, and $TIR_{CGM}$ are conceptually and statistically related[10, 29]. Sixth, our study population is predominately Caucasian, which might limit the generalizability of our results. Last, although the models were adjusted for a large number of cardiovascular risk and lifestyle factors, residual confounding could still be present.

Our findings support the concept that greater daily GV and lower $TIR_{CGM}$ are determinants of worse aortic stiffness, but do not support this for other arterial measures. Interestingly, the fully adjusted associations of $SD_{CGM}$ and $MSG_{CGM}$ with cf-PWV were comparable. Taken together, this study further underscores the pathophysiological relevance of daily GV, irrespective of mean glycaemia, in the context of macrovascular complications. Future studies should explore this association prospectively and assess whether interventions that specifically target CGM-measured GV or $TIR_{CGM}$ can prevent CVD.

6

# References

1.  Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical Update: Cardiovascular Disease in Diabetes Mellitus: Atherosclerotic Cardiovascular Disease and Heart Failure in Type 2 Diabetes Mellitus - Mechanisms, Management, and Clinical Considerations. Circulation. 2016;133(24): 2459-502.
2.  Cai X, Zhang Y, Li M, Wu JH, Mai L, Li J, et al. Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis. BMJ. 2020;370:m2297.
3.  Stehouwer CD, Henry RM, Ferreira I. Arterial stiffness in diabetes and the metabolic syndrome: a pathway to cardiovascular disease. Diabetologia. 2008;51(4):527-39.
4.  Prenner SB, Chirinos JA. Arterial stiffness in diabetes mellitus. Atherosclerosis. 2015;238(2):370-9.
5.  Zieman SJ, Melenovsky V, Kass DA. Mechanisms, pathophysiology, and therapy of arterial stiffness. Arterioscler Thromb Vasc Biol. 2005;25(5):932-43.
6.  Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev. 2013;93(1):137-88.
7.  van Sloten TT, Henry RM, Dekker JM, Nijpels G, Unger T, Schram MT, et al. Endothelial dysfunction plays a key role in increasing cardiovascular risk in type 2 diabetes: the Hoorn study. Hypertension. 2014;64(6):1299-305.
8.  Gorst C, Kwok CS, Aslam S, Buchan I, Kontopantelis E, Myint PK, et al. Long-term Glycemic Variability and Risk of Adverse Outcomes: A Systematic Review and Meta-analysis. Diabetes Care. 2015;38(12):2354-69.
9.  Ceriello A, Monnier L, Owens D. Glycaemic variability in diabetes: clinical and therapeutic implications. Lancet Diabetes Endocrinol. 2019;7(3):221-30.
10. Rodbard D. Glucose Variability: A Review of Clinical Applications and Research Developments. Diabetes Technol Ther. 2018;20(S2):S25-S215.
11. Bancks MP, Carson AP, Lewis CE, Gunderson EP, Reis JP, Schreiner PJ, et al. Fasting glucose variability in young adulthood and incident diabetes, cardiovascular disease and all-cause mortality. Diabetologia. 2019;62(8):1366-74.
12. Slieker RC, van der Heijden A, Nijpels G, Elders PJM, t Hart LM, Beulens JWJ. Visit-to-visit variability of glycemia and vascular complications: the Hoorn Diabetes Care System cohort. Cardiovasc Diabetol. 2019;18(1):170.
13. Wang A, Liu X, Xu J, Han X, Su Z, Chen S, et al. Visit-to-Visit Variability of Fasting Plasma Glucose and the Risk of Cardiovascular Disease and All-Cause Mortality in the General Population. J Am Heart Assoc. 2017;6(12).
14. Akirov A, Diker-Cohen T, Masri-Iraqi H, Shimon I. High Glucose Variability Increases Mortality Risk in Hospitalized Patients. J Clin Endocrinol Metab. 2017;102(7):2230-41.
15. Lee DY, Han K, Park S, Yu JH, Seo JA, Kim NH, et al. Glucose variability and the risks of stroke, myocardial infarction, and all-cause mortality in individuals with diabetes: retrospective cohort study. Cardiovasc Diabetol. 2020;19(1):144.
16. Zhou Z, Sun B, Huang S, Zhu C, Bian M. Glycemic variability: adverse clinical outcomes and how to improve it? Cardiovasc Diabetol. 2020;19(1):102.
17. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care. 2017;40(12):1631-40.
18. Rodbard D. The challenges of measuring glycemic variability. J Diabetes Sci Technol. 2012;6(3): 712-5.
19. Lu J, Ma X, Shen Y, Wu Q, Wang R, Zhang L, et al. Time in Range Is Associated with Carotid Intima-Media Thickness in Type 2 Diabetes. Diabetes Technol Ther. 2019.
20. Su G, Mi S, Tao H, Li Z, Yang H, Zheng H, et al. Association of glycemic variability and the presence and severity of coronary artery disease in patients with type 2 diabetes. Cardiovasc Diabetol. 2011;10:19.
21. Mo Y, Zhou J, Li M, Wang Y, Bao Y, Ma X, et al. Glycemic variability is associated with subclinical atherosclerosis in Chinese type 2 diabetic patients. Cardiovasc Diabetol. 2013;12:15.

22.  Magri CJ, Mintoff D, Camilleri L, Xuereb RG, Galea J, Fava S. Relationship of Hyperglycaemia, Hypoglycaemia, and Glucose Variability to Atherosclerotic Disease in Type 2 Diabetes. J Diabetes Res. 2018;2018:7464320.

23.  Zhang C, Tang M, Lu X, Zhou Y, Zhao W, Liu Y, et al. Relationship of ankle-brachial index, vibration perception threshold, and current perception threshold to glycemic variability in type 2 diabetes. Medicine (Baltimore). 2020;99(12):e19374.

24.  Siegelaar SE, Kerr L, Jacober SJ, Devries JH. A decrease in glucose variability does not reduce cardiovascular event rates in type 2 diabetic patients after acute myocardial infarction: a reanalysis of the HEART2D study. Diabetes Care. 2011;34(4):855-7.

25.  Foreman YD, Brouwers MCGJ, Berendschot TTJM, van Dongen MCJM, Eussen SJPM, van Greevenbroek MMJ, et al. The oral glucose tolerance test-derived incremental glucose peak is associated with greater arterial stiffness and maladaptive arterial remodeling: The Maastricht Study. Cardiovascular Diabetology. 2019;18(1):152.

26.  Laurent S, Cockcroft J, Van Bortel L, Boutouyrie P, Giannattasio C, Hayoz D, et al. Expert consensus document on arterial stiffness: methodological issues and clinical applications. Eur Heart J. 2006;27(21):2588-605.

27.  Vlachopoulos C, Aznaouridis K, Stefanadis C. Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. J Am Coll Cardiol. 2010;55(13):1318-27.

28.  Sequi-Dominguez I, Cavero-Redondo I, Alvarez-Bueno C, Pozuelo-Carrascosa DP, Nunez de Arenas-Arroyo S, Martinez-Vizcaino V. Accuracy of Pulse Wave Velocity Predicting Cardiovascular and All-Cause Mortality. A Systematic Review and Meta-Analysis. J Clin Med. 2020;9(7).

29.  Rodbard D. Glucose Time In Range, Time Above Range, and Time Below Range Depend on Mean or Median Glucose or HbA1c, Glucose Coefficient of Variation, and Shape of the Glucose Distribution. Diabetes Technol Ther. 2020.

30.  Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.

31.  Foreman YD, Brouwers MCGJ, van der Kallen CJH, Pagen DME, van Greevenbroek MMJ, Henry RMA, et al. Glucose Variability Assessed with Continuous Glucose Monitoring: Reliability, Reference Values, and Correlations with Established Glycemic Indices-The Maastricht Study. Diabetes Technol Ther. 2020;22(5):395-403.

32.  Geijselaers SL, Sep SJ, Schram MT, van Boxtel MP, van Sloten TT, Op Het Roodt J, et al. Carotid circumferential wall stress is not associated with cognitive performance among individuals in late middle age: The Maastricht Study. Atherosclerosis. 2018;276:15-22.

33.  van Sloten TT, Czernichow S, Houben AJ, Protogerou AD, Henry RM, Muris DM, et al. Association Between Arterial Stiffness and Skin Microvascular Function: The SUVIMAX2 Study and The Maastricht Study. Am J Hypertens. 2015;28(7):868-76.

34.  Willekes C, Hoeks AP, Bots ML, Brands PJ, Willigers JM, Reneman RS. Evaluation of off-line automated intima-media thickness detection of the common carotid artery based on M-line signal processing. Ultrasound Med Biol. 1999;25(1):57-64.

35.  Reneman RS, Meinders JM, Hoeks AP. Non-invasive ultrasound in arterial wall dynamics in humans: what have we learned and what remains to be solved. Eur Heart J. 2005;26(10):960-6.

36.  Henry RM, Kostense PJ, Dekker JM, Nijpels G, Heine RJ, Kamp O, et al. Carotid arterial remodeling: a maladaptive phenomenon in type 2 diabetes but not in impaired glucose metabolism: the Hoorn study. Stroke. 2004;35(3):671-6.

37.  Aboyans V, Criqui MH, Abraham P, Allison MA, Creager MA, Diehm C, et al. Measurement and interpretation of the ankle-brachial index: a scientific statement from the American Heart Association. Circulation. 2012;126(24):2890-909.

38.  WHO. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. WHO. 2006.

39.  Looman M, Feskens EJ, de Rijk M, Meijboom S, Biesbroek S, Temme EH, et al. Development and evaluation of the Dutch Healthy Diet index 2015. Public Health Nutr. 2017;20(13):2289-99.

6

40. van Dongen MC, Wijckmans-Duysens NEG, den Biggelaar LJ, Ocke MC, Meijboom S, Brants HA, et al. The Maastricht FFQ: Development and validation of a comprehensive food frequency questionnaire for the Maastricht study. Nutrition. 2019;62:39-46.
41. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. Epidemiology (Sunnyvale). 2016;6(2).
42. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carre G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography. 2013;36(1):27-46.
43. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care. 2019;42(8):1593-603.
44. Schiebinger L, Leopold SS, Miller VM. Editorial policies for sex and gender analysis. Lancet. 2016;388(10062):2841-2.
45. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009;20(4):488-95.
46. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J. 2003;20(5):453-8.
47. Bland JM, Altman DG. Correlation in restricted ranges of data. BMJ. 2011;342:d556.
48. Haffner SM, Stern MP, Hazuda HP, Mitchell BD, Patterson JK. Cardiovascular risk factors in confirmed prediabetic individuals. Does the clock for coronary heart disease start ticking before the onset of clinical diabetes? JAMA. 1990;263(21):2893-8.
49. Wong MS, Gu K, Heng D, Chew SK, Chew LS, Tai ES. The Singapore impaired glucose tolerance follow-up study: does the ticking clock go backward as well as forward? Diabetes Care. 2003;26(11):3024-30.
50. Soupal J, Skrha J, Jr., Fajmon M, Horova E, Mraz M, Skrha J, et al. Glycemic variability is higher in type 1 diabetes patients with microvascular complications irrespective of glycemic control. Diabetes Technol Ther. 2014;16(4):198-203.
51. Rizzo MR, Marfella R, Barbieri M, Boccardi V, Vestini F, Lettieri B, et al. Relationships between daily acute glucose fluctuations and cognitive performance among aged type 2 diabetic patients. Diabetes Care. 2010;33(10):2169-74.
52. Ceriello A, Kilpatrick ES. Glycemic variability: both sides of the story. Diabetes Care. 2013;36 Suppl 2:S272-5.
53. Monnier L, Mas E, Ginet C, Michel F, Villon L, Cristol JP, et al. Activation of oxidative stress by acute glucose fluctuations compared with sustained chronic hyperglycemia in patients with type 2 diabetes. JAMA. 2006;295(14):1681-7.
54. Brownlee M. The pathobiology of diabetic complications: a unifying mechanism. Diabetes. 2005;54(6):1615-25.
55. van Eupen MG, Schram MT, van Sloten TT, Scheijen J, Sep SJ, van der Kallen CJ, et al. Skin Autofluorescence and Pentosidine Are Associated With Aortic Stiffening: The Maastricht Study. Hypertension. 2016;68(4):956-63.
56. Baumann M, Richart T, Sollinger D, Pelisek J, Roos M, Kouznetsova T, et al. Association between carotid diameter and the advanced glycation end product N-epsilon-carboxymethyllysine (CML). Cardiovasc Diabetol. 2009;8:45.
57. Jones SC, Saunders HJ, Qi W, Pollock CA. Intermittent high glucose enhances cell growth and collagen synthesis in cultured human tubulointerstitial cells. Diabetologia. 1999;42(9):1113-9.
58. Ceriello A, Esposito K, Piconi L, Ihnat MA, Thorpe JE, Testa R, et al. Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in normal and type 2 diabetic patients. Diabetes. 2008;57(5):1349-54.
59. International Hypoglycaemia Study G. Hypoglycaemia, cardiovascular disease, and mortality in diabetes: epidemiology, pathogenesis, and management. Lancet Diabetes Endocrinol. 2019;7(5):385-96.
60. Gimenez M, Gilabert R, Monteagudo J, Alonso A, Casamitjana R, Pare C, et al. Repeated episodes of hypoglycemia as a potential aggravating factor for preclinical atherosclerosis in subjects with type 1 diabetes. Diabetes Care. 2011;34(1):198-203.

61. McEniery CM, Yasmin, Hall IR, Qasem A, Wilkinson IB, Cockcroft JR, et al. Normal vascular aging: differential effects on wave reflection and aortic pulse wave velocity: the Anglo-Cardiff Collaborative Trial (ACCT). J Am Coll Cardiol. 2005;46(9):1753-60.
62. Ceriello A. Glucose Variability and Diabetic Complications: Is It Time to Treat? Diabetes Care. 2020;43(6):1169-71.
63. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. BMJ. 2010;340:c2289.
64. Muhammad IF, Borne Y, Ostling G, Kennback C, Gottsater M, Persson M, et al. Arterial Stiffness and Incidence of Diabetes: A Population-Based Cohort Study. Diabetes Care. 2017;40(12):1739-45.
65. Althouse AD. Adjust for Multiple Comparisons? It's Not That Simple. Ann Thorac Surg. 2016;101(5):1644-5.
66. Rothman KJ. No Adjustments Are Needed for Multiple Comparisons. Epidemiology. 1990;1(1):43-6.

6

# Supplemental materials

## *Supplemental methods*

*Ridge regression*

The method used in R package glmnet (version 4.0.2) estimates regression coefficients according to the formula: $Ridge = RRS + \frac{1}{n} * \lambda * \sum_{j=1}^{p} \beta_j^2$ , where RSS is the residual sum of squares, n is the sample size, $\lambda$ is the chosen amount of penalization, and $\sum_{j=1}^{p} \beta_j^2$ represents the sum of all squared regression coefficients.

*Supplemental results*
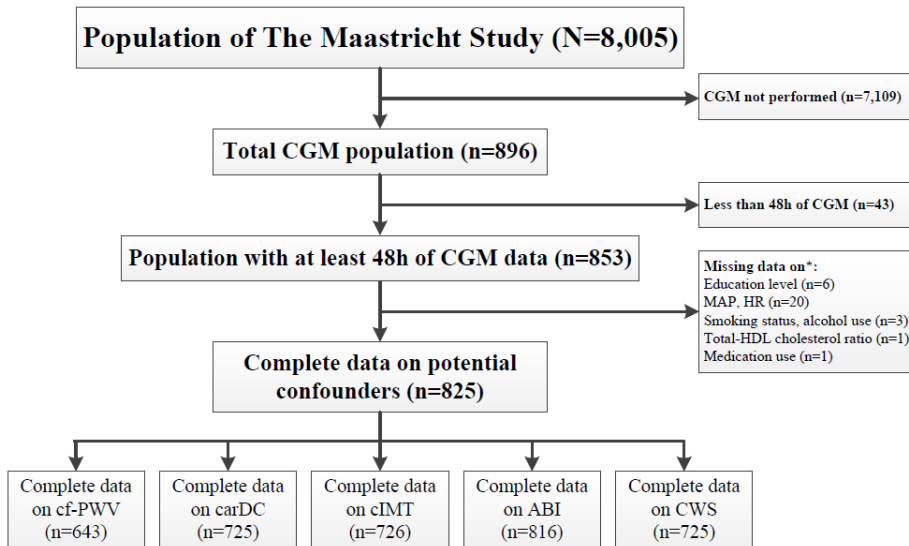


**Figure S6.1 Flowchart delineating the derivation of the study populations.** CGM, continuous glucose monitoring; MAP, mean arterial pressure; HR, heart rate; HDL, high-density lipoprotein; cf-PWV, carotid-femoral pulse wave velocity; carDC, carotid distensibility coefficient; cIMT, carotid intima-media thickness; ABI, ankle-brachial index; CWS, circumferential wall stress. * Not mutually exclusive.

6
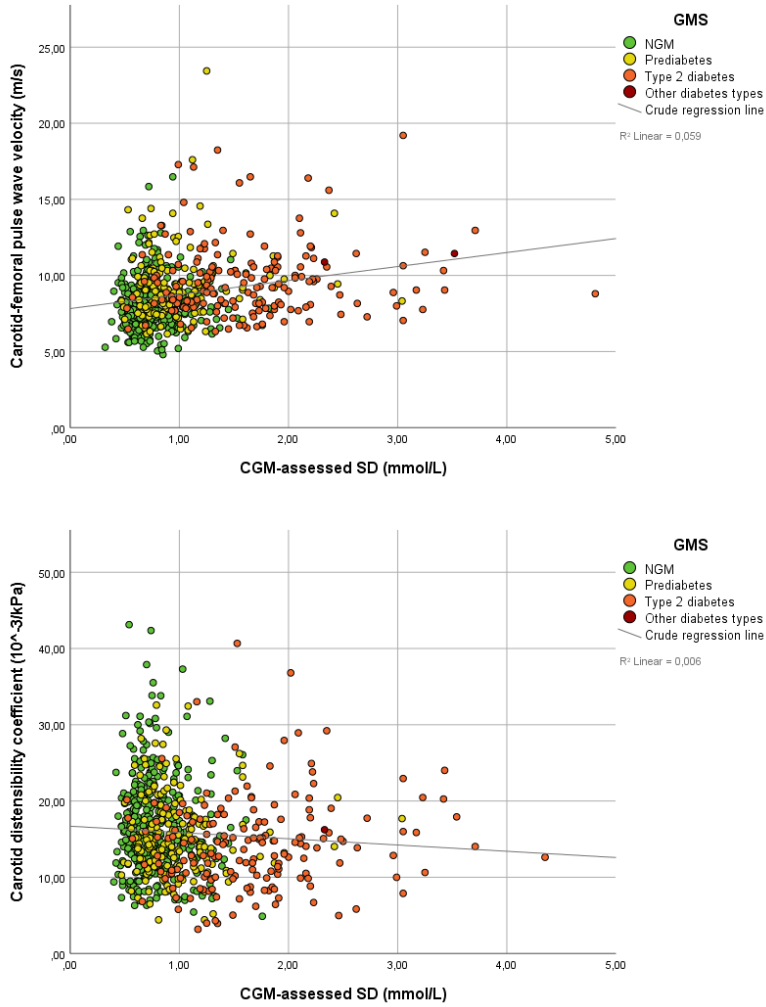
**Figure S6.2** **GMS-highlighted scatter plots of the associations of CGM-assessed SD (SD$_{CGM}$) with carotid-femoral pulse wave velocity and carotid distensibility coefficient.** CGM, continuous glucose monitoring; SD, standard deviation; GMS, glucose metabolism status; NGM, normal glucose metabolism.
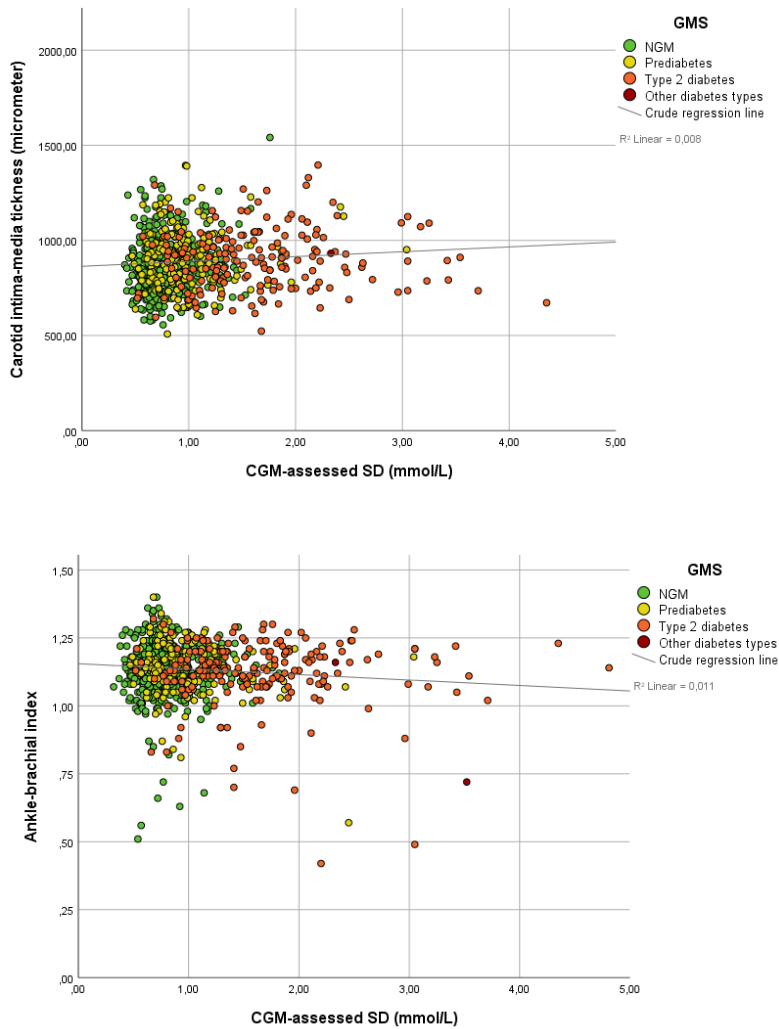
Figure S6.3  **GMS-highlighted scatter plots of the associations of CGM-assessed SD (SD$_{CGM}$) with carotid intima-media thickness and ankle-brachial index.** CGM, continuous glucose monitoring; SD, standard deviation; GMS, glucose metabolism status; NGM, normal glucose metabolism.

6
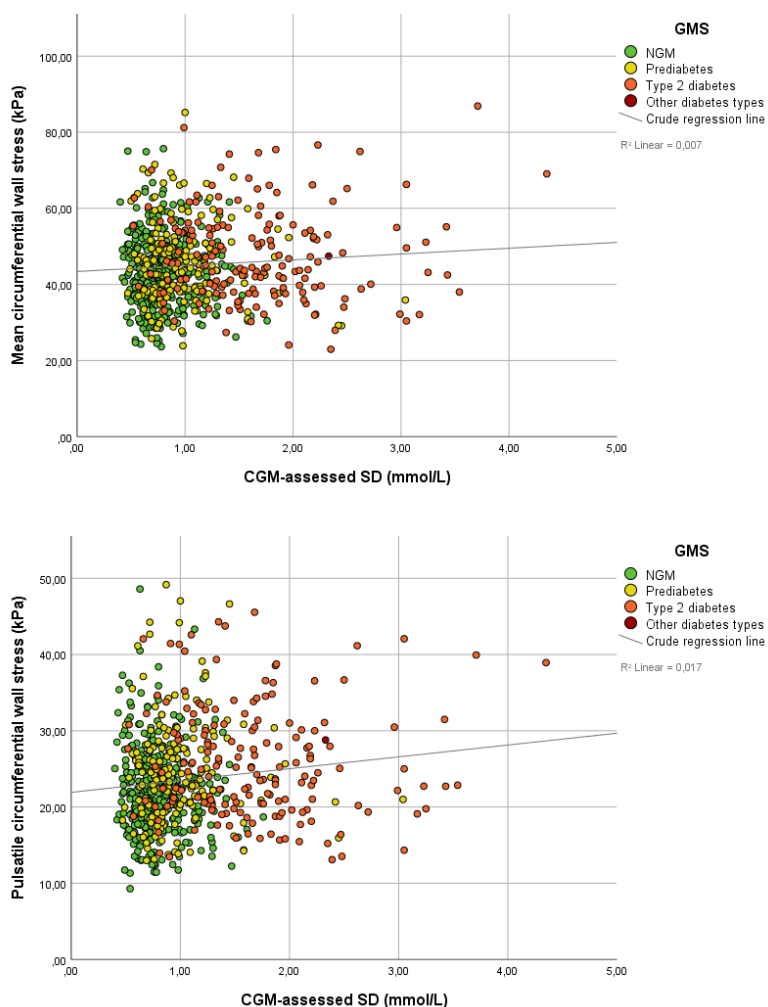
**Figure S6.4  GMS-highlighted scatter plots of the associations of CGM-assessed SD (SD$_{CGM}$) with mean and pulsatile circumferential wall stress.** CGM, continuous glucose monitoring; SD, standard deviation; GMS, glucose metabolism status; NGM, normal glucose metabolism.

**Table S6.1  Participant characteristics of the different study populations and the number of missing values.**

| Characteristic | Complete data on potential confounders (n=825) | Missing data | cf-PWV (n=643) | Missing data | carDC, CWS (n=725) | Missing data | cIMT (n=726) | Missing data | ABI (n=816) | Missing data |
|---|---|---|---|---|---|---|---|---|---|---|
| **Demographics** | | | | | | | | | | |
| Age, y | 59.7 ± 8.6 | 0 | 60.0 ± 8.5 | 0 | 60.1 ± 8.6 | 0 | 60.1 ± 8.6 | 0 | 59.7 ± 8.6 | 0 |
| Women | 403 (48.8) | 0 | 295 (45.9) | 0 | 350 (48.3) | 0 | 351 (48.3) | 0 | 398 (48.8) | 0 |
| Educational level (low/medium/high), n | 261/231/333 | 0 | 197/184/262 | 0 | 226/211/288 | 0 | 226/211/289 | 0 | 256/227/333 | 0 |
| Educational level (low/medium/high), % | 31.6/28.0/40.4 | | 30.6/28.6/40.7 | | 31.2/29.1/39.7 | | 31.1/29.1/39.8 | | 31.4/27.8/40.8 | |
| **Glycaemic parameters** | | | | | | | | | | |
| **Glucose metabolism status** | | | | | | | | | | |
| NGM/PreD/T2D/T1D, n | 455/179/189/2 | 0 | 333/148/160/2 | 0 | 384/169/171/1 | 0 | 384/169/172/1 | 0 | 454/175/185/2 | 0 |
| NGM/PreD/T2D/T1D, % | 55.2/21.7/22.9/0.2 | | 51.8/23.0/24.9/0.3 | | 53.0/23.3/23.6/0.1 | | 52.9/23.3/23.7/0.1 | | 55.6/21.4/22.7/0.2 | |
| Newly diagnosed T2D, n | 68 (8.2) | 0 | 57 (8.9) | 0 | 59 (8.1) | 0 | 59 (8.1) | 0 | 68 (8.3) | 0 |
| Fasting plasma glucose, mmol/L | 5.4 [5.0-6.2] | 0 | 5.5 [5.0-6.4] | 0 | 5.5 [5.0-6.3] | 0 | 5.5 [5.0-6.3] | 0 | 5.4 [5.0-6.2] | 0 |
| 2-hour post-load glucose, mmol/L | 6.6 [5.2-9.1] | 39 | 6.9 [5.3-9.4] | 33 | 6.8 [5.3-9.3] | 39 | 6.9 [5.3-9.3] | 39 | 6.6 [5.2-9.1] | 39 |
| $MSC_{CGM}$, mmol/L | 6.1 [5.7-6.7] | 0 | 6.1 [5.7-6.8] | 0 | 6.1 [5.7-6.7] | 0 | 6.1 [5.7-6.7] | 0 | 6.1 [5.7-6.7] | 0 |
| $SD_{CGM}$, mmol/L | 0.84 [0.68-1.19] | 0 | 0.86 [0.69-1.22] | 0 | 0.86 [0.69-1.21] | 0 | 0.86 [0.69-1.21] | 0 | 0.84 [0.68-1.17] | 0 |
| $CV_{CGM}$, % | 14.0 [11.6-17.6] | 0 | 14.3 [11.6-17.9] | 0 | 14.3 [11.7-17.9] | 0 | 14.3 [11.7-17.9] | 0 | 14.0 [11.6-17.6] | 0 |
| $TIR_{CGM}$, % | 99.8 [97.3-100.0] | 0 | 99.7 [96.5-100.0] | 0 | 99.7 [96.8-100.0] | 0 | 99.7 [96.8-100.0] | 0 | 99.8 [97.6-100.0] | 0 |
| $HbA_{1c}$, % | 5.5 [5.3-5.9] | 0 | 5.6 [5.4-5.9] | 0 | 5.6 [5.3-5.9] | 0 | 5.6 [5.4-5.9] | 0 | 5.5 [5.3-5.9] | 0 |
| $HbA_{1c}$, mmol/mol | 37.0 [34.0-41.0] | 0 | 38.0 [35.0-41.0] | 0 | 38.0 [35.0-41.0] | 0 | 38.0 [35.0-41.0] | 0 | 37.0 [34.0-41.0] | 0 |
| Diabetes medication use, n | 106 (12.8) | 0 | 90 (14.0) | 0 | 98 (13.5) | 0 | 99 (13.6) | 0 | 102 (12.5) | 0 |
| Insulin | 21 (2.5) | | 18 (2.8) | | 17 (2.3) | | 17 (2.3) | | 20 (2.5) | |
| Metformin | 100 (12.1) | | 85 (13.2) | | 94 (13.0) | | 95 (13.1) | | 97 (11.9) | |
| Sulfonylureas | 21 (2.5) | | 19 (3.0) | | 20 (2.8) | | 20 (2.8) | | 21 (2.6) | |
| GLP-1 analogs | 4 (0.5) | | 4 (0.6) | | 3 (0.4) | | 3 (0.4) | | 4 (0.5) | |
| DDP-4 inhibitors | 1 (0.1) | | 1 (0.2) | | 1 (0.1) | | 1 (0.1) | | 1 (0.1) | |
| SGLT-2 inhibitors | 1 (0.1) | | 1 (0.2) | | 1 (0.1) | | 1 (0.1) | | 1 (0.1) | |
| **Lifestyle factors** | | | | | | | | | | |
| BMI, kg/m² | 27.2 ± 4.4 | 0 | 27.3 ± 4.4 | 0 | 27.2 ± 4.4 | 0 | 27.2 ± 4.4 | 0 | 27.0 ± 4.2 | 0 |
| Waist circumference (men), cm | 102.4 ± 11.8 | 0 | 102.7 ± 11.8 | 0 | 102.5 ± 11.9 | 0 | 102.5 ± 11.9 | 0 | 102.1 ± 11.5 | 0 |
| Waist circumference (women), cm | 90.7 ± 12.2 | 0 | 91.0 ± 11.9 | 0 | 90.7 ± 12.1 | 0 | 90.7 ± 12.1 | 0 | 90.4 ± 12.0 | 0 |
| Physical activity, hours/week | 12.3 [7.5-18.3] | 95 | 12.0 [7.5-18.3] | 63 | 12.3 [7.5-18.4] | 81 | 12.3 [7.5-18.5] | 81 | 12.3 [7.5-18.4] | 92 |

6

**Table S6.1**  (continued)

| Characteristic | Complete data on potential confounders (n=825) | Missing data | cf-PWV (n=643) | Missing data | carDC, CWS (n=725) | Missing data | cIMT (n=726) | Missing data | ABI (n=816) | Missing data |
|---|---|---|---|---|---|---|---|---|---|---|
| Dutch healthy diet index, (range: 0-150) | 83.6 ± 16.1 | 200 | 83.4 ± 16.1 | 114 | 83.6 ± 15.8 | 154 | 83.6 ± 15.9 | 154 | 83.6 ± 16.0 | 199 |
| Alcohol use (none/low/high), n | 145/529/151 | 0 | 114/410/119 | 0 | 130/465/130 | 0 | 130/466/130 | 0 | 143/523/150 | 0 |
| Alcohol use (none/low/high), % | 17.6/64.1/18.3 | | 17.7/63.8/18.5 | | 17.9/64.1/19.9 | | 17.9/64.2/19.9 | | 17.5/64.1/18.4 | |
| Smoking (never/former/current), n | 321/402/102 | 0 | 254/310/79 | 0 | 278/356/91 | 0 | 279/356/91 | 0 | 317/397/102 | 0 |
| Smoking (never/former/current), % | 38.9/48.7/12.4 | | 39.5/48.2/12.3 | | 38.3/49.1/12.6 | | 38.4/49.0/12.5 | | 38.8/48.7/12.5 | |
| Cardiovascular risk factors | | | | | | | | | | |
| History of CVD | 125 (15.2) | 2 | 102 (15.9) | 2 | 109 (15.1) | 2 | 109 (15.1) | 2 | 122 (15.0) | 2 |
| Office systolic BP, mmHg | 133.2 ± 18.0 | 0 | 133.6 ± 18.1 | 0 | 133.5 ± 18.1 | 0 | 133.5 ± 18.1 | 0 | 133.1 ± 18.0 | 0 |
| Office diastolic BP, mmHg | 75.1 ± 10.2 | 0 | 75.3 ± 10.5 | 0 | 75.2 ± 10.2 | 0 | 75.2 ± 10.2 | 0 | 75.0 ± 10.2 | 0 |
| Mean arterial pressure, mmHg | 97.0 ± 10.8 | 0 | 97.0 ± 10.8 | 0 | 97.2 ± 10.9 | 0 | 97.2 ± 10.9 | 0 | 97.0 ± 10.8 | 0 |
| Mean heart rate, beats/minute | 61.0 ± 8.7 | 0 | 61.1 ± 8.7 | 0 | 61.1 ± 8.6 | 0 | 61.1 ± 8.6 | 0 | 60.9 ± 8.7 | 0 |
| Antihypertensive medication use, n | 290 (35.2) | 0 | 235 (36.5) | 0 | 261 (36.0) | 0 | 262 (36.1) | 0 | 284 (34.8) | 0 |
| Total-to-HDL cholesterol ratio | 3.5 [2.8-4.3] | 0 | 3.5 [2.9-4.4] | 0 | 3.5 [2.8-4.3] | 0 | 3.5 [2.8-4.3] | 0 | 3.5 [2.8-4.3] | 0 |
| Triglycerides, mmol/L | 1.3 [0.9-1.8] | 0 | 1.3 [0.9-1.8] | 0 | 1.3 [0.9-1.8] | 0 | 1.3 [0.9-1.8] | 0 | 1.3 [0.9-1.7] | 0 |
| Lipid-modifying medication use, n | 205 (24.8) | 0 | 171 (26.6) | 0 | 187 (25.8) | 0 | 187 (25.8) | 0 | 201 (24.6) | 0 |
| eGFR, mL/min/1.73 m² | 80.5 ± 13.7 | 0 | 80.5 ± 13.8 | 0 | 80.1 ± 13.7 | 0 | 80.1 ± 13.7 | 0 | 80.5 ± 13.7 | 0 |
| Albuminuria, n | 63 (7.9) | 3 | 52 (8.1) | 3 | 59 (8.2) | 3 | 59 (8.1) | 3 | 63 (7.8) | 3 |
| Outcome measures | | | | | | | | | | |
| cf-PWV, m/s | 8.8 ± 2.2 | 182 | 8.8 ± 2.2 | 0 | 8.8 ± 2.2 | 125 | 8.8 ± 2.2 | 125 | 8.8 ± 2.2 | 179 |
| carDC, 10⁻³/kPa | 15.8 ± 6.0 | 100 | 16.0 ± 6.1 | 43 | 15.8 ± 6.0 | 0 | 15.8 ± 6.0 | 1 | 15.8 ± 6.0 | 99 |
| cIMT, µm | 890.5 ± 152.8 | 100 | 891.9 ± 151.6 | 43 | 890.5 ± 152.8 | 0 | 890.5 ± 152.7 | 0 | 891.0 ± 153.3 | 99 |
| ABI | 1.14 ± 0.11 | 8 | 1.14 ± 0.12 | 8 | 1.14 ± 0.11 | 7 | 1.14 ± 0.11 | 7 | 1.13 ± 0.11 | 0 |
| ABI <0.9, n | 24 (2.9) | 100 | 16 (2.5) | 43 | 20 (2.8) | 0 | 20 (2.8) | 0 | 24 (2.9) | 0 |
| CWS_mean, kPa | 44.0 [37.9-50.6] | 100 | 43.7 [37.7-50.6] | 43 | 44.0 [37.9-50.6] | 0 | 44.0 [37.9-50.6] | 1 | 44.0 [37.9-50.6] | 99 |
| CWS_puls, kPa | 22.5 [19.1-27.0] | 100 | 22.4 [19.0-26.8] | 43 | 22.5 [19.1-27.0] | 0 | 22.5 [19.1-27.0] | 1 | 22.5 [19.0-27.0] | 99 |

Data are reported as mean ± SD, median [interquartile range], or number (percentage %) as appropriate. Data represent the study population of participants with complete data on determinant and confounders or on determinant, outcome and confounders (in case of cf-PWV; carDC, CWS; cIMT; and ABI). NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; T1D, type 1 diabetes; MSG$_{CGM}$, mean sensor glucose; SD$_{CGM}$, standard deviation; CV$_{CGM}$, coefficient of variation; TIR$_{CGM}$, time in range; HbA1c, glycated hemoglobin A1c; GLP-1 glucagon-like peptide-1; DPP-4 dipeptidase-4; SGLT2, sodium/glucose cotransporter 2; BMI, body mass index; CVD, cardiovascular disease; BP, blood pressure; HDL high-density lipoprotein; eGFR, estimated glomerular filtration rate; CWS$_{mean}$, mean circumferential wall stress; CWS$_{puls}$, pulsatile circumferential wall stress.

Table S6.2   GMS-stratified distributions of CGM-derived indices in the ankle-brachial index study population.

| CGM-derived indices* | | NGM (n=454) | PreD (n=175) | T2D (n=185) |
|---|---|---|---|---|
| $MSG_{CGM}$, mmol/L | Median [IQR] | 5.8 [5.5 − 6.1] | 6.2 [5.8 − 6.6] | 7.5 [6.8 − 8.7] |
| | $5^{th} − 95^{th}$ percentile | 5.15 - 6.54 | 5.25 - 7.60 | 5.73 − 11.85 |
| $SD_{CGM}$, mmol/L | Median [IQR] | 0.72 [0.61 − 0.87] | 0.89 [0.73 − 1.12] | 1.55 [1.16 − 1.98] |
| | $5^{th} − 95^{th}$ percentile | 0.48 - 1.28 | 0.60 - 1.61 | 0.79 - 3.13 |
| $CV_{CGM}$, % | Median [IQR] | 12.6 [10.7 − 14.9] | 14.7 [12.2 − 17.5] | 19.3 [16.1 − 24.1] |
| | $5^{th} − 95^{th}$ percentile | 8.51 - 20.57 | 10.13 − 24.08 | 12.42 − 33.36 |
| $TIR_{CGM}$, % | Median [IQR] | 100.0 [99.5 − 100.0] | 99.8 [98.5 − 100.0] | 91.8 [79.1 − 98.3] |
| | $5^{th} − 95^{th}$ percentile | 95.91 − 100.0 | 87.96 − 100.0 | 24.18 − 100.0 |

GMS, glucose metabolism status; CGM, continuous glucose monitoring; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; $MSG_{CGM}$, mean sensor glucose; $SD_{CGM}$, standard deviation; $CV_{CGM}$, coefficient of variation; $TIR_{CGM}$, time in range; IQR, interquartile range. *Because of the small sample size (n=2), these values are not reported for individuals with type 1 diabetes.

6

**Table S6.3** Multivariable-adjusted associations of daily glucose variability (expressed as $SD_{CGM}$) with arterial outcome variables.

| Model | B (95%CI) | p-value | VIF |
|---|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=643) | | | |
| Crude | 0.920 (0.636; 1.205) | <0.001 | 1 |
| Model 1 | 0.647 (0.384; 0.909) | <0.001 | 1.044 |
| Model 2* | 0.413 (0.147; 0.679) | 0.002 | 1.322 |
| Model 2 + $MSG_{CGM}$ | 0.270 (-0.125; 0.666) | 0.180 | 2.918 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=725) | | | |
| Crude | -0.822 (-1.615; -0.029) | 0.042 | 1 |
| Model 1 | -0.184 (-0.917; 0.549) | 0.622 | 1.036 |
| Model 2* | 0.684 (-0.052; 1.420) | 0.068 | 1.291 |
| Model 2 + $MSG_{CGM}$ | -0.071 (-1.204; 1.063) | 0.903 | 3.070 |
| Carotid intima-media thickness (cIMT), μm (n=726) | | | |
| Crude | 25.441 (5.286; 45.595) | 0.013 | 1 |
| Model 1 | 11.907 (-7.253; 31.066) | 0.223 | 1.036 |
| Model 2† | 14.679 (-6.257; 35.615) | 0.169 | 1.286 |
| Model 2 + $MSG_{CGM}$ | -1.648 (-33.984; 30.688) | 0.920 | 3.071 |
| Ankle-brachial index (ABI) (n=816) | | | |
| Crude | -0.020 (-0.033; -0.007) | 0.003 | 1 |
| Model 1 | -0.023 (-0.036; -0.010) | <0.001 | 1.048 |
| Model 2† | -0.011 (-0.026; 0.003) | 0.126 | 1.339 |
| Model 2 + $MSG_{CGM}$ | -0.017 (-0.039; 0.005) | 0.121 | 3.001 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=725) | | | |
| Crude | 1.530 (0.183; 2.877) | 0.026 | 1 |
| Model 1 | 1.009 (-0.320; 2.338) | 0.136 | 1.036 |
| Model 2‡ | 0.077 (-1.313; 1.467) | 0.913 | 1.287 |
| Model 2 + $MSG_{CGM}$ | -1.126 (-3.271; 1.019) | 0.303 | 3.070 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=725) | | | |
| Crude | 1.551 (0.690; 2.413) | <0.001 | 1 |
| Model 1 | 1.014 (0.161; 1.867) | 0.020 | 1.036 |
| Model 2* | -0.202 (-1.019; 0.614) | 0.627 | 1.291 |
| Model 2 + $MSG_{CGM}$ | -0.602 (-1.862; 0.658) | 0.349 | 3.070 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose. VIF, variance inflation factor

**Table S6.4**  Multivariable-adjusted associations of daily glucose variability (expressed as $CV_{CGM}$) with arterial outcome variables.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=643) | | |
| Crude | 0.732 (0.443; 1.021) | <0.001 |
| Model 1 | 0.488 (0.223; 0.752) | <0.001 |
| Model 2* | 0.303 (0.046; 0.559) | 0.021 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=725) | | |
| Crude | -0.790 (-1.587; 0.006) | 0.052 |
| Model 1 | -0.160 (-0.893; 0.573) | 0.669 |
| Model 2* | 0.274 (-0.436; 0.984) | 0.449 |
| Carotid intima-media thickness (cIMT), µm (n=726) | | |
| Crude | 18.784 (-1.497; 39.064) | 0.069 |
| Model 1 | 7.487 (-11.684; 26.659) | 0.443 |
| Model 2† | 9.021 (-11.182; 29.223) | 0.381 |
| Ankle-brachial index (ABI) (n=816) | | |
| Crude | -0.020 (-0.033; -0.006) | 0.003 |
| Model 1 | -0.021 (-0.034; -0.008) | 0.002 |
| Model 2† | -0.010 (-0.024; 0.004) | 0.166 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=725) | | |
| Crude | 0.494 (-0.863; 1.851) | 0.475 |
| Model 1 | 0.131 (-1.200; 1.462) | 0.847 |
| Model 2‡ | -0.440 (-1.779; 0.899) | 0.519 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=725) | | |
| Crude | 0.919 (0.049; 1.789) | 0.038 |
| Model 1 | 0.436 (-0.419; 1.292) | 0.317 |
| Model 2* | -0.388 (-1.174; 0.398) | 0.333 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 10% increase in $CV_{CGM}$. Crude: $CV_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs.

6

**Table S6.5** Multivariable-adjusted associations of time in range (TIR$_{CGM}$) with arterial outcome variables.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=643) | | |
|   Crude | -0.303 (-0.424; -0.182) | <0.001 |
|   Model 1 | -0.247 (-0.356; -0.139) | <0.001 |
|   Model 2* | -0.145 (-0.252; -0.038) | 0.008 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=725) | | |
|   Crude | 0.107 (-0.227; 0.441) | 0.529 |
|   Model 1 | 0.028 (-0.277; 0.333) | 0.857 |
|   Model 2* | -0.350 (-0.646; -0.055) | 0.020 |
| Carotid intima-media thickness (cIMT), μm (n=726) | | |
|   Crude | -9.743 (-18.211; -1.276) | 0.024 |
|   Model 1 | -7.380 (-15.336; 0.576) | 0.069 |
|   Model 2† | -8.144 (-16.563; 0.275) | 0.058 |
| Ankle-brachial index (ABI) (n=816) | | |
|   Crude | 0.005 (-0.001; 0.010) | 0.118 |
|   Model 1 | 0.006 (0.001; 0.012) | 0.026 |
|   Model 2† | 0.002 (-0.005; 0.008) | 0.620 |
| Mean circumferential wall stress (CWS$_{mean}$), kPa (n=725) | | |
|   Crude | -0.785 (-1.349; -0.220) | 0.007 |
|   Model 1 | -0.564 (-1.116; -0.013) | 0.045 |
|   Model 2‡ | -0.179 (-0.739; 0.380) | 0.530 |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa (n=725) | | |
|   Crude | -0.552 (-0.915; -0.190) | 0.003 |
|   Model 1 | -0.433 (-0.788; -0.079) | 0.017 |
|   Model 2* | 0.049 (-0.280; 0.378) | 0.768 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 10% increase in TIR$_{CGM}$. Crude: TIR$_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs.

**Table S6.6** Multivariable-adjusted associations of time in range above 70% (TIR$_{CGM}$≥70%) with arterial outcome variables.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=643) | | |
| Crude | -1.943 (-2.707; -1.179) | <0.001 |
| Model 1 | -1.732 (-2.416; -1.048) | <0.001 |
| Model 2* | -1.098 (-1.745; -0.451) | 0.001 |
| Model 2 + HbA1c | -0.775 (-1.504; -0.047) | 0.037 |
| Carotid distensibility coefficient (carDC), 10$^{-3}$/kPa (n=725) | | |
| Crude | 0.618 (-1.513; 2.749) | 0.569 |
| Model 1 | 0.455 (-1.482; 2.393) | 0.645 |
| Model 2* | -1.759 (-3.583; 0.066) | 0.059 |
| Model 2 + HbA1c | -1.828 (-3.886; 0.230) | 0.082 |
| Carotid intima-media thickness (cIMT), µm (n=726) | | |
| Crude | -69.366 (-123.377; -15.354) | 0.012 |
| Model 1 | -62.914 (-113.401; -12.426) | 0.015 |
| Model 2† | -63.722 (-115.422; -12.023) | 0.016 |
| Model 2 + HbA1c | -48.116 (-106.285; 10.054) | 0.105 |
| Ankle-brachial index (ABI) (n=816) | | |
| Crude | 0.055 (0.018; 0.091) | 0.003 |
| Model 1 | 0.063 (0.028; 0.099) | 0.001 |
| Model 2† | 0.041 (0.004; 0.077) | 0.030 |
| Model 2 + HbA1c | 0.045 (0.004; 0.086) | 0.037 |
| Mean circumferential wall stress (CWS$_{mean}$), kPa (n=725) | | |
| Crude | -4.480 (-8.087; -0.872) | 0.015 |
| Model 1 | -3.400 (-6.909; 0.108) | 0.057 |
| Model 2‡ | -1.394 (-4.834; 2.047) | 0.427 |
| Model 2 + HbA1c | -2.242 (-6.119; 1.635) | 0.257 |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa (n=725) | | |
| Crude | -2.686 (-5.006; -0.365) | 0.023 |
| Model 1 | -2.224 (-4.481; 0.203) | 0.053 |
| Model 2* | 0.523 (-1.502; 2.549) | 0.612 |
| Model 2 + HbA1c | 0.374 (-1.911; 2.659) | 0.748 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 10% increase in TIR$_{CGM}$. Crude: TIR$_{CGM}$≥70%. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + HbA1c: additionally adjusted for HbA1c.

6

**Table S6.7** P-values for interaction for sex, age, and type 2 diabetes status for the associations between $SD_{CGM}$ and arterial outcome variables.

| Outcome | Sex | Age | Diabetes status[a] | Type 2 diabetes status[b] |
|---|---|---|---|---|
| cf-PWV* | 0.41 | 0.10 | 0.57 | 0.41 |
| carDC* | 0.07 | 0.69 | 0.96 | 0.93 |
| cIMT[†] | 0.044 | 0.94 | 0.34 | 0.36 |
| ABI[†] | 0.68 | 0.80 | 0.86 | 0.15 |
| $CWS_{mean}$[‡] | 0.06 | 0.60 | 0.54 | 0.46 |
| $CWS_{puls}$* | 0.07 | 0.67 | 0.99 | 0.90 |

All models were adjusted for age, education level, mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. a For the associations with '$SD_{CGM}$*diabetes status' as interaction term, the associations were additionally adjusted for diabetes status. Bold denotes statistical significance; b For the associations with '$SD_{CGM}$*type 2 diabetes status' as interaction term, the associations were additionally adjusted for type 2 diabetes status. Bold denotes statistical significance.

**Table S6.8** Sex-stratified associations of daily glucose variability (expressed as $SD_{CGM}$) with carotid intima-media thickness.

| Model | B (95%CI) | p-value | B (95%CI) | p-value |
|---|---|---|---|---|
| | *cIMT, men (n=375)* | | *cIMT, women (n=351)* | |
| Crude | 3.693 (-23.829; 31.215) | 0.792 | 51.562 (21.871; 81.254) | 0.001 |
| Model 1 | -7.205 (-33.606; 19.197) | 0.592 | 39.873 (11.918; 67.829) | 0.005 |
| Model 2 | 1.003 (-27.653; 29.659) | 0.945 | 33.853 (2.814; 64.891) | 0.033 |
| Model 2 + $MSG_{CGM}$ | -7.146 (-51.779; 37.487) | 0.753 | 1.448 (-46.749; 49.644) | 0.953 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age and education level. Model 2: additionally adjusted for office systolic blood pressure, body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

**Table S6.9** Diabetes status-stratified associations of daily glucose variability (expressed as $SD_{CGM}$) with arterial outcome measures.

| Model | B (95%CI) | p-value | B (95%CI) | p-value |
|---|---|---|---|---|
| cf-PWV, m/s | No diabetes (n=481) | | Diabetes (n=162) | |
| Crude | 1.031 (0.450; 1.613) | 0.001 | 0.271 (-0.260; 0.802) | 0.316 |
| Model 1 | 0.483 (-0.057; 1.023) | 0.079 | 0.345 (-0.148; 0.839) | 0.169 |
| Model 2* | 0.505 (-0.008; 1.017) | 0.054 | 0.276 (-0.171; 0.723) | 0.224 |
| Model 2 + $MSG_{CGM}$ | 0.249 (-0.373; 0.872) | 0.432 | 0.367 (-0.247; 0.982) | 0.239 |
| carDC, $10^{-3}$/kPa | No diabetes (n=553) | | Diabetes (n=172) | |
| Crude | -1.096 (-2.794; 0.602) | 0.205 | 0.947 (-0.375; 2.270) | 0.159 |
| Model 1 | 0.820 (-0.727; 2.368) | 0.298 | 0.576 (-0.733; 1.884) | 0.386 |
| Model 2* | 1.164 (-0.278; 2.605) | 0.113 | 0.868 (-0.322; 2.058) | 0.152 |
| Model 2 + $MSG_{CGM}$ | 0.965 (-0.779; 2.708) | 0.278 | -0.575 (-2.254; 1.103) | 0.499 |
| cIMT, μm | No diabetes (n=553) | | Diabetes (n=173) | |
| Crude | 69.320 (26.604; 112.037) | 0.002 | 4.316 (-30.835; 39.467) | 0.809 |
| Model 1 | 36.178 (-4.282; 76.638) | 0.080 | 13.354 (-19.747; 46.455) | 0.427 |
| Model 2† | 37.967 (-3.157; 79.091) | 0.070 | 13.855 (-19.914; 47.625) | 0.419 |
| Model 2 + $MSG_{CGM}$ | 18.148 (-31.489; 67.784) | 0.473 | -7.393 (-56.083; 41.297) | 0.765 |
| ABI | No diabetes (n=629) | | Diabetes (n=187) | |
| Crude | -0.026 (-0.053; 0.001) | 0.060 | -0.014 (-0.040; 0.011) | 0.289 |
| Model 1 | -0.023 (-0.049; 0.004) | 0.091 | -0.015 (-0.041; 0.011) | 0.252 |
| Model 2† | -0.013 (-0.039; 0.014) | 0.354 | -0.007 (-0.035; 0.020) | 0.588 |
| Model 2 + $MSG_{CGM}$ | -0.013 (-0.045; 0.019) | 0.425 | -0.017 (-0.055; 0.020) | 0.363 |
| $CWS_{mean}$, kPa | No diabetes (n=553) | | Diabetes (n=172) | |
| Crude | -1.280 (-3.991; 1.431) | 0.354 | -0.032 (-2.661; 2.124) | 0.981 |
| Model 1 | -1.523 (-4.232; 1.187) | 0.270 | -0.434 (-2.992; 2.124) | 0.738 |
| Model 2‡ | -1.508 (-4.094; 1.078) | 0.253 | -0.196 (-2.741; 2.349) | 0.879 |
| Model 2 + $MSG_{CGM}$ | -0.955 (-4.083; 2.172) | 0.549 | -1.763 (-5.421; 1.895) | 0.343 |
| $CWS_{puls}$, kPa | No diabetes (n=553) | | Diabetes (n=172) | |
| Crude | 0.549 (-1.203; 2.300) | 0.539 | -0.344 (-1.970; 1.282) | 0.677 |
| Model 1 | -0.558 (-2.295; 1.179) | 0.528 | -0.315 (-1.961; 1.330) | 0.706 |
| Model 2* | -0.876 (-2.412; 0.660) | 0.263 | -0.725 (-2.218; 0.768) | 0.339 |
| Model 2 + $MSG_{CGM}$ | -0.698 (-2.556; 1.160) | 0.461 | -1.069 (-3.211; 1.072) | 0.326 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

6

**Table S6.10** Multivariable-adjusted associations of daily glucose variability (expressed as $SD_{CGM}$) with arterial outcome variables after adjustment for glucose metabolism status, $HbA_{1c}$, or fasting plasma glucose instead of mean sensor glucose.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=643) | | |
| Model 2 + GMS | 0.260 (-0.058; 0.579) | 0.109 |
| Model 2 + HbA$_{1c}$ | 0.244 (-0.081; 0.569) | 0.140 |
| Model 2 + FPG | 0.294 (-0.014; 0.603) | 0.061 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=725) | | |
| Model 2 + GMS | 0.827 (-0.062; 1.716) | 0.068 |
| Model 2 + HbA$_{1c}$ | 0.792 (-0.116; 1.700) | 0.087 |
| Model 2 + FPG | 0.687 (-0.157; 1.531) | 0.111 |
| Carotid intima-media thickness (cIMT), μm (n=726) | | |
| Model 2 + GMS | 22.231 (-3.061; 47.524) | 0.085 |
| Model 2 + HbA$_{1c}$ | 2.242 (-23.487; 27.972) | 0.864 |
| Model 2 + FPG | 9.738 (-14.259; 33.735) | 0.426 |
| Ankle-brachial index (ABI) (n=816) | | |
| Model 2 + GMS | -0.009 (-0.026; 0.008) | 0.307 |
| Model 2 + HbA$_{1c}$ | -0.013 (-0.031; 0.004) | 0.141 |
| Model 2 + FPG | -0.009 (-0.025; 0.008) | 0.290 |
| Mean circumferential wall stress (CWS$_{mean}$), kPa (n=725) | | |
| Model 2 + GMS | -0.670 (-2.349; 1.010) | 0.434 |
| Model 2 + HbA$_{1c}$ | 0.408 (-1.307; 2.123) | 0.641 |
| Model 2 + FPG | 0.275 (-1.321; 1.871) | 0.735 |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa (n=725) | | |
| Model 2 + GMS | -0.868 (-1.853; 0.117) | 0.084 |
| Model 2 + HbA$_{1c}$ | -0.130 (-1.138; 0.879) | 0.801 |
| Model 2 + FPG | -0.147 (-1.084; 0.791) | 0.759 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Model 2 + represents the fully adjusted model that was additionally adjusted for glucose metabolism status (GMS), $HbA_{1c}$, or fasting plasma glucose (FPG).

**Table S6.11** Multivariable-adjusted associations of $SD_{CGM}$ with arterial outcome variables, additionally adjusted for physical activity and Dutch healthy diet adherence.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=487) | | |
| Crude | 0.970 (0.639; 1.301) | <0.001 |
| Model 1 | 0.669 (0.356; 0.981) | <0.001 |
| Model 2* | 0.487 (0.169; 0.804) | 0.003 |
| Model 3 | 0.487 (0.169; 0.805) | 0.003 |
| Model 3 + $MSG_{CGM}$ | 0.461 (-0.014; 0.935) | 0.057 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=519) | | |
| Crude | -0.895 (-1.806; 0.017) | 0.054 |
| Model 1 | -0.256 (-1.112; 0.601) | 0.558 |
| Model 2* | 0.628 (-0.219; 1.475) | 0.146 |
| Model 3 | 0.641 (-0.209; 1.490) | 0.139 |
| Model 3 + $MSG_{CGM}$ | 0.866 (-0.461; 2.194) | 0.200 |
| Carotid intima-media thickness (cIMT), µm (n=520) | | |
| Crude | 18.139 (-5.039; 41.317) | 0.125 |
| Model 1 | 3.957 (-18.619; 26.534) | 0.731 |
| Model 2† | 6.775 (-18.133; 31.683) | 0.593 |
| Model 3 | 7.294 (-17.684; 32.271) | 0.566 |
| Model 3 + $MSG_{CGM}$ | -10.304 (-49.368; 28.760) | 0.605 |
| Ankle-brachial index (ABI) (n=562) | | |
| Crude | -0.025 (-0.040; -0.009) | 0.002 |
| Model 1 | -0.029 (-0.044; -0.013) | <0.001 |
| Model 2† | -0.014 (-0.031; 0.003) | 0.115 |
| Model 3 | -0.014 (-0.031; 0.003) | 0.112 |
| Model 3 + $MSG_{CGM}$ | -0.021 (-0.047; 0.005) | 0.114 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=519) | | |
| Crude | 2.785 (1.184; 4.386) | 0.001 |
| Model 1 | 1.952 (0.348; 3.556) | 0.017 |
| Model 2‡ | 1.242 (-0.445; 2.928) | 0.149 |
| Model 3 | 1.227 (-0.465; 2.918) | 0.155 |
| Model 3 + $MSG_{CGM}$ | -0.040 (-2.688; 2.607) | 0.976 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=519) | | |
| Crude | 2.024 (0.978; 3.071) | <0.001 |
| Model 1 | 1.414 (0.367; 2.462) | 0.008 |
| Model 2* | 0.110 (-0.888; 1.109) | 0.828 |
| Model 3 | 0.059 (-0.938; 1.056) | 0.908 |
| Model 3 + $MSG_{CGM}$ | -0.605 (-2.161; 0.951) | 0.445 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 3: additionally adjusted for physical activity and Dutch healthy diet adherence. Model 3 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose. Of note, the alcohol component of the Dutch healthy diet adherence index was subtracted from the sum score prior to performing the regression analyses to avoid multicollinearity.

6

**Table S6.12** Multivariable-adjusted associations of $SD_{CGM}$ with arterial outcome variables, additionally adjusted for history of cardiovascular disease, estimated glomerular filtration rate, and urinary albumin excretion.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=638) | | |
| Crude | 0.885 (0.599; 1.171) | < 0.001 |
| Model 1 | 0.609 (0.347; 0.871) | < 0.001 |
| Model 2* | 0.390 (0.124; 0.656) | 0.004 |
| Model 3 | 0.315 (0.048; 0.583) | 0.021 |
| Model 3 + $MSG_{CGM}$ | 0.225 (-0.169; 0.618) | 0.263 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=720) | | |
| Crude | -0.832 (-1.631; -0.033) | 0.041 |
| Model 1 | -0.187 (-0.925; 0.551) | 0.620 |
| Model 2* | 0.668 (-0.072; 1.409) | 0.077 |
| Model 3 | 0.697 (-0.061; 1.455) | 0.071 |
| Model 3 + $MSG_{CGM}$ | -0.015 (-1.167; 1.138) | 0.980 |
| Carotid intima-media thickness (cIMT), μm (n=721) | | |
| Crude | 25.173 (4.872; 45.474) | 0.015 |
| Model 1 | 11.635 (-7.671; 30.942) | 0.237 |
| Model 2† | 14.680 (-6.391; 35.751) | 0.172 |
| Model 3 | 15.638 (-5.910; 37.187) | 0.155 |
| Model 3 + $MSG_{CGM}$ | 1.648 (-31.181; 34.477) | 0.922 |
| Ankle-brachial index (ABI) (n=811) | | |
| Crude | -0.021 (-0.034; -0.008) | 0.002 |
| Model 1 | -0.024 (-0.037; -0.011) | < 0.001 |
| Model 2† | -0.012 (-0.027; 0.002) | 0.094 |
| Model 3 | -0.007 (-0.022; 0.007) | 0.333 |
| Model 3 + $MSG_{CGM}$ | -0.012 (-0.034; 0.009) | 0.258 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=720) | | |
| Crude | 1.430 (0.079; 2.782) | 0.038 |
| Model 1 | 0.911 (-0.423; 2.245) | 0.180 |
| Model 2‡ | -0.036 (-1.431; 1.359) | 0.960 |
| Model 3 | -0.129 (-1.556; 1.297) | 0.859 |
| Model 3 + $MSG_{CGM}$ | -1.300 (-3.472; 0.872) | 0.240 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=720) | | |
| Crude | 1.528 (0.661; 2.395) | 0.001 |
| Model 1 | 0.989 (0.131; 1.846) | 0.024 |
| Model 2* | -0.215 (-1.036; 0.605) | 0.607 |
| Model 3 | -0.306 (-1.143; 0.532) | 0.474 |
| Model 3 + $MSG_{CGM}$ | -0.787 (-2.062; 0.488) | 0.226 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 3: additionally adjusted for history of cardiovascular disease, estimated glomerular filtration rate, and urinary albumin excretion. Model 3 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

**Table S6.13** Multivariable-adjusted associations of $SD_{CGM}$ with arterial structure adjusted for ambulatory systolic blood pressure instead of office systolic blood pressure.

| Model | B (95%CI) | p-value |
| --- | --- | --- |
| Carotid intima-media thickness (cIMT), μm (n=649) | | |
| Crude | 29.322 (7.349; 51.294) | 0.009 |
| Model 1 | 15.560 (-5.410; 36.530) | 0.146 |
| Model 2 | 18.314 (-4.281; 40.909) | 0.112 |
| Model 2 + $MSG_{CGM}$ | 8.592 (-25.341; 42.525) | 0.619 |
| Ankle-brachial index (ABI) (n=728) | | |
| Crude | -0.026 (-0.040; -0.012) | <0.001 |
| Model 1 | -0.030 (-0.044; -0.016) | <0.001 |
| Model 2 | -0.019 (-0.034; -0.004) | 0.015 |
| Model 2 + $MSG_{CGM}$ | -0.021 (-0.043; 0.001) | 0.062 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for ambulatory systolic blood pressure and heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

6

**Table S6.14** Multivariable-adjusted associations of $SD_{CGM}$ with arterial outcome variables after exclusion of individuals with type 1 diabetes.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=641) | | |
| Crude | 0.910 (0.619; 1.200) | <0.001 |
| Model 1 | 0.629 (0.362; 0.896) | <0.001 |
| Model 2* | 0.402 (0.132; 0.673) | 0.004 |
| Model 2 + $MSG_{CGM}$ | 0.256 (-0.146; 0.658) | 0.211 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=724) | | |
| Crude | -0.831 (-1.627; -0.034) | 0.041 |
| Model 1 | -0.187 (-0.923; 0.549) | 0.618 |
| Model 2* | 0.679 (-0.058; 1.416) | 0.071 |
| Model 2 + $MSG_{CGM}$ | -0.068 (-1.203; 1.067) | 0.906 |
| Carotid intima-media thickness (cIMT), μm (n=725) | | |
| Crude | 25.389 (5.144; 45.634) | 0.014 |
| Model 1 | 11.724 (-7.511; 30.960) | 0.232 |
| Model 2† | 14.708 (-6.267; 35.682) | 0.169 |
| Model 2 + $MSG_{CGM}$ | -1.779 (-34.168; 30.610) | 0.914 |
| Ankle-brachial index (ABI) (n=814) | | |
| Crude | -0.017 (-0.030; -0.003) | 0.014 |
| Model 1 | -0.020 (-0.033; -0.007) | 0.003 |
| Model 2† | -0.007 (-0.022; 0.007) | 0.313 |
| Model 2 + $MSG_{CGM}$ | -0.011 (-0.033; 0.011) | 0.328 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=724) | | |
| Crude | 1.527 (0.174; 2.880) | 0.027 |
| Model 1 | 1.026 (-0.308; 2.360) | 0.132 |
| Model 2‡ | 0.103 (-1.289; 1.495) | 0.884 |
| Model 2 + $MSG_{CGM}$ | -1.170 (-3.317; 0.977) | 0.285 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=724) | | |
| Crude | 1.532 (0.667; 2.397) | 0.001 |
| Model 1 | 0.999 (0.143; 1.855) | 0.022 |
| Model 2* | -0.209 (-1.027; 0.610) | 0.617 |
| Model 2 + $MSG_{CGM}$ | -0.596 (-1.858; 0.666) | 0.354 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

**Table S6.15  Multivariable-adjusted associations of daily glucose variability (expressed as $SD_{CGM}$) with arterial outcome variables after exclusion of individuals with CGM data gaps.**

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=591) | | |
| Crude | 0.984 (0.671; 1.297) | <0.001 |
| Model 1 | 0.734 (0.446; 1.022) | <0.001 |
| Model 2* | 0.501 (0.207; 0.796) | 0.001 |
| Model 2 + $MSG_{CGM}$ | 0.368 (-0.095; 0.831) | 0.119 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=672) | | |
| Crude | -0.628 (-1.470; 0.214) | 0.144 |
| Model 1 | -0.043 (-0.821; 0.736) | 0.914 |
| Model 2* | 0.909 (0.127; 1.691) | 0.023 |
| Model 2 + $MSG_{CGM}$ | 0.139 (-1.102; 1.380) | 0.826 |
| Carotid intima-media thickness (cIMT), µm (n=673) | | |
| Crude | 20.012 (-1.404; 41.428) | 0.067 |
| Model 1 | 8.587 (-11.835; 29.009) | 0.409 |
| Model 2† | 12.413 (-9.973; 34.798) | 0.277 |
| Model 2 + $MSG_{CGM}$ | -2.126 (-37.762; 33.510) | 0.907 |
| Ankle-brachial index (ABI) (n=749) | | |
| Crude | -0.023 (-0.037; -0.008) | 0.002 |
| Model 1 | -0.026 (-0.040; -0.011) | <0.001 |
| Model 2† | -0.015 (-0.031; 0.001) | 0.071 |
| Model 2 + $MSG_{CGM}$ | -0.026 (-0.051; 0.000) | 0.047 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=672) | | |
| Crude | 1.751 (0.311; 3.192) | 0.017 |
| Model 1 | 1.254 (-0.165; 2.672) | 0.131 |
| Model 2‡ | 0.180 (-1.309; 1.670) | 0.812 |
| Model 2 + $MSG_{CGM}$ | -1.027 (-3.396; 1.341) | 0.395 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=672) | | |
| Crude | 1.693 (0.773; 2.614) | <0.001 |
| Model 1 | 1.230 (0.315; 2.145) | 0.008 |
| Model 2* | -0.011 (-0.886; 0.864) | 0.980 |
| Model 2 + $MSG_{CGM}$ | -0.277 (-1.668; 1.114) | 0.696 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

6

**Table S6.16** Multivariable-adjusted associations of daily glucose variability (expressed as $SD_{CGM}$) with arterial outcome variables after exclusion of individuals with a suboptimal CGM recording period.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=620) | | |
| Crude | 0.931 (0.642; 1.221) | <0.001 |
| Model 1 | 0.649 (0.381; 0.917) | <0.001 |
| Model 2* | 0.425 (0.151; 0.698) | 0.002 |
| Model 2 + $MSG_{CGM}$ | 0.299 (-0.109; 0.706) | 0.151 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=698) | | |
| Crude | -0.711 (-1.519; 0.098) | 0.085 |
| Model 1 | -0.017 (-0.765; 0.730) | 0.964 |
| Model 2* | 0.782 (0.028; 1.537) | 0.044 |
| Model 2 + $MSG_{CGM}$ | 0.066 (-1.116; 1.248) | 0.913 |
| Carotid intima-media thickness (cIMT), μm (n=699) | | |
| Crude | 26.580 (6.011; 47.149) | 0.011 |
| Model 1 | 12.627 (-6.950; 32.204) | 0.206 |
| Model 2† | 16.739 (-4.712; 38.190) | 0.126 |
| Model 2 + $MSG_{CGM}$ | 4.126 (-29.577; 37.829) | 0.810 |
| Ankle-brachial index (ABI) (n=785) | | |
| Crude | -0.020 (-0.034; -0.007) | 0.003 |
| Model 1 | -0.024 (-0.037; -0.010) | 0.001 |
| Model 2† | -0.011 (-0.026; 0.004) | 0.163 |
| Model 2 + $MSG_{CGM}$ | -0.019 (-0.041; 0.004) | 0.109 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=698) | | |
| Crude | 1.514 (0.155; 2.872) | 0.029 |
| Model 1 | 0.944 (-0.396; 2.285) | 0.167 |
| Model 2‡ | 0.097 (-1.316; 1.510) | 0.893 |
| Model 2 + $MSG_{CGM}$ | -1.544 (-3.759; 0.671) | 0.172 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=698) | | |
| Crude | 1.579 (0.707; 2.452) | <0.001 |
| Model 1 | 0.996 (0.133; 1.859) | 0.024 |
| Model 2* | -0.175 (-1.006; 0.656) | 0.679 |
| Model 2 + $MSG_{CGM}$ | -0.619 (-1.923; 0.685) | 0.352 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

**Table S6.17** Multivariable-adjusted associations of daily glucose variability (expressed as $SD_{CGM}$) with arterial outcome variables after exclusion of individuals who underwent CGM as part of a 'catch-up visit'.

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=418) | | |
| Crude | 0.724 (0.337; 1.107) | <0.001 |
| Model 1 | 0.408 (0.062; 0.754) | 0.021 |
| Model 2* | 0.188 (-0.160; 0.536) | 0.289 |
| Model 2 + $MSG_{CGM}$ | -0.200 (-0.673; 0.273) | 0.406 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=500) | | |
| Crude | -0.728 (-1.882; 0.425) | 0.215 |
| Model 1 | 0.147 (-0.904; 1.198) | 0.784 |
| Model 2* | 0.521 (-0.513; 1.555) | 0.323 |
| Model 2 + $MSG_{CGM}$ | -0.355 (-1.758; 1.047) | 0.619 |
| Carotid intima-media thickness (cIMT), μm (n=500) | | |
| Crude | 42.209 (12.158; 72.261) | 0.006 |
| Model 1 | 22.183 (-5.946; 50.312) | 0.122 |
| Model 2† | 27.960 (-2.331; 58.251) | 0.070 |
| Model 2 + $MSG_{CGM}$ | 4.808 (-36.285; 45.902) | 0.818 |
| Ankle-brachial index (ABI) (n=577) | | |
| Crude | 0.000 (-0.019; 0.019) | 0.998 |
| Model 1 | -0.004 (-0.022; 0.015) | 0.684 |
| Model 2† | 0.011 (-0.009; 0.031) | 0.266 |
| Model 2 + $MSG_{CGM}$ | -0.002 (-0.028; 0.025) | 0.902 |
| Mean circumferential wall stress ($CWS_{mean}$), kPa (n=500) | | |
| Crude | -0.767 (-2.710; 1.176) | 0.438 |
| Model 1 | -0.987 (-2.909; 0.936) | 0.314 |
| Model 2‡ | -1.740 (-3.697; 0.217) | 0.081 |
| Model 2 + $MSG_{CGM}$ | -1.867 (-4.531; 0.797) | 0.169 |
| Pulsatile circumferential wall stress ($CWS_{puls}$), kPa (n=500) | | |
| Crude | 0.434 (-0.824; 1.692) | 0.498 |
| Model 1 | -0.139 (-1.382; 1.105) | 0.827 |
| Model 2* | -0.872 (-2.060; 0.315) | 0.150 |
| Model 2 + $MSG_{CGM}$ | -0.473 (-2.088; 1.142) | 0.565 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Crude: $SD_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + $MSG_{CGM}$: additionally adjusted for mean sensor glucose.

6

**Table S6.18 Multivariable-adjusted associations of mean sensor glucose with arterial outcome variables.**

| Model | B (95%CI) | p-value |
|---|---|---|
| Carotid femoral pulse wave velocity (cf-PWV), m/s (n=643) | | |
|   Crude | 0.437 (0.315; 0.560) | <0.001 |
|   Model 1 | 0.298 (0.184; 0.412) | <0.001 |
|   Model 2* | 0.177 (0.057; 0.297) | 0.004 |
|   Model 2 + SD$_{CGM}$ | 0.087 (-0.091; 0.265) | 0.338 |
| Carotid distensibility coefficient (carDC), $10^{-3}$/kPa (n=725) | | |
|   Crude | -0.423 (-0.759; -0.087) | 0.014 |
|   Model 1 | -0.143 (-0.456; 0.171) | 0.372 |
|   Model 2* | 0.412 (0.089; 0.734) | 0.012 |
|   Model 2 + SD$_{CGM}$ | 0.435 (-0.062; 0.933) | 0.086 |
| Carotid intima-media thickness (cIMT), µm (n=726) | | |
|   Crude | 14.612 (6.085; 23.139) | 0.001 |
|   Model 1 | 7.526 (-0.670; 15.722) | 0.072 |
|   Model 2† | 8.847 (-0.330; 18.024) | 0.059 |
|   Model 2 + SD$_{CGM}$ | 9.398 (-4.793; 23.589) | 0.194 |
| Ankle-brachial index (ABI) (n=816) | | |
|   Crude | -0.005 (-0.011; 0.000) | 0.068 |
|   Model 1 | -0.008 (-0.014; -0.002) | 0.006 |
|   Model 2† | -0.002 (-0.009; 0.004) | 0.507 |
|   Model 2 + SD$_{CGM}$ | 0.004 (-0.0006 0.013) | 0.477 |
| Mean circumferential wall stress (CWS$_{mean}$), kPa (n=725) | | |
|   Crude | 1.096 (0.528; 1.664) | <0.001 |
|   Model 1 | 0.823 (0.256; 1.390) | 0.005 |
|   Model 2‡ | 0.317 (-0.293; 0.927) | 0.308 |
|   Model 2 + SD$_{CGM}$ | 0.694 (-0.248; 1.635) | 0.149 |
| Pulsatile circumferential wall stress (CWS$_{puls}$), kPa (n=725) | | |
|   Crude | 0.905 (0.542; 1.268) | <0.001 |
|   Model 1 | 0.657 (0.293; 1.020) | <0.001 |
|   Model 2* | 0.029 (-0.330; 0.388) | 0.873 |
|   Model 2 + SD$_{CGM}$ | 0.230 (-0.323; 0.784) | 0.414 |

Regression coefficients (B) indicate the mean difference (95% confidence interval) associated with 1 unit (mmol/L) increase of MSG$_{CGM}$. Crude: SD$_{CGM}$. Model 1: adjusted for age, sex, and education level. Model 2: additionally adjusted for mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, and use of antihypertensive and lipid-modifying drugs. Model 2 + SD$_{CGM}$: additionally adjusted for standard deviation.

**Table S6.19 Regression coefficients of standard deviation and mean sensor glucose in the fully adjusted models with arterial outcome variables estimated with different degrees of penalization.**

| VIF | SD (B, 95%CI) | $SD_{CGM}$ (st.β, 95%CI) | p-value | $MSG_{CGM}$ (st.β, 95%CI) | p-value |
|---|---|---|---|---|---|
| Carotid femoral pulse wave velocity* (cf-PWV) (n=643) | | | | | |
| Model 2 + $MSG_{CGM}$ (λ=0) | 0.259 (-0.127; 0.771) | 0.070 (-0.037; 0.190) | 0.198 | 0.055 (-0.086; 0.174) | 0.426 |
| Halfway (λ=0.04) | 0.246 (-0.113; 0.723) | 0.065 (-0.031; 0.183) | 0.192 | 0.054 (-0.065; 0.174) | 0.378 |
| Model 2 (λ=0.11) | 0.250 (-0.067; 0.663) | 0.065 (-0.018; 0.167) | 0.160 | 0.059 (-0.043; 0.164) | 0.272 |
| Carotid distensibility coefficient* (carDC) (n=725) | | | | | |
| Model 2 + $MSG_{CGM}$ (λ=0) | -0.087 (-1.154; 1.088) | -0.008 (-0.106; 0.102) | 0.864 | 0.095 (-0.024; 0.205) | 0.110 |
| Halfway (λ=0.04) | -0.069 (-1.179; 1.055) | -0.006 (-0.106; 0.095) | 0.904 | 0.089 (-0.017; 0.201) | 0.086 |
| Model 2 (λ=0.12) | -0.035 (-1.065; 1.023) | -0.003 (-0.097; 0.092) | 0.952 | 0.088 (-0.014; 0.184) | 0.102 |
| Carotid intima-media thickness† (cIMT) (n=726) | | | | | |
| Model 2 + $MSG_{CGM}$ (λ=0) | -0.871 (-35.641; 31.033) | -0.003 (-0.128; 0.109) | 0.954 | 0.080 (-0.038; 0.201) | 0.196 |
| Halfway (λ=0.04) | -0.215 (-34.456; 32.768) | -0.001 (-0.128; 0.120) | 0.986 | 0.075 (-0.045; 0.204) | 0.206 |
| Model 2 (λ=0.12) | -1.872 (-32.526; 30.975) | -0.007 (-0.123; 0.111) | 0.916 | 0.078 (-0.038; 0.207) | 0.198 |
| Ankle-brachial index† (ABI) (n=816) | | | | | |
| Model 2 + $MSG_{CGM}$ (λ=0) | -0.017 (-0.047; 0.007) | -0.086 (-0.232; 0.038) | 0.184 | 0.042 (-0.057; 0.147) | 0.438 |
| Halfway (λ=0.04) | -0.009 (-0.021; 0.001) | -0.045 (-0.105; 0.007) | 0.104 | 0.003 (-0.040; 0.042) | 0.898 |
| Model 2 (λ=0.11) | -0.006 (-0.014; 0.000) | -0.033 (-0.071; 0.002) | 0.060 | -0.008 (-0.032; 0.017) | 0.548 |
| Mean circumferential wall stress‡ ($CWS_{mean}$) (n=725) | | | | | |
| Model 2 + $MSG_{CGM}$ (λ=0) | -1.138 (-3.408; 1.304) | -0.061 (-0.181; 0.073) | 0.344 | 0.085 (-0.045; 0.220) | 0.186 |
| Halfway (λ=0.04) | -1.131 (-3.278; 1.165) | -0.060 (-0.172; 0.063) | 0.312 | 0.090 (-0.039; 0.202) | 0.160 |
| Model 2 (λ=0.12) | -1.077 (-3.067; 1.218) | -0.059 (-0.169; 0.066) | 0.318 | 0.082 (-0.044; 0.204) | 0.180 |
| Pulsatile circumferential wall stress* ($CWS_{puls}$) (n=725) | | | | | |
| Model 2 + $MSG_{CGM}$ (λ=0) | -0.673 (-1.785; 0.707) | -0.057 (-0.150; 0.058) | 0.354 | 0.049 (-0.065; 0.144) | 0.420 |
| Halfway (λ=0.04) | -0.594 (-1.790; 0.788) | -0.049 (-0.148; 0.066) | 0.350 | 0.045 (-0.065; 0.146) | 0.434 |
| Model 2 (λ=0.12) | -0.532 (-1.702; 0.643) | -0.045 (-0.145; 0.053) | 0.374 | 0.042 (-0.055; 0.138) | 0.410 |

Regression coefficients (B) indicate the median difference (95% confidence interval) associated with 1 unit (mmol/L) increase of $SD_{CGM}$. Standardized regression coefficients (st.β) indicate the median difference (95% confidence interval) associated with 1 SD higher $SD_{CGM}$ or $MSG_{CGM}$. All coefficients were estimated by use of ridge regression. We pragmatically chose the level of penalization based on the λ required to reduce the variance inflation factor (VIF) of Model 2 + $MSG_{CGM}$ back to the VIF of Model 2 (or halfway back). Point estimates and 95% confidence intervals were calculated by use of 1,000 bootstrap estimates. The associations were adjusted for age, sex, educational level, mean arterial pressure (*), office systolic blood pressure (†) or brachial pulse pressure (‡), heart rate (in case of cf-PWV and ABI only), body mass index, smoking status, alcohol use, total-to-HDL cholesterol levels, use of antihypertensive and lipid-modifying drugs, and the other CGM-assessed index.

6

# Chapter 7

(Pre)diabetes, greater glycemia, and greater daily glucose variability are associated with lower retinal nerve fiber layer thickness, an index of neurodegeneration – The Maastricht Study

Frank C.T. van der Heide*, Yuri D. Foreman*, Iris W.M. Franken,
Ronald M.A. Henry, Alfons J.H.M. Houben, Abraham A. Kroon, Pieter C. Dagnelie,
Simone J.P.M. Eussen, Tos T.J.M. Berendschot, Jan S.A.G. Schouten,
Carroll A.B. Webers, Miranda T. Schram, Carla J.H. van der Kallen,
Marleen M.J. van Greevenbroek, Anke Wesselius, Casper G. Schalkwijk,
Nicolaas C. Schaper, Martijn C.G.J. Brouwers, and Coen D.A. Stehouwer

* First authors contributed equally

# Abstract

### Objective

Retinopathy and neuropathy in type 2 diabetes are preceded by retinal nerve fiber layer (RNFL) thinning, an index of neurodegeneration. We investigated whether glucose metabolism status (GMS), measures of glycemia, and daily glucose variability (GV) are associated with RNFL thickness over the entire range of glucose tolerance.

### Research design and methods

We used cross-sectional data from The Maastricht Study (up to 5,455 participants, 48.9% men, mean age 59.5 years and 22.7% with type 2 diabetes) to investigate the associations of GMS, measures of glycemia (fasting plasma glucose [FPG], 2-hour post-load glucose [2-h PG], HbA$_{1c}$, advanced glycation endproducts [AGEs] assessed as skin autofluorescence [SAF]) and indices of daily GV (incremental glucose peak [IGP] and continuous glucose monitoring [CGM]-assessed standard deviation [SD]) with mean RNFL thickness. We used linear regression analyses and, for GMS, P for trend analyses. We adjusted associations for demographic, cardiovascular risk and lifestyle factors, and, only for measures of GV, for indices of mean glycemia.

### Results

After full adjustment, type 2 diabetes and prediabetes (versus normal glucose metabolism) were associated with lower RNFL thickness (standardized beta [95%CI], respectively -0.16[-0.25;-0.08]; -0.05[-0.13;0.03]; P$_{trend}$=0.001). Greater FPG, 2-h PG, HbA$_{1c}$, SAF, IGP and CGM-assessed SD were also associated with lower RNFL thickness (per SD, respectively -0.05 [-0.08; -0.01]; -0.06 [-0.09; -0.02]; -0.05 [-0.08; -0.02]; -0.04 [-0.07; -0.01]; -0.06 [-0.12; -0.01]; and -0.07 [-0.21; 0.07]).

### Conclusion

In this population-based study, a more adverse GMS and, over the entire range of glucose tolerance, greater glycemia and daily GV were associated with lower RNFL thickness. Hence, early identification of individuals with hyperglycemia, early glucose-lowering treatment, and early monitoring of daily GV may contribute to the prevention of RNFL thinning, an index of neurodegeneration and precursor of retinopathy and neuropathy.

## Introduction

Retinopathy and neuropathy, both hallmark microvascular complications of type 2 diabetes, are preceded by subtle neurodegenerative changes. Such changes include retinal nerve fiber layer (RNFL) thinning, which can be non-invasively assessed.[1-3] RNFL thinning reflects a gradual loss of retinal ganglion cell axons, which transmit visual information from the retina to the brain.[1,2] Mechanistically, elevated glucose concentrations are thought to be toxic for retinal ganglion cells as well as for retinal endothelial and glia cells, which contribute to local metabolic regulation.[4,5] Thus, hyperglycemia can lead to retinal ganglion cell apoptosis and loss of retinal ganglion cell axons both directly and through impairment of endothelial and glia cell function.[6]

As postulated in the "ticking clock hypothesis", hyperglycemia-mediated damage is thought to be a continuous (i.e., linear) process that already starts before the onset of type 2 diabetes.[7,8] Indeed, our group observed linear associations of more adverse glucose metabolism status (GMS) and higher glycemia (estimated by various measures) with lower heart rate variability,[9] more structural brain abnormalities,[10,11] and worse peripheral nerve function,[12] all of which are measures of neurodegeneration.

The current literature on the associations of hyperglycemia with RNFL thickness has some important limitations. First, in previous population-based studies, several important confounders were not included, such as age,[13-15] sex,[14,15] socioeconomic status,[13-23] cardiovascular risk factors,[13-15,19,21-24] and lifestyle factors (e.g., alcohol use,[13,16-19,21] diet,[13-24] and physical activity).[13-19,21-24] Second, no population-based studies have yet investigated whether advanced glycation endproducts (AGEs) or daily glucose variability are associated with RNFL thickness. It is important to establish to what extent hyperglycemia (including AGEs) and daily glucose variability contribute to RNFL thinning over the entire range of glucose tolerance because early diagnosis and treatment of hyperglycemia as well as novel strategies to monitor glycemic exposure may contribute to the early prevention of hyperglycemia-mediated RNFL thinning.[5,25,26]

In view of above, we investigated, using a large, well-characterized population-based cohort study, whether more adverse GMS, greater glycemia, and greater daily glucose variability are associated with lower RNFL thickness.

## Methods

### *Study population and design*

We used data from The Maastricht Study, a prospectively designed, population-based observational cohort study. The rationale and methodology have been described previously.[27] In brief, the study focuses on the etiology, pathophysiology, complications and comorbidities of type 2 diabetes mellitus and is characterized by an extensive phenotyping approach. Eligible for participation were all individuals aged between 40 and 75 years and living in the southern part of the Netherlands. Participants were

7

recruited through mass media campaigns and from the municipal registries and the regional Diabetes Patient Registry via mailings. Recruitment was stratified according to known type 2 diabetes status, with an oversampling of individuals with type 2 diabetes, for reasons of efficiency.[27] The present report includes cross-sectional data of 8,005 participants who completed the baseline survey between November 2010 and September 2018. The examinations of each participant were performed within a time window of three months. From 19 September 2016 until 13 September 2018, participants were invited to also undergo CGM.[26] During this period, a selected group of recently included participants was invited to return for CGM ('catch-up visit'). For these participants only there was a median time interval ('visit interval') of 2.1 years between CGM and all other measurements (more details are provided in the Supplemental Material). The study has been approved by the institutional medical ethical committee (NL31329.068.10) and the Minister of Health, Welfare and Sports of the Netherlands (Permit 131088-105234-PG). All participants gave written informed consent.[27]

## Glucose metabolism status

After an overnight fast, participants underwent a standardized seven-point oral glucose tolerance test (OGTT) as part of which venous samples were collected at 15, 30, 45, 60, 90, and 120 minutes post ingestion of a 75g glucose drink. All participants underwent an OGTT except for the participants who used insulin or had a fasting plasma glucose (FPG) concentration above 11.0 mmol/L. Based on FPG and 2-hour post-load glucose, GMS was determined as normal glucose metabolism (NGM), prediabetes (impaired fasting glucose, impaired glucose tolerance, or both), type 2 diabetes, or other types of diabetes (including type 1 diabetes) in accordance with the World Health Organization 2006 criteria.[28]

## Measures of glycemia

FPG (mmol/L) and hemoglobin $A_{1c}$ (Hb$A_{1c}$; mmol/mol; %) were determined in venous plasma samples collected after an overnight fast. Two-hour post-load glucose (mmol/L) was determined in venous plasma collected at 120 minutes post glucose drink ingestion. AGEs were assessed with the AGE Reader (DiagnOptics Technologies BV, Groningen, the Netherlands). In brief, the AGE Reader is a desktop device that uses the characteristic fluorescent properties of certain AGEs to quantify their accumulation in the skin as skin autofluorescence (SAF; arbitrary units [AU]).[29] The AGE Reader illuminates a skin surface of 4 cm$^2$, shielded from other light, and uses the ratio of the reflection of fluorescent light (wavelength 420 to 600 nm) to non-fluorescent light (300–420 nm) to calculate SAF (more details are provided in the Supplemental Materials).

## Indices of daily glucose variability

The incremental glucose peak (IGP; mmol/L), a recently validated OGTT-based index of daily glucose variability, was calculated by subtracting FPG from the maximum glucose peak value measured during a complete seven-point OGTT.[30] At the time of analysis, data on IGP were available in a subset of participants (n=2,407). We used CGM (iPro2 and Enlite Glucose Sensor; Medtronic, Tolochenaz, Switzerland) to assess daily glucose variability during a 1-week period, which was expressed as standard deviation (mmol/L) of mean sensor glucose (mmol/L).[26] CGM-assessed data were available in a subset of participants (n=622). More details on the assessment of glycemic indices with OGTT and CGM are provided in the Supplemental Methods.

## RNFL thickness

We assessed peripapillary RNFL thickness (µm) in both eyes using optical coherence tomography (OCT) (Spectralis unit and Eye Explorer version 5.7.5.0 software; Heidelberg Engineering, Heidelberg, Germany; 3.45-mm-diameter–circle scan, manually centered on optic nerve head, 12°, 768 voxels, 100 automatic real-time tracking). Intra- and interindividual reliability, expressed as intraclass correlation coefficients, are 0.97 and 0.96 respectively.[31] At least 15 minutes before the examination, pupils were dilated with topical 0.5% tropicamide and 2.5% phenylephrine. Experienced graders masked to clinical information on the participants reviewed the OCT scans and graded their quality. OCT images were excluded if one of the following criteria was present: scan error (i.e., incomplete scan, poor centering of the circular scan on the optic nerve head, RNFL layer incorrectly defined, or technical problem with the OCT device) or poor imaging quality (signal-to-noise ratio<15 dB).[23] If data from both eyes were available (n=2,796 participants) we averaged RNFL thickness in order to reduce measurement error. If data from only one eye were available (n=2,755 participants), we used the RNFL thickness of that eye in the analyses. More details, including on quality criteria, are shown in the Supplemental Methods.

## Covariates

As described previously,[27] we assessed educational level (low, intermediate, high), socio-economic status (income level and occupational status),[32] smoking status (never, former, current), alcohol use (none, low, high), history of cardiovascular disease, and duration of diabetes by questionnaire; assessed dietary habits ("dietary intake") with the Dutch Healthy Diet index sum score, a measure of adherence to the Dutch dietary guidelines 2015,[33] based on a validated food frequency questionnaire;[34] assessed lipid-modifying, antihypertensive, intraocular pressure-lowering, and glucose-lowering medication use as part of a medication interview; assessed weight, height, and waist circumference during a physical examination; calculated body mass index (BMI) based on body weight and height; measured office and 24-hour ambulatory blood pressure

7

(BP); measured total daily physical activity (hours/day) with an accelerometer;[35] measured lipid profile and plasma biomarkers of low-grade inflammation[36] (i.e., high-sensitive C-reactive protein, serum amyloid A, interleukin-6, interleukin-8 and tumor necrosis factor alpha) in fasting venous blood samples; measured urinary albumin excretion in two 24-hour urine collections; calculated the estimated glomerular filtration rate (eGFR) based on serum creatinine only, since cystatin C was not presently available in all study participants; assessed presence of retinopathy in both eyes via fundus photography; and used an automated refractor and noncontact tonometer (Tonoref II; Nidek, Gamagordi, Japan) to assess spherical equivalent and intraocular pressure in both eyes. Glaucoma was defined as use of intraocular pressure-lowering medication, intraocular pressure higher than 21 mm Hg in any eye (91.3% of all participants had data on intraocular pressure available for at least 1 eye), or both. Spherical equivalent was defined as the mean spherical equivalent of both eyes (available for 91.1% of all participants) or as the spherical equivalent of the eye for which data were available.

## *Statistical analysis*

We used multivariable linear regression analyses to investigate the associations of GMS (entered as dummy variables of prediabetes, type 2 diabetes, or other types of diabetes versus NGM) and standardized FPG, 2-hour post-load glucose, HbA$_{1c}$, SAF, IGP, and CGM-assessed standard deviation (determinants) with standardized mean RNFL thickness (outcome).

We performed P for trend analyses to test for linear trend with more adverse GMS. To test for trend, we entered GMS into the model as an ordinal variable (i.e., GMS was coded as NGM=0, prediabetes=1, type 2 diabetes=2).In P for trend analyses we excluded participants with other types of diabetes because other types of diabetes (such as type 1 diabetes) do not constitute part of the spectrum of deterioration of GMS from NGM to prediabetes and type 2 diabetes. Then, we checked whether we could assume a linear trend by comparing the statistical variance explained by the statistical model in which GMS was entered as dummy variables to the statistical model in which GMS was entered as an ordinal variable. We used a likelihood ratio test to assess whether the amount of variance explained by both models differed statistically significantly. A p-value >0.05 indicates that both models are not different and, thus, that a linear trend can be assumed. For all analyses under study, the p-value for the likelihood ratio test was >0.05 (data not shown) and therefore a linear trend could be assumed.

Model 1 shows crude results. In model 2, we adjusted for age, sex, educational status (low, medium, high). We chose these variables because they are key potential confounders.[17] In model 3, we additionally adjusted for variables of which their status as confounder has been less firmly established (office systolic BP, use of antihypertensive medication [yes/no], waist circumference, total cholesterol / HDL cholesterol ratio, lipid-modifying medication [yes/no], smoking status [current, former,

never], and alcohol consumption status [none, low, high]).[17] Then, and only for IGP and CGM-assessed standard deviation, we additionally adjusted for HbA$_{1c}$ or mean sensor glucose, respectively (model 3 + HbA$_{1c}$; model 3 + mean sensor glucose), so that we could differentiate between daily glucose variability and average glycemia, both of which are strongly related.[37] Additionally, and only for CGM-assessed indices, we entered 'visit interval' in model 1. The associations were expressed as standardized regression coefficient (stβ) and corresponding 95% confidence interval (95%CI). Collinearity diagnostics (i.e., tolerance <0.10 and/or variance inflation factor >10) were used to detect excessive multicollinearity between covariates.

We tested for interaction to assess whether associations differed by GMS (i.e., between individuals with type 2 diabetes, individuals with prediabetes, and individuals with NGM) or by sex. For interaction analyses with GMS, we excluded participants with other types of diabetes from the interaction analyses because the number of these participants was small.

To assess the robustness of our findings we performed several sensitivity analyses. First, we repeated the analyses with additional adjustment for lifestyle factors (dietary intake, physical activity) or ocular variables (spherical equivalent, intraocular pressure).[38-40] Adjustment for these potential confounders was not included in the main analyses because data were missing for a relatively large number of participants (up to n=768 had missing data on one or more of these variables). Second, we additionally adjusted for kidney variables (eGFR and urinary albumin excretion), history of cardiovascular disease, plasma biomarkers of low-grade inflammation, retinopathy, and glaucoma. We adjusted for these covariates in a separate model because they may be confounders but may also (in part) be mediators or descendants of the outcome.[41] Third, we performed analyses in which participants with retinopathy, glaucoma, or other types of diabetes were excluded. Fourth, we replaced IGP with other OGTT-based indices of daily glucose variability (i.e., maximum glucose peak and 1-hour post-load glucose). We did not include maximum glucose peak and 1-hour post-load glucose in the main analyses because they are known to correlate less strongly than IGP with CGM-assessed indices of daily glucose variability in GMS-stratified groups.[30] Fifth, we replaced CGM-assessed standard deviation with other CGM-assessed measures of daily glucose variability (more details are provided in the Supplemental Methods section). Sixth, we studied the association between mean sensor glucose and retinal nerve fiber layer thickness. Although the sample size of the CGM study population was relatively small (n=622), CGM-based glycemic mean sensor glucose may be less susceptible to measurement error than other measures of glycemia under study, as mean sensor glucose constitutes the average of a large number of glucose concentrations.[30,42] Seventh, we replaced waist circumference with BMI; educational status with occupational status or income level; and office systolic BP with office diastolic BP, systolic or diastolic 24-hour ambulatory BP. Eighth, and only for SAF, we additionally adjusted the association between SAF and RNFL thickness for FPG, 2-hour

7

post-load glucose, or HbA$_{1c}$. We additionally adjusted for these measures to investigate whether the association between SAF and RNFL thickness was independent of short to middle long-term exposure to hyperglycemia. Ninth, and only for IGP and CGM-assessed standard deviation, we respectively replaced HbA$_{1c}$ with FPG or SAF, and mean sensor glucose with FPG, SAF, or HbA$_{1c}$. Tenth, and only for CGM-assessed standard deviation, we repeated the analyses after exclusion of individuals with insufficient CGM recording days, with CGM recording data gaps, or with a visit interval. Eleventh, because of the strong correlation between CGM-assessed standard deviation and mean sensor glucose,[37] we repeated the main analysis using ridge regression, which is a valid method to counter potential instability caused by multicollinearity (additional information is provided in the Supplemental Methods).[43] Last, we studied the association between duration of diabetes and RNFL thickness. Data on duration of diabetes were only available for individuals with type 2 diabetes (n=982).

All analyses were performed with Statistical Package for Social Sciences version 25.0 (IBM SPSS, IBM Corp, Armonk, NY, USA). For all analyses, a p-value <0.05 was considered statistically significant.

## Results

### *Selection and characteristics of the study population*

**Figure 7.1** gives an overview of the study population selection. Participants in whom OCT data were missing or of insufficient quality were excluded first (n=2,454). Next, individuals with missing data on confounders were excluded (n=96). The sample size of the final study populations depended on the availability of data on the main determinant (n=5,455 for GMS, n=982 to 5,454 for measures of glycemia, and n=622 to 2,407 for indices of daily glucose variability).

**Table 7.1** and **Supplemental Table S7.1** show general participant characteristics according to tertiles of RNFL thickness. Overall, participants with a thinner RNFL were more often men, were older, and had a more adverse cardiovascular risk profile. General characteristics of participants included in the study were comparable to those of participants with missing data (**Supplemental Table S7.2**).

### *Glucose metabolism status and RNFL thickness*

After full adjustment (model 3), a more adverse GMS was associated with lower RNFL thickness (standardized beta [95%CI], type 2 diabetes versus NGM -0.16 [-0.25; -0.08]; prediabetes versus NGM -0.05 [-0.13; 0.03], P for trend =0.001; **Table 7.2** and **Figure 7.2**).
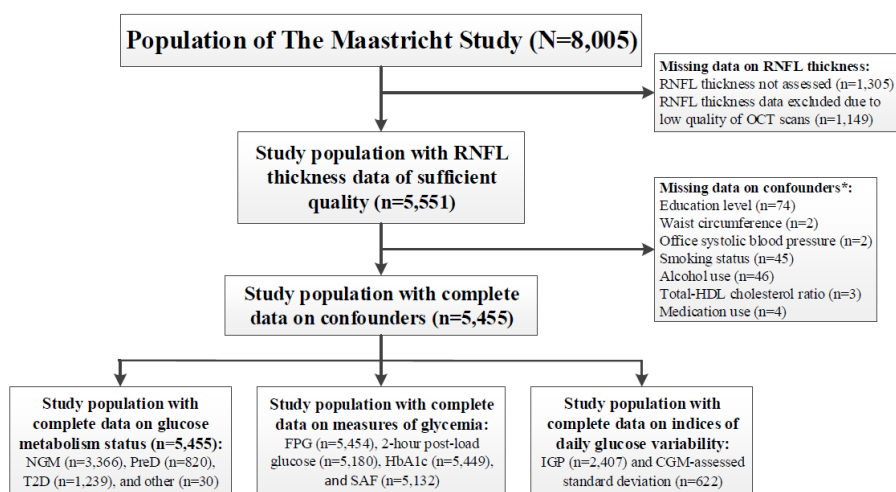
**Figure 7.1    RNFL study population selection.** OCT, optical coherence tomography; RNFL, retinal nerve fiber layer; HDL, high density lipoprotein; NGM, normal glucose metabolism; PreD, prediabetes; T2D, type 2 diabetes; FPG, fasting plasma glucose; HbA₁c, hemoglobin A1C; SAF, skin autofluorescence; IGP, incremental glucose peak; CGM, continuous glucose monitoring. * Not mutually exclusive.

**Table 7.1    General study population characteristics according to tertiles of the retinal nerve fiber layer thickness in the study population with complete data on glucose metabolism status.**

| Characteristic | RNFL thickness | | | |
| | Total study population (n=5,455) | Tertile 1 (high) (n=1,818) | Tertile 2 (middle) (n=1,819) | Tertile 3 (low) (n=1,818) |
|---|---|---|---|---|
| Age (years) | 59.5 ± 8.6 | 59.1 ± 8.7 | 59.3 ± 8.7 | 60.0 ± 8.5 |
| Men | 2,665 (48.9) | 805 (48.9) | 847 (46.6) | 1,013 (55.7) |
| Educational level | | | | |
| Low | 1,914 (35.1) | 669 (36.8) | 651 (35.8) | 594 (32.7) |
| Medium | 1,519 (27.8) | 525 (28.9) | 505 (27.8) | 489 (26.9) |
| High | 2,022 (37.1) | 624 (34.3) | 663 (36.4) | 735 (40.4) |
| Glucose metabolism status | | | | |
| NGM | 3,366 (61.7) | 1,174 (64.6) | 1,144 (62.9) | 1,048 (57.6) |
| Prediabetes | 820 (15.0) | 266 (14.6) | 280 (15.4) | 274 (15.1) |
| Type 2 diabetes | 1,239 (22.7) | 370 (20.4) | 383 (21.1) | 486 (26.7) |
| Other type of diabetes | 30 (0.5) | 8 (0.4) | 12 (0.7) | 10 (0.6) |
| Measures of glycemia | | | | |
| Fasting plasma glucose (mmol/L)* | 5.9 ± 1.5 | 5.8 ± 1.4 | 5.8 ± 1.5 | 6.0 ± 1.7 |
| 2-hour post-load glucose (mmol/L)* | 6.2 [4.9-8.6] | 6.1 [4.9-8.2] | 6.1 [4.9-8.5] | 6.3 [5.1-9.2] |
| HbA₁c (mmol/mol)* | 39.2 ± 9.1 | 38.7 ± 8.5 | 39.1 ± 9.3 | 39.9 ± 9.5 |
| HbA₁c (%)* | 5.7 ± 0.8 | 5.7 ± 0.8 | 5.7 ± 0.8 | 5.8 ± 0.9 |
| Skin autofluorescence (AU)* | 2.2 ± 0.5 | 2.1 ± 0.5 | 2.2 ± 0.5 | 2.2 ± 0.5 |
| Indices of daily glucose variability | | | | |
| Incremental glucose peak (mmol/L)* | 4.1 [2.7-6.5] | 3.9 [2.6-5.9] | 4.1 [2.8-6.5] | 4.4 [2.8-7.0] |

**Table 7.1**     (continued)

| Characteristic | RNFL thickness | | | |
| --- | --- | --- | --- | --- |
| | Total study population (n=5,455) | Tertile 1 (high) (n=1,818) | Tertile 2 (middle) (n=1,819) | Tertile 3 (low) (n=1,818) |
| CGM-assessed standard deviation (mmol/L)* | 0.86 [0.68-1.21] | 0.86 [0.67-1.18] | 0.85 [0.67-1.17] | 0.88 [0.70-1.29] |
| Use of glucose-modifying medication, yes vs. no | 927 (17.0) | 266 (14.6) | 289 (15.9) | 372 (20.5) |
| Waist circumference, men (cm) | 100.4 ± 11.6 | 100.2 ± 11.7 | 99.9 ± 11.7 | 101.0 ± 11.5 |
| Waist circumference, women (cm) | 89.1 ± 12.5 | 89.2 ± 12.8 | 88.9 ± 12.1 | 89.2 ± 12.8 |
| Total-to-HDL cholesterol ratio | 3.4 [2.8-4.2] | 3.3 [2.7-4.1] | 3.3 [2.8-4.2] | 3.4 [2.8-4.2] |
| Use of lipid-modifying medication, yes vs. no | 1,687 (30.9) | 535 (29.4) | 541 (29.7) | 611 (33.6) |
| Office systolic blood pressure (mmHg) | 133.2 ± 17.7 | 132.0 ± 18.0 | 133.0 ± 17.6 | 134.5 ± 17.5 |
| Office diastolic blood pressure (mmHg) | 75.5 ± 9.8 | 75.0 ± 9.8 | 75.3 ± 9.7 | 76.4 ± 9.9 |
| Use of antihypertensive medication, yes vs. no | 1,983 (36.4) | 601 (33.1) | 637 (35.0) | 745 (41.0) |
| Smoking status | | | | |
|   Never | 2,101 (38.5) | 717 (39.4) | 684 (37.6) | 700 (38.5) |
|   Former | 2,666 (48.9) | 843 (44.4) | 911 (50.1) | 912 (50.2) |
|   Current | 688 (12.6) | 258 (14.2) | 224 (12.3) | 206 (11.3) |
| Alcohol consumption | | | | |
|   None | 995 (18.2) | 344 (18.9) | 356 (19.6) | 295 (16.2) |
|   Moderate | 3,181 (58.3) | 1,070 (58.9) | 1,042 (57.3) | 1,069 (58.8) |
|   High | 1,279 (23.4) | 404 (22.2) | 421 (23.1) | 454 (25.0) |
| RNFL thickness (μm) | 94.8 ± 10.8 | 106.0 ± 6.3 | 95.2 ± 2.5 | 83.2 ± 6.8 |

Data are presented as mean ± standard deviation, median [interquartile range] or number (%). * Data shown in the study population with complete data on fasting plasma glucose (n=5,454), 2-hour post-load glucose (n=5,180), HbA$_{1c}$ (n=5,449), skin autofluorescence (n=5,132), incremental glucose peak (n=2,407), and CGM-assessed standard deviation (n=622). CGM, continuous glucose monitoring; HbA$_{1c}$, hemoglobin A1c; HDL, high-density lipoprotein; NGM, normal glucose metabolism; RNFL, retinal nerve fiber layer; AU, arbitrary units.

## *Measures of glycemia and RNFL thickness*

After full adjustment (model 3), greater FPG, 2-hour post-load glucose, HbA$_{1c}$, and SAF were statistically significantly associated with lower RNFL thickness (per SD, respectively -0.05 [-0.08; -0.01]; -0.06 [-0.09; -0.02]; -0.05 [-0.08; -0.02]; and -0.04 [-0.07; -0.01]; **Table 7.2** and **Figure 7.2**).
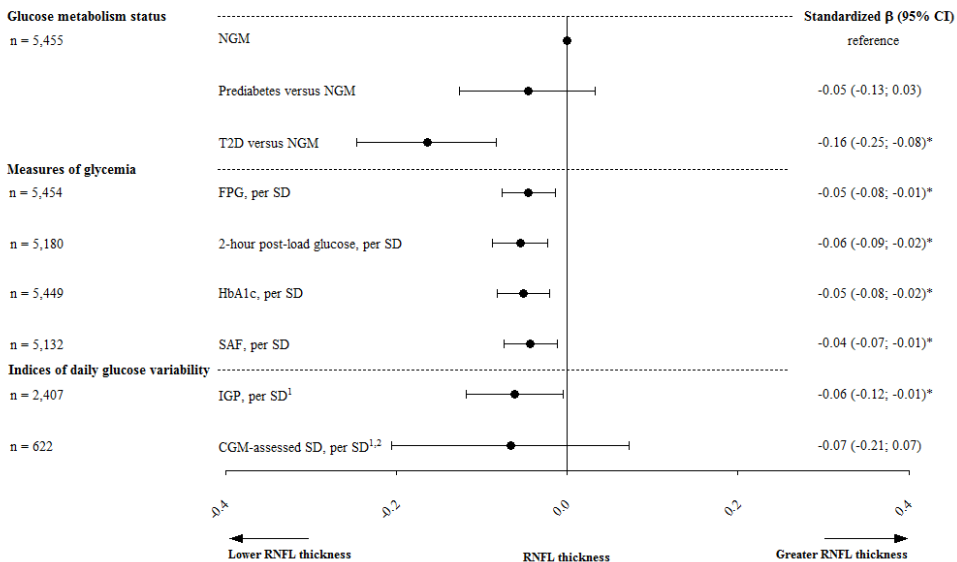
## *Indices of daily glucose variability and RNFL thickness*

After full adjustment (model 3), greater IGP was statistically significantly associated with lower RNFL thickness (per SD, -0.06 [-0.11; -0.01]; **Table 7.2** and **Figure 7.2**). The association remained statistically significant after additional adjustment for HbA$_{1c}$ (per SD, -0.06 [-0.12; -0.01]). After full adjustment (model 3), CGM-assessed standard deviation was also negatively associated with RNFL thickness, but not statistically significantly (per SD, -0.08 [-0.17; 0.01]). The association was similar after further adjustment for mean sensor glucose.

**Table 7.2    Associations of glucose metabolism status, measures of glycemia, and indices of daily glucose variability with RNFL thickness.**

| | Number of participants | RNFL, per SD | | | |
|---|---|---|---|---|---|
| | | Model 1 stβ (95% CI) | Model 2 stβ (95% CI) | Model 3 stβ (95% CI) | Model 3 + HbA1c/MSG stβ (95% CI) |
| **Glucose metabolism status** | | | | | |
| Prediabetes versus NGM | 5,455 | -0.07 (-0.15; 0.003) | -0.05 (-012; 0.03) | -0.05 (-0.13; 0.03) | NA |
| Type 2 diabetes versus NGM | 5,455 | **-0.19 (-0.25; -0.12)** | **-0.16 (-0.26; -0.09)** | **-0.16 (-0.25; -0.08)** | NA |
| **Measures of glycemia** | | | | | |
| Fasting plasma glucose, per SD | 5,454 | **-0.07 (-0.09; -0.04)** | **-0.05 (-0.08; -0.02)** | **-0.05 (-0.08; -0.01)** | NA |
| 2-hour post-load glucose, per SD | 5,180 | **-0.07 (-0.09; -0.04)** | **-0.06 (-0.09; -0.03)** | **-0.06 (-0.09; -0.02)** | NA |
| HbA1c, per SD | 5,449 | **-0.06 (-0.09; -0.03)** | **-0.05 (-0.08; -0.02)** | **-0.05 (-0.08; -0.02)** | NA |
| Skin autofluorescence, per SD | 5,132 | **-0.06 (-0.08; -0.03)** | **-0.04 (-0.04; -0.01)** | **-0.04 (-0.07; -0.01)** | NA |
| **Indices of daily glucose variability** | | | | | |
| Incremental glucose peak, per SD | 2,407 | **-0.09 (-0.13; -0.05)** | **-0.07 (-0.11; -0.03)** | **-0.06 (-0.11; -0.01)** | **-0.06 (-0.12; -0.01)** |
| CGM-assessed standard deviation, per SD | 622 | **-0.09 (-0.18; -0.01)** | **-0.09 (-0.17; -0.001)** | -0.08 (-0.17; 0.01) | -0.07 (-0.21; 0.07) |

Standardized regression coefficient (stβ) represents the difference in RNFL thickness in SD for type 2 diabetes or prediabetes versus NGM or per SD greater measure of glycemia or index of daily glucose variability. In the GMS, fasting plasma glucose, 2-hour post-load glucose, HbA1c, and skin autofluorescence study populations, 1 SD corresponds with 10.8 μm for RNFL thickness, 1.5 mmol/L for fasting plasma glucose, 4.0 mmol/L for 2-hour post-load glucose, 0.8% or 9.1 mmol/mol for HbA1c, and 0.5 AU for skin autofluorescence. For incremental glucose peak, 1 SD corresponds with 11.1 μm for RNFL thickness and 2.9 mmol/L for incremental glucose peak. For the CGM-assessed standard deviation, 1 SD corresponds with 10.8 μm for RNFL thickness and 0.58 mmol/L for CGM-assessed standard deviation. Bold denotes p<0.05. Variables entered in the models in addition to glucose metabolism status, measures of glycemia, or indices of daily glucose variability: model 1: none (crude results); model 2: age, sex, and educational status (low, medium, high); model 3: model 2 + office systolic blood pressure, total cholesterol to HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication (yes/no), waist circumference, smoking status (current, ever, never), and alcohol consumption status (none, low, high). In addition, only for incremental glucose peak, model 3 was additionally adjusted for HbA1c (model 3 + HbA1c), and only for CGM-assessed standard deviation, model 3 was additionally adjusted for MSG (model 3 + MSG). Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1. AU, arbitrary unit; CI, confidence interval; CGM, continuous glucose monitoring; GMS, glucose metabolism status; HbA1c, hemoglobin A1c; HDL, high-density lipoprotein; NA, not applicable; NGM, normal glucose metabolism; SD, standard deviation; RNFL, retinal nerve fiber layer; MSG, mean sensor glucose.

7

**Figure 7.2** **Associations between glucose metabolism status, measures of glycemia and indices of daily glucose variability with RNFL thickness (per SD).** Standardized regression coefficient (stβ) represents the difference in RNFL thickness in SD for type 2 diabetes or prediabetes status versus NGM status, or per SD greater measure of glycemia or index of daily glucose variability. In the GMS, fasting plasma glucose, 2-hour post-load glucose, HbA$_{1c}$, and SAF study populations, 1 SD corresponds with 10.8 µm for RNFL thickness, 1.5 mmol/L for fasting plasma glucose, 4.0 mmol/L for 2-hour post-load glucose, 0.8% or 9.1 mmol/mol for HbA$_{1c}$, and 0.5 AU for SAF. For incremental glucose peak, 1 SD corresponds with 11.1 µm for RNFL thickness and 2.9 mmol/L for incremental glucose peak. For CGM-assessed standard deviation, 1 SD corresponds with 10.8 µm for RNFL thickness and 0.58 mmol/L for CGM-assessed standard deviation. Variables entered in the models in addition to glucose metabolism status, measures of glycemia, or indices of daily glucose variability: age, sex, and educational status (low, medium, high), office systolic blood pressure, total cholesterol to HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication (yes/no), waist circumference, smoking status (current, ever, never), and alcohol consumption status (none, low, high). SD, standard deviation; CI, confidence interval; NGM, normal glucose metabolism; T2D, type 2 diabetes; FPG, fasting plasma glucose; HbA$_{1c}$, hemoglobin A1c; SAF, skin autofluorescence; IGP, incremental glucose peak; CGM, continuous glucose monitoring. * Indicates statistically significant (p<0.05). [1] The associations of indices of daily glucose variability with RNFL thickness were additionally adjusted for HbA$_{1c}$ (IGP) or mean sensor glucose (CGM-assessed SD). [2] Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1.

### Interaction analyses

GMS and sex did not modify any of the associations under study. All p-values for interaction are shown in **Supplemental Table S7.4**.

### Additional analyses

Quantitatively similar results were observed in a range of sensitivity analyses and are presented in the Supplemental Results section.

## Discussion

The present population-based study had two main findings. First, a more adverse GMS, greater glycemia (estimated from FPG, 2-hour post-load glucose, HbA$_{1c}$, and SAF) and greater daily glucose variability (estimated from IGP and CGM-assessed standard deviation) were all linearly and—except for CGM-assessed standard deviation—statistically significantly associated with lower RNFL thickness. Second, the associations between indices of daily glucose variability and RNFL thickness did not materially change after additional adjustment for measures of glycemia.

Our findings are in line with and extend observations from most previous studies.[13,19-24] The present study is the first large population-based study to comprehensively report associations of GMS, measures of glycemia, and indices of daily glucose variability with RNFL thickness, and also adjust for an extensive set of potential confounders. Additionally, the present study is the first to present associations of SAF, duration of diabetes, and indices of daily glucose variability with RNFL thickness.

Mechanistically, the linearity of the associations of GMS, measures of glycemia, and indices of daily glucose variability with RNFL thickness most likely reflects the increasing loss of retinal ganglion cells due to both hyperglycemia-induced neurotoxicity and impairment of functioning of retinal cells that contribute to metabolic regulation.[4,5] Such impairment of metabolic regulation can predispose retinal ganglion cells to ischemia.[5,44] Importantly, retinal ganglion cells are thought to be highly susceptible to ischemia, since they are highly active and have an energy demand that exceeds that of brain cells.[44]

These findings extend our previous work on the "ticking clock hypothesis",[5,9-12] which postulates that hyperglycemia-induced microvascular and neuronal deterioration is a continuous, gradual process that starts in prediabetes, progresses with the onset of type 2 diabetes, and continues during type 2 diabetes.[8] Indeed, we observed that the regression estimate for prediabetes was in between the estimate for type 2 diabetes and the reference category (i.e., NGM), and was directionally and numerically comparable to our previous findings.[9-12] However, the association between prediabetes and RNFL thickness was not statistically significant, which is most likely due to insufficient statistical power. We, therefore, additionally tested for a linear

7

trend with GMS deterioration by using the statistically more powerful P for trend analysis,[45] which was consistent with a linear decrease in RNFL thickness with more adverse GMS. In support, all measures of glycemia, regardless of whether they reflect shorter (i.e., FPG, 2-h PG, and HbA$_{1c}$) or longer (i.e., SAF and duration of diabetes) exposure, were consistently linearly associated with RNFL thickness.

Similarly, a likely explanation why the association between CGM-assessed standard deviation and RNFL thickness was not statistically significant is that the statistical power to detect any such association was too low.[46] Indeed, we observed that the association between IGP and RNFL thickness, which included almost fourfold the number of participants (n=2,407 versus n=622), was statistically significant. Moreover, the strength of the associations of IGP and CGM-assessed standard deviation with RNFL thickness were numerically analogous.

A probable explanation why the association between daily glucose variability and RNFL thickness was not materially altered after additional adjustment for measures of average glycemia is that daily glucose variability, measures of glycemia, and GMS represent different underlying constructs.[26] While daily glucose variability reflects oscillating glucose levels, other measures under study reflect exposure to average chronic levels of glycemia. Mechanistically, substantial glucose fluctuations entail hyperglycemic peaks, hypoglycemic nadirs (in individuals with type 2 diabetes treated with agents that can induce hypoglycemia), or both, which are thought to be potent inducers of retinal ganglion cell apoptosis.[26,44] Whereas hyperglycemic peaks may be highly neurotoxic, hypoglycemic nadirs likely hamper retinal ganglion cell metabolism as their key nutrient is glucose.[44]

Our findings can have several implications for clinical practice. First, the strength of the association between type 2 diabetes and RNFL thickness corresponds with 15 years of aging and, thus, indicates that with respect to neurodegeneration substantial "additional aging" occurs in individuals with type 2 diabetes (**Supplemental Table S7.14** shows how this comparison was calculated). Second, RNFL thickness may be a biomarker for the identification of individuals at risk of retinopathy and neuropathy. Use of RNFL thickness measurement is feasible because RNFL thickness assessment is non-invasive,[2] relatively inexpensive[2] and easier to perform than other tests of early neuronal dysfunction such as 24-hour electrocardiogram,[9] magnetic resonance imaging,[10,11] or electromyography.[12] Indeed, RNFL thickness has been found to be a promising early biomarker for other neurodegenerative diseases (e.g., multiple sclerosis).[47] Third, early glycemic control, possibly already in prediabetes, is likely crucial in the early prevention of microvascular and neuronal complications.[5] Last, our findings add to growing evidence that control of daily glucose variability besides mean glucose concentrations may be important to prevent microvascular complications.[48]

Strengths of this study are 1) the large size of this population-based cohort with oversampling of individuals with type 2 diabetes, which enabled accurate comparison of individuals with and without diabetes; 2) the extensive number of potential

confounders that were considered; 3) the use of state-of-the-art and novel methods (e.g., CGM)[26] to assess all variables included in this study; and 4) the considerable number of additional analyses, which generally yielded consistent findings.

The study has certain limitations. First, due to the cross-sectional nature of the study, causal inferences should be made with caution.[49] Mechanistically, hyperglycemia may not only lead to neurodegeneration but the reverse may also be true, thus causing a vicious cycle. Intact neurovascular interaction is required for normal microvascular function and impaired microvascular function may aggravate hyperglycemia.[5, 50] Second, we may have underestimated the strength of the associations of GMS, measures of glycemia, and daily glucose variability with RNFL thickness if such an association was similar or stronger in participants that were excluded from the study population (who generally tend to be less healthy).[51] The 2-hour post-load glucose and IGP results are most susceptible to this form of selection bias, as no data was available in individuals with the most therapy-intensive diabetes because they were excluded from undergoing an OGTT. Such range restriction may lead to underestimated associations.[51] Third, a single OGTT may misclassify GMS, especially in individuals with prediabetes. Because individuals classified with prediabetes based on their first OGTT are relatively more prone to receive a NGM classification based on their second OGTT,[52] this would likely lead to an underestimation of the association with RNFL thickness in the prediabetes group. Fourth, although we took an extensive set of confounders into account, we cannot fully exclude bias due to unmeasured confounding (e.g., environmental factors such as air pollution).[53] Fifth, due to the relatively low numbers of participants with data on CGM-based glycemic indices (n=622)), and—to a lesser extent—IGP (n=2,407), statistical power of analyses with these determinants was reduced compared to statistical power of analyses with GMS and measures of glycemia (n=5,132 to n=5,455).[46] Last, we studied Caucasian individuals aged 40-75 years with access to high-quality diabetes care. Therefore, the generalizability of our results to other populations requires further study.

## Conclusions

In summary, the present population-based study demonstrated that adverse GMS, greater  glycemia, and greater daily glucose variability are associated with lower RNFL thickness, i.e., neurodegeneration, independent of demographics, cardiovascular risk factors, and lifestyle risk factors. Hence, these results suggest that RNFL thickness may be an early biomarker for the identification of individuals at risk of retinopathy and neuropathy. Additionally, the combination of early glycemic monitoring and early glucose-lowering treatment, possibly already in prediabetes, may contribute to the prevention of RNFL thinning and, ultimately, retinopathy and neuropathy.

7

# References

1. Zakin E, Abrams R, Simpson DM. Diabetic Neuropathy. Semin Neurol. 2019;39(5):560-9.
2. Srinivasan S, Dehghani C, Pritchard N, Edwards K, Russell AW, Malik RA, et al. Optical coherence tomography predicts 4-year incident diabetic neuropathy. Ophthalmic Physiol Opt. 2017;37(4):451-9.
3. De Clerck EE, Schouten JS, Berendschot TT, Kessels AG, Nuijts RM, Beckers HJ, et al. New ophthalmologic imaging techniques for detection and monitoring of neurodegenerative changes in diabetes: a systematic review. Lancet Diabetes Endocrinol. 2015;3(8):653-63.
4. Tomlinson DR, Gardiner NJ. Glucose neurotoxicity. Nat Rev Neurosci. 2008;9(1):36-45.
5. Stehouwer CDA. Microvascular Dysfunction and Hyperglycemia: A Vicious Cycle With Widespread Consequences. Diabetes. 2018;67(9):1729-41.
6. Brownlee M. The pathobiology of diabetic complications: a unifying mechanism. Diabetes. 2005;54(6):1615-25.
7. Haffner SM, Stern MP, Hazuda HP, Mitchell BD, Patterson JK. Cardiovascular risk factors in confirmed prediabetic individuals. Does the clock for coronary heart disease start ticking before the onset of clinical diabetes? JAMA. 1990;263(21):2893-8.
8. Wong MS, Gu K, Heng D, Chew SK, Chew LS, Tai ES. The Singapore impaired glucose tolerance follow-up study: does the ticking clock go backward as well as forward? Diabetes Care. 2003;26(11):3024-30.
9. Coopmans C, Zhou TL, Henry RMA, Heijman J, Schaper NC, Koster A, et al. Both Prediabetes and Type 2 Diabetes Are Associated With Lower Heart Rate Variability: The Maastricht Study. Diabetes Care. 2020;43(5):1126-33.
10. van Agtmaal MJM, Houben A, de Wit V, Henry RMA, Schaper NC, Dagnelie PC, et al. Prediabetes Is Associated With Structural Brain Abnormalities: The Maastricht Study. Diabetes Care. 2018;41(12):2535-43.
11. Vergoossen LW, Schram MT, de Jong JJ, Stehouwer CD, Schaper NC, Henry RM, et al. White Matter Connectivity Abnormalities in Prediabetes and Type 2 Diabetes: The Maastricht Study. Diabetes Care. 2020;43(1):201-8.
12. van der Velde J, Koster A, Strotmeyer ES, Mess WH, Hilkman D, Reulen JPH, et al. Cardiometabolic risk factors as determinants of peripheral nerve function: the Maastricht Study. Diabetologia. 2020;63(8):1648-58.
13. Huru J, Leiviska I, Saarela V, Liinamaa MJ. Prediabetes influences the structure of the macula: thinning of the macula in the Northern Finland Birth Cohort. Br J Ophthalmol. 2020.
14. Hashimoto S, Yasuda M, Ninomiya T, Hata J, Yoshida D, Tahara-Asakuma T, et al. Foveal and Macular Thickness in a Japanese Population: The Hisayama Study. Ophthalmic Epidemiol. 2016;23(3):202-8.
15. Cheung CY, Chen D, Wong TY, Tham YC, Wu R, Zheng Y, et al. Determinants of quantitative optic nerve measurements using spectral domain optical coherence tomography in a population-based sample of non-glaucomatous subjects. Invest Ophthalmol Vis Sci. 2011;52(13):9629-35.
16. Mauschitz MM, Holz FG, Finger RP, Breteler MMB. Determinants of Macular Layers and Optic Disc Characteristics on SD-OCT: The Rhineland Study. Transl Vis Sci Technol. 2019;8(3):34.
17. Mauschitz MM, Bonnemaijer PWM, Diers K, Rauscher FG, Elze T, Engel C, et al. Systemic and Ocular Determinants of Peripapillary Retinal Nerve Fiber Layer Thickness Measurements in the European Eye Epidemiology (E3) Population. Ophthalmology. 2018;125(10):1526-36.
18. Lamparter J, Schmidtmann I, Schuster AK, Siouli A, Wasielica-Poslednik J, Mirshahi A, et al. Association of ocular, cardiovascular, morphometric and lifestyle parameters with retinal nerve fibre layer thickness. PLoS One. 2018;13(5):e0197682.
19. Ho H, Tham YC, Chee ML, Shi Y, Tan NYQ, Wong KH, et al. Retinal Nerve Fiber Layer Thickness in a Multiethnic Normal Asian Population: The Singapore Epidemiology of Eye Diseases Study. Ophthalmology. 2019;126(5):702-11.

20. Fujiwara K, Yasuda M, Hata J, Hirakawa Y, Hashimoto S, Ueda E, et al. Glucose Tolerance Levels and Circumpapillary Retinal Nerve Fiber Layer Thickness in a General Japanese Population: The Hisayama Study. Am J Ophthalmol. 2019;205:140-6.

21. Duan XR, Liang YB, Friedman DS, Sun LP, Wong TY, Tao QS, et al. Normal macular thickness measurements using optical coherence tomography in healthy eyes of adult Chinese persons: the Handan Eye Study. Ophthalmology. 2010;117(8):1585-94.

22. De Clerck EEB, Schouten J, Berendschot T, Goezinne F, Dagnelie PC, Schaper NC, et al. Macular thinning in prediabetes or type 2 diabetes without diabetic retinopathy: the Maastricht Study. Acta Ophthalmol. 2018;96(2):174-82.

23. De Clerck EEB, Schouten JSAG, Berendschot TTJM, Beckers HJM, Schaper NC, Schram MT, et al. Loss of Temporal Peripapillary Retinal Nerve Fibers in Prediabetes or Type 2 Diabetes Without Diabetic Retinopathy: The Maastricht Study. Invest Ophth Vis Sci. 2017;58(2).

24. Khawaja AP, Chua S, Hysi PG, Georgoulas S, Currant H, Fitzgerald TW, et al. Comparison of Associations with Different Macular Inner Retinal Thickness Parameters in a Large Cohort: The UK Biobank. Ophthalmology. 2020;127(1):62-71.

25. Jud P, Sourij H. Therapeutic options to reduce advanced glycation end products in patients with diabetes mellitus: A review. Diabetes Res Clin Pract. 2019;148:54-63.

26. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care. 2017;40(12):1631-40.

27. Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.

28. World Health Organization. Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycemia: Report of a WHO/IDF Consultation. . Geneva, Switzerland, . 2006.

29. van Eupen MG, Schram MT, van Sloten TT, Scheijen J, Sep SJ, van der Kallen CJ, et al. Skin Autofluorescence and Pentosidine Are Associated With Aortic Stiffening: The Maastricht Study. Hypertension. 2016;68(4):956-63.

30. Foreman YD, Brouwers MCGJ, van der Kallen CJH, Pagen DME, van Greevenbroek MMJ, Henry RMA, et al. Glucose Variability Assessed with Continuous Glucose Monitoring: Reliability, Reference Values, and Correlations with Established Glycemic Indices-The Maastricht Study. Diabetes Technol Ther. 2020;22(5):395-403.

31. Hong JT, Sung KR, Cho JW, Yun SC, Kang SY, Kook MS. Retinal nerve fiber layer measurement variability with spectral domain optical coherence tomography. Korean J Ophthalmol. 2012;26(1):32-8.

32. Qi Y, Koster A, van Boxtel M, Kohler S, Schram M, Schaper N, et al. Adulthood Socioeconomic Position and Type 2 Diabetes Mellitus-A Comparison of Education, Occupation, Income, and Material Deprivation: The Maastricht Study. Int J Environ Res Public Health. 2019;16(8).

33. Looman M, Feskens EJ, de Rijk M, Meijboom S, Biesbroek S, Temme EH, et al. Development and evaluation of the Dutch Healthy Diet index 2015. Public Health Nutr. 2017;20(13):2289-99.

34. van Dongen MC, Wijckmans-Duysens NEG, den Biggelaar LJ, Ocke MC, Meijboom S, Brants HA, et al. The Maastricht FFQ: Development and validation of a comprehensive food frequency questionnaire for the Maastricht study. Nutrition. 2019;62:39-46.

35. van der Berg JD, Willems PJ, van der Velde JH, Savelberg HH, Schaper NC, Schram MT, et al. Identifying waking time in 24-h accelerometry data in adults using an automated algorithm. J Sports Sci. 2016;34(19):1867-73.

36. van Dooren FE, Schram MT, Schalkwijk CG, Stehouwer CD, Henry RM, Dagnelie PC, et al. Associations of low grade inflammation and endothelial dysfunction with depression - The Maastricht Study. Brain Behav Immun. 2016;56:390-6.

37. Rodbard D. The challenges of measuring glycemic variability. J Diabetes Sci Technol. 2012;6(3):712-5.

38. Bikbov MM, Kazakbaeva GM, Gilmanshin TR, Zainullin RM, Arslangareeva, II, Salavatova VF, et al. Axial length and its associations in a Russian population: The Ural Eye and Medical Study. PLoS One. 2019;14(2):e0211186.

39.    Haeger A, Costa AS, Schulz JB, Reetz K. Cerebral changes improved by physical activity during cognitive decline: A systematic review on MRI studies. Neuroimage Clin. 2019;23:101933.
40.    Poulose SM, Miller MG, Scott T, Shukitt-Hale B. Nutritional Factors Affecting Adult Neurogenesis and Cognitive Function. Adv Nutr. 2017;8(6):804-11.
41.    Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009;20(4):488-95.
42.    Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care. 2019;42(8):1593-603.
43.    Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. Epidemiology (Sunnyvale). 2016;6(2).
44.    Yu DY, Cringle SJ, Balaratnasingam C, Morgan WH, Yu PK, Su EN. Retinal ganglion cells: Energetics, compartmentation, axonal transport, cytoskeletons and vulnerability. Prog Retin Eye Res. 2013;36:217-46.
45.    Greenland S. Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. Epidemiology. 1995;6(4):450-4.
46.    Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J. 2003;20(5):453-8.
47.    Britze J, Frederiksen JL. Optical coherence tomography in multiple sclerosis. Eye (Lond). 2018;32(5):884-8.
48.    Ceriello A. Glucose Variability and Diabetic Complications: Is It Time to Treat? Diabetes Care. 2020;43(6):1169-71.
49.    Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health. 2005;95 Suppl 1:S144-50.
50.    Toth P, Tarantini S, Csiszar A, Ungvari Z. Functional vascular contributions to cognitive impairment and dementia: mechanisms and consequences of cerebral autoregulatory dysfunction, endothelial impairment, and neurovascular uncoupling in aging. Am J Physiol Heart Circ Physiol. 2017;312(1):H1-H20.
51.    Bland JM, Altman DG. Correlation in restricted ranges of data. BMJ. 2011;342:d556.
52.    Mooy JM, Grootenhuis PA, de Vries H, Kostense PJ, Popp-Snijders C, Bouter LM, et al. Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. Diabetologia. 1996;39(3):298-305.
53.    Sharma S, Wakode S, Sharma A, Nair N, Dhobi M, Wani MA, et al. Effect of environmental toxicants on neuronal functions. Environ Sci Pollut Res Int. 2020;27(36):44906-21.

# Supplemental materials

## *Supplemental Methods*

### *Performance of optical coherence tomography (OCT) scans*

For a subset of participants (n=227) OCT scans were performed as part of a catch up visit. For these participants only there was a median time interval of 4.4 (interquartile range 4.0-5.9) years between OCT and all other measurements. We checked whether this affected the associations and this was not the case (data not shown).

### *Grading of OCT circle scans*

OCT scans were considered of sufficient quality if all the following criteria were met: good centering of the circular scan on the optic nerve head (examples of good, poor and very poor centering are shown in **Supplemental Figure S7.1**); complete (data of all 768 voxels was available); automatic quality ≥15 dB (an example of a scan with poor quality imaging is shown in Supplemental Figure S2); no measurement error present (examples of all assessed measurement errors are shown in **Supplemental Figure S7.2**). The percentage of agreement for selection of scans with sufficient quality ranged between 90% and 94% for four trained graders and was 70% for one grader (n=50 OCT scans per comparison).

### *Assessment of advanced glycation endproducts*

Advanced glycation endproducts (AGEs) were assessed with the AGE Reader (DiagnOptics Technologies BV, Groningen, the Netherlands). In brief, the AGE Reader is a desktop device that uses the characteristic fluorescent properties of certain AGEs to quantify their accumulation in the skin as skin autofluorescence (SAF; arbitrary units [AU]).[1] The AGE Reader illuminates a skin surface of 4 cm$^2$, shielded from other light, and uses the ratio of the reflection of fluorescent light (wavelength 420 to 600 nm) to non-fluorescent light (300–420 nm) to calculate SAF.

Due to the use of different versions of calibration software absolute SAF values assessed before the 4th September 2012 were 0.5 arbitrary units higher than the SAF values assessed after this date. To realign absolute SAF values we recalculated the SAF values assessed before the 4th of September 2012 by subtracting 0.5 arbitrary units from the SAF values assessed before the 4th September 2012.

### *Assessment of OGTT-based and CGM-based glycemic indices*

Incremental glucose peak (IGP), maximum glucose peak (i.e., the highest of the seven oral glucose tolerance test [OGTT] time points) and 1-h post-load glucose value were assessed in venous OGTT-derived samples. We have recently shown that the 1-hour post-load glucose value, maximum glucose peak, and IGP correlated most strongly with continuous glucose monitoring (CGM)-based indices of daily glucose variability[2].

Of these indices, the association between IGP and CGM-measured daily glucose variability indices was most consistent across different GMS strata, and, hence, IGP is considered the preferred OGTT-based index of daily glucose variability.

The rationale and methodology of CGM (iPro2 and Enlite Glucose Sensor; Medtronic, Tolochenaz, Switzerland) in The Maastricht Study have been described previously.[2] From 19 September 2016 until 13 September 2018, all new participants were invited to undergo CGM as part of their regular measurements at The Maastricht Study. To accelerate the inclusion process and to ensure the inclusion of a sufficient number of participants with prediabetes and type 2 diabetes, a selected group of participants who had recently visited The Maastricht Study was re-invited to undergo CGM as a separate research visit ('catch-up visit'). For individuals who participated in the catch-up visit, there was an average time period ("visit interval") of 2.1 years between CGM and other measurements. The CGM device was worn on the lower abdomen and recorded subcutaneous interstitial glucose values (range: 2.2 - 22.2 mmol/L) every five minutes for a 7-day period. Participants were asked to perform self-measurements of blood glucose four times daily (Contour Next; Ascensia Diabetes Care, Mijdrecht, the Netherlands) for retrospective calibration of the CGM device. Participants were blinded to the CGM recording, but not to the self-measured values. Diabetes medication use was allowed during the CGM period, and no dietary or physical activity instructions were given.

The first 24 hours of CGM were excluded, because of insufficient calibration. Next, we excluded individuals of whom less than 24 hours of recording (less than one data day) remained. Then, we calculated mean sensor glucose (mmol/L), standard deviation (mmol/L), coefficient of variation (standard deviation / mean sensor glucose * 100%), and time in range (TIR; i.e., % of time between 3.9 – 10.0 mmol/L) for the total recording period. Based on international consensus, we used standard deviation and coefficient of variation as indices of daily glucose variability.[3] The CGM-assessed coefficient of variation, which is intrinsically adjusted for mean sensor glucose, was not used as main determinant, as use of a ratio variable may introduce bias.[4,5] TIR is an emerging glycemic index that is partly determined by daily glucose variability.[6]

*Ridge regression*

Because we presumed the reliability of 'model 3 + mean sensor glucose' to be negatively impacted by multicollinearity, due to the strong correlation between CGM-assessed standard deviation and mean sensor glucose (rho = 0.69),[7] we additionally performed ridge regression. It is a L2-regularized form of linear regression and a valid statistical method to counter a degree of model instability caused by multicollinearity.[8] Ridge regression estimates are computed according to the combination of the residual sum of squares characteristic of regular linear regression and predefined penalization

of the coefficients (i.e., $Ridge = RRS + \frac{1}{n} * \lambda * \sum_{j=1}^{p} \beta_j^2$, where RSS is the residual sum of squares, n is the sample size, $\lambda$ is the chosen amount of penalization, and $\sum_{j=1}^{p} \beta_j^2$ represents the sum of all squared regression coefficients). As such, it slightly biases the regression coefficients and can strongly reduce inflated variances that arise when high multicollinearity is present. We pragmatically chose the level of penalization based on the lambda ($\lambda$) required to reduce the variance inflation factor (VIF) of model 3.1 back to the VIF of model 3 (or halfway back). The ridge regression results are presented as: standardized regression coefficient (st.$\beta$) (95%CI), p-value. The median st.$\beta$s (95%CIs) were estimated with use of resampling (1,000 bootstraps).

## Supplemental results

### Additional analyses

Quantitatively similar results were observed in a range of sensitivity analyses. First, associations remained similar after we additionally adjusted for physical activity and dietary habits, or spherical equivalent and intraocular pressure (**Supplemental Table S7.5**). Second, associations were generally comparable when we additionally adjusted for kidney variables, history of cardiovascular disease, plasma biomarkers of low-grade inflammation, retinopathy, and glaucoma (**Supplemental Table S7.5**). Third, associations were similar when we excluded participants with retinopathy, glaucoma, or other types of diabetes (**Supplemental Table S7.6**). Fourth, associations were numerically comparable when we replaced IGP with other OGTT-based indices of glucose variability and CGM-assessed standard deviation with other CGM-assessed measures of daily glucose variability (**Supplemental Table S7.7**). Fifth, the strength of the association of CGM-assessed mean sensor glucose with RNFL thickness was numerically comparable to the strength of the associations of measures of glycemia with RNFL thickness (**Supplemental Table S7**). Sixth, associations remained similar after replacement of waist circumference with BMI; of educational status with occupational status or income level; and of office systolic BP with office diastolic BP, or systolic or diastolic 24-hour ambulatory BP (**Supplemental Tables S7.8** and **S7.9**). Seventh, the association between SAF was not altered after additional adjustment for FPG, 2-hour post-load glucose, or HbA$_{1c}$ (**Supplemental Table S7.10**). Eighth, associations of IGP and CGM-assessed standard deviation with RNFL thickness were similar after replacement of HbA$_{1c}$ or mean sensor glucose with other measures of glycemia (**Supplemental Table S7.11**). Ninth, the associations between CGM-assessed standard deviation and RNFL thickness were similar after exclusion of individuals with insufficient recording days or recording data gaps (**Supplemental Table S7.12**). Exclusion of participants with a visit interval, most of whom had type 2 diabetes,[2] strongly attenuated the association between CGM-assessed standard deviation and

7

RNFL thickness. Tenth, ridge regression did not yield materially different results for the association of CGM-assessed standard deviation with RNFL thickness (**Supplemental Table S7.13**). Last, after full adjustment (model 3), longer duration of diabetes was significantly associated with lower RNFL thickness (per SD, -0.07 [-0.13; -0.001]; **Supplemental Table S7.14**).

## *References*

1. van Eupen MG, Schram MT, van Sloten TT, Scheijen J, Sep SJ, van der Kallen CJ, et al. Skin Autofluorescence and Pentosidine Are Associated With Aortic Stiffening: The Maastricht Study. Hypertension. 2016;68(4):956-63.
2. Foreman YD, Brouwers MCGJ, van der Kallen CJH, Pagen DME, van Greevenbroek MMJ, Henry RMA, et al. Glucose Variability Assessed with Continuous Glucose Monitoring: Reliability, Reference Values, and Correlations with Established Glycemic Indices-The Maastricht Study. Diabetes Technol Ther. 2020;22(5):395-403.
3. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care. 2017;40(12):1631-40.
4. Curran-Everett D. Explorations in statistics: the analysis of ratios and normalized data. Adv Physiol Educ. 2013;37(3):213-9.
5. Rodbard D. Glycemic variability: challenges in interpretation. Diabetes Technol Ther. 2015;17(6):370-2.
6. Rodbard D. Glucose Time In Range, Time Above Range, and Time Below Range Depend on Mean or Median Glucose or $HbA_{1c}$, Glucose Coefficient of Variation, and Shape of the Glucose Distribution. Diabetes Technol Ther. 2020.
7. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. Epidemiology (Sunnyvale). 2016;6(2).
8. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carre G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography. 2013;36(1):27-46.

## Supplemental figures
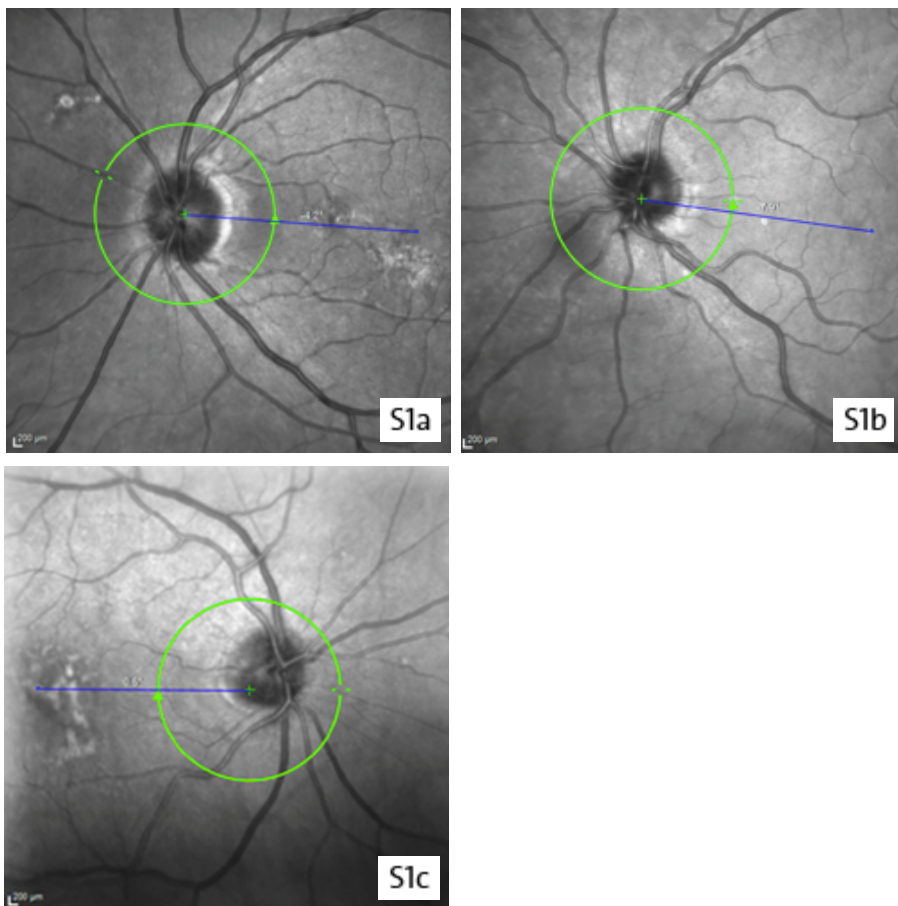


**Figure S7.1  Examples of quality of centering of circular scans on the optic nerve head.**
Supplemental Figure S7.1 shows examples of quality of centering of the circular scan on the optic nerve head: S1a shows good quality, S1b shows poor quality, and S1c shows very poor quality.
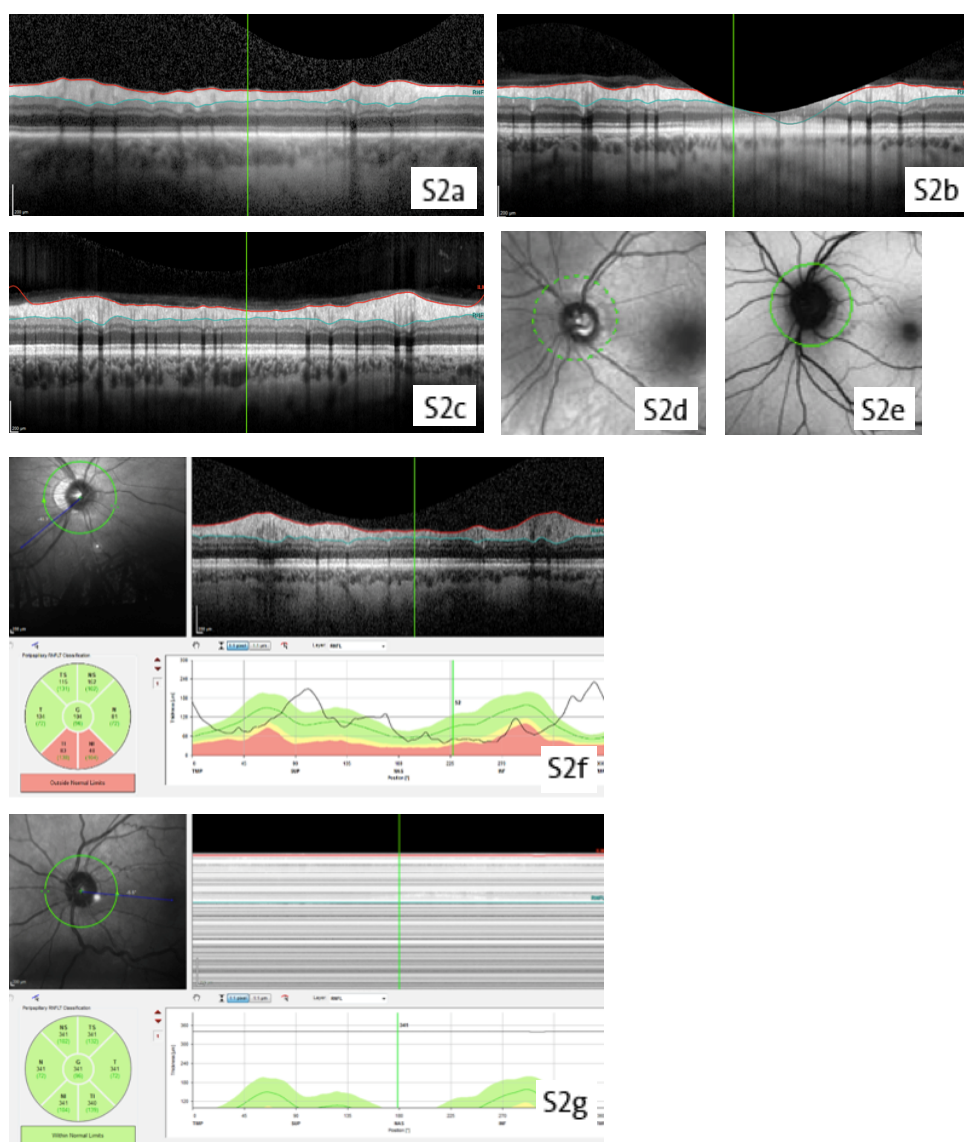
**Figure S7.2  Examples of poor quality and scan errors.** S2a: Example of poor imaging quality (Signal-to-noise ratio<15 dB); S2b: OCT device too close to the eye; S2c: RNFL layer incorrectly defined; S2d: incorrect circle position (dashed line); S2e: autofluorescence on; S2f: participant does not look in the correct direction; S2g: technical problem with OCT device. OCT, optical coherence tomography; RNFL, retinal nerve fiber layer thickness

## *Supplemental tables*

**Table S7.1** Additional general study population characteristics according to tertiles of retinal nerve fiber layer thickness in the study population with complete data on glucose metabolism status.

| Characteristic | Total study population | RNFL thickness | | | Number of participants with missing data |
| --- | --- | --- | --- | --- | --- |
| | | Tertile 1 (high) | Tertile 2 (middle) | Tertile 3 (low) | |
| | (n=5,455) | (n=1,818) | (n=1,819) | (n=1,818) | |
| BMI (kg/m2) | 26.8 ± 4.4 | 26.8 ± 4.5 | 26.7 ± 4.3 | 26.9 ± 4.3 | 1 |
| Income level (euro) | 1,856 [1,502-2,386] | 1,856 [1,502-2,386] | 1,875 [1,503-2,386] | 1,856 [1,503-2,386] | 1,288 |
| Occupational status | | | | | 3,618 |
| Low | 572 (31.1) | 195 (34.2) | 190 (31.2) | 187 (28.4) | 535 |
| Intermediate | 666 (36.3) | 207 (36.3) | 231 (37.9) | 228 (34.7) | 535 |
| High | 599 (32.6) | 168 (29.5) | 188 (30.9) | 243 (36.9) | 388 |
| Systolic ambulatory 24-hour blood pressure (mmHg) | 118.5 ± 11.3 | 117.8 ± 11.6 | 118.6 ± 11.2 | 119.3 ± 11.3 | 768 |
| Diastolic ambulatory 24-hour blood pressure (mmHg) | 72.9 ± 7.1 | 72.4 ± 7.2 | 73.0 ± 7.1 | 73.3 ± 6.9 | 275 |
| Dutch Healthy diet score (points) | 84.0 ± 15.0 | 84.5 ± 15.1 | 84.6 ± 15.1 | 82.8 ± 14.9 | 319 |
| Physical activity (hours/day) | 2.0 ± 0.7 | 2.0 ± 0.7 | 2.0 ± 0.7 | 2.0 ± 0.7 | 319 |
| Spherical equivalent (diopter) | 0.13 [-1.19-1.06] | 0.6 [-0.4-1.6] | 0.1 [-1.1-1.1] | -0.5 [-2.9-0.6] | 137 |
| Intraocular pressure (mmHg) | 14.0 [11.7-16.0] | 13.5 [11.5-15.7] | 13.8 [11.5-15.7] | 14.3 [12.0-16.5] | 101 |
| Glaucoma | 234 (4.6) | 63 (3.7) | 53 (3.1) | 118 (6.9) | 3 |
| Retinopathy | 79 (1.5) | 8 (1.3) | 5 (0.8) | 12 (1.7) | 23 |
| History of CVD | 1,337 (16.9) | 279 (15.4) | 293 (16.2) | 324 (17.9) | 23 |
| eGFR, ml/min/1.73m2 | 82.3 ± 14.0 | 83.0 ± 13.8 | 82.1 ± 13.9 | 81.9 ± 14.1 | 3,049 |
| Urinary albumin excretion (mg/24 hours) | 5.4 [3.4-10.0] | 5.2 [3.4-9.8] | 5.4 [3.4-9.5] | 5.5 [3.4-10.6] | |
| Albuminuria | 424 (7.8) | 130 (7.2) | 137 (7.6) | 157 (8.7) | |
| Biomarkers of low-grade inflammation | | | | | |
| C-reactive protein, μg/ml | 1.2 [0.6-2.7] | 1.2 [0.6-2.7] | 1.2 [0.6-2.6] | 1.2 [0.7-2.8] | |
| Serum amyloid A, μg/ml | 3.2 [2.0-5.4] | 3.2 [2.0-5.5] | 3.2 [2.1-5.2] | 3.3 [2.1-5.6] | |
| Interleukin-6, pg/ml | 0.6 [0.4-0.9] | 0.6 [0.4-0.9] | 0.5 [0.4-0.8] | 0.6 [0.4-1.0] | |
| Interleukin-8, pg/ml | 4.0 [3.2-5.2] | 3.9 [3.2-4.9] | 4.1 [3.2-5.2] | 4.2 [3.2-5.3] | |
| Tumor necrosis factor α, pg/ml | 2.2 [1.9-2.6] | 2.1 [1.8-2.6] | 2.2 [1.9-2.5] | 2.2 [1.9-2.6] | 257 |
| Duration of diabetes (years) | 4.0 [0-9.0] | 3.0 [0-9.0] | 3.0 [0-8.0] | 4.0 [1.0-10.8] | 3,048 |
| Maximum glucose peak (mmol/L) | 9.5 [7.9-12.4] | 9.3 [7.8-11.7] | 9.5 [7.9-12.2] | 9.8 [8.1-13.4] | 2,924 |
| 1-hour post-load glucose (mmol/L) | 8.6 [6.5-11.9] | 8.4 [6.3-11.2] | 8.6 [6.6-11.7] | 8.8 [6.7-12.6] | |

7

**Table S7.1** (continued)

| Characteristic | Total study population | RNFL thickness | | | Number of participants with missing data |
| --- | --- | --- | --- | --- | --- |
| | | Tertile 1 (high) | Tertile 2 (middle) | Tertile 3 (low) | |
| | (n=5,455) | (n=1,818) | (n=1,819) | (n=1,818) | |
| CGM: indices and methodology | | | | | 4,833 |
| Mean sensor glucose, mmol/L | 6.1 [5.9-6.8] | 6.1 [5.7-6.6] | 6.0 [5.7-6.7] | 6.1 [5.7-6.9] | |
| Coefficient of variation, % | 14.5 [11.7-17.8] | 14.6 [11.6-17.4] | 14.1 [11.5-17.6] | 14.8 [11.9-18.7] | |
| Time in range, % | 99.7 [96.8-100.0] | 99.7 [96.8-100.0] | 99.7 [97.8-100.0] | 99.7 [96.1-100.0] | |
| CGM during regular visit | 409 (65.8) | 123 (63.4) | 135 (65.9) | 151 (67.7) | |
| CGM during 'catch-up visit' | 213 (34.2) | 71 (36.6) | 70 (34.1) | 72 (32.2) | |
| Insufficient recording day, yes vs. no | 8 (1.3) | 1 (0.5) | 2 (1.0) | 5 (2.2) | |
| CGM recording data gap, yes vs. no | 49 (7.9) | 18 (9.3) | 14 (6.8) | 17 (7.6) | |
| CGM visit interval* (years) | 2.1 [2.0-2.2] | 2.1 [2.0-2.2] | 2.1 [2.1-2.2] | 2.1 [2.0-2.2] | |

Data are presented as mean ± standard deviation, median [interquartile range] or number (%). BMI, body-mass index; CGM, continuous glucose monitoring; CVD, cardiovascular disease; eGFR, estimated glomerular filtration rate. * Of 'catch-up visit' participants only (n=213).

**Table S7.2** General study population characteristics of the included and excluded participants for the study population of glucose metabolism status.

| Characteristic | Included study population (n= 5,455) | Missing data in/excluded | Excluded study population (n=2,550) |
|---|---|---|---|
| Age (years) | 59.5 ± 8.6 | 0/0 | 60.6 ± 8.7 |
| Men | 2,665 (48.9) | 0/0 | 1356 (53.2) |
| Educational level | | 0/116 | |
|   Low | 1,914 (35.1) | | 821 (33.7) |
|   Medium | 1,519 (27.8) | | 647 (26.6) |
|   High | 2,022 (37.1) | | 966 (39.7) |
| Glucose metabolism status | | 0/0 | |
| NGM | 3,366 (61.7) | | 1,479 (58.0) |
| Prediabetes | 820 (15.0) | | 370 (14.5) |
| Type 2 diabetes | 1,239 (22.7) | | 681 (26.7) |
| Other type of diabetes | 30 (0.5) | | 20 (0.8) |
| Measures of glycemia | | | |
| Fasting plasma glucose (mmol/L) | 5.9 ± 1.5 | 1/0 | 6.0 ± 1.8 |
| 2-hour post-load glucose (mmol/L) | 6.2 [4.9-8.6] | 275/212 | 6.2 [4.9-8.8] |
| HbA$_{1c}$ (mmol/mol) | 39.2 ± 9.1 | 6/8 | 40.6 ± 10.4 |
| HbA$_{1c}$ (%) | 5.7 ± 0.8 | 6/8 | 5.9 ± 1.0 |
| SAF (AU) | 2.2 ± 0.5 | 323/296 | 2.2 ± 0.5 |
| Indices of daily glucose variability | | | |
| Incremental glucose peak (mmol/L) | 4.1 [2.7-6.5] | 3048/1368 | 4.2 [2.7-6.6] |
| CGM-assessed standard deviation (mmol/L) | 0.86 [0.68-1.21] | 4833/2319 | 0.79 [0.68-1.08] |
| Use of glucose-modifying medication, yes vs. no | 927 (17.0) | 0/6 | 536 (21.1) |
| Duration of diabetes (years) | 4.0 [0-9.0] | 257/155 | 4.5 [1.0-10.0] |
| Waist circumference, men (cm) | 100.4 ± 11.6 | 0/2 | 102.1 ± 12.5 |
| Waist circumference, women (cm) | 89.1 ± 12.5 | 0/3 | 90.7 ± 13.7 |
| Total-to-HDL cholesterol ratio | 3.4 [2.8-4.2] | 0/4 | 3.5 [2.8-4.4] |
| Use of lipid-modifying medication, yes vs. no | 1,687 (30.9) | 0/6 | 864 (34.0) |
| Office systolic blood pressure (mmHg) | 133.2 ± 17.7 | 0/3 | 134.6 ± 18.3 |
| Office diastolic blood pressure (mmHg) | 75.5 ± 9.8 | 1/3 | 75.4 ± 9.8 |
| Use of antihypertensive medication, yes vs. no | 1,983 (36.4) | 0/6 | 1039 (40.8) |
| Smoking status | | 0/65 | |
|   Never | 2,101 (38.5) | | 858 (34.5) |
|   Former | 2,666 (48.9) | | 1253 (50.4) |
|   Current | 688 (12.6) | | 374 (15.1) |
| Alcohol consumption | | 0/66 | |
|   None | 995 (18.2) | | 478 (19.2) |
|   Moderate | 3,181 (58.3) | | 1,447 (58.3) |
|   High | 1,279 (23.4) | | 559 (22.5) |
| RNFL thickness (µm) | 94.8 ± 10.8 | 0/2454* | 95.1 ± 11.5 |

Data are presented as mean ± standard deviation, median [interquartile range] or number (%). * Number of participants with missing data represents the number of participants that missed RNFL thickness assessment in both eyes or had a RNFL thickness assessment of insufficient quality for both eyes. CGM, continues glucose monitoring; HbA$_{1c}$, hemoglobin A1c; HDL, high-density lipoprotein; NGM, normal glucose metabolism; RNFL, retinal nerve fiber layer; SAF, skin autofluorescence AU, arbitrary units; NA, not applicable.

7

**Table S7.3   General study population characteristics of the included and excluded participants for the study populations of indices of daily glucose variability (incremental glucose peak and continuous glucose monitoring-assessed standard deviation).**

| Characteristic | Incremental glucose peak | | | Continuous glucose monitoring-assessed standard deviation | | |
|---|---|---|---|---|---|---|
| | Included study Population (n=2,407) | Missing data in/excluded | Excluded study population (n=5,598) | Included study population (n=622) | Missing data in/excluded | Excluded study population (n=7,383) |
| Age (years) | 59.5 ± 8.2 | 0/0 | 60.0 ± 8.9 | 59.8 ± 8.5 | 0/0 | 59.8 ± 8.7 |
| Men | 1,211 (50.3) | 0/0 | 2810 (50.2) | 318 (51.1) | 0/0 | 3703 (50.2) |
| Educational level | | 0/116 | | | 0/116 | |
| Low | 794 (33.0) | | 1941 (35.4) | 205 (33.0) | | 2530 (34.8) |
| Medium | 691 (28.7) | | 1475 (26.9) | 168 (27.0) | | 1998 (27.5) |
| High | 922 (38.3) | | 2066 (37.7) | 249 (40.0) | | 2739 (37.7) |
| Glucose metabolism status | | 0/0 | | | 0/0 | |
| NGM | 1,455 (60.4) | | 3,390 (60.6) | 328 (52.7) | | 4517 (61.2) |
| Prediabetes | 430 (17.9) | | 760 (13.6) | 140 (22.5) | | 1050 (14.2) |
| Type 2 diabetes | 522 (21.7) | | 1398 (25.0) | 152 (24.4) | | 1768 (23.9) |
| Other type of diabetes | 0 (0.0) | | 50 (0.9) | 2 (0.3) | | 48 (0.7) |
| Measures of glycemia | | | | | | |
| Fasting plasma glucose (mmol/L) | 5.8 ± 1.1 | 0/1 | 6.0 ± 1.8 | 5.9 ± 1.5 | 0/1 | 5.9 ± 1.6 |
| 2-hour post-load glucose (mmol/L) | 6.4 [5.1-9.1] | 0/487 | 6.1 [4.9-8.4] | 6.8 [5.3-9.4] | 29/458 | 6.1 [4.9-8.6] |
| HbA1c (mmol/mol) | 38.9 ± 6.9 | 0/14 | 40.0 ± 10.5 | 39.6 ± 8.8 | 0/14 | 39.7 ± 9.6 |
| HbA1c (%) | 5.7 ± 0.6 | 0/14 | 5.8 ± 1.0 | 5.8 ± 0.8 | 0/14 | 5.8 ± 0.9 |
| SAF (AU) | 2.2 ± 0.4 | 112/507 | 2.2 ± 0.5 | 2.1 ± 0.5 | 49/570 | 2.2 ± 0.5 |
| Indices of daily glucose variability | | | | | | |
| Incremental glucose peak (mmol/L) | 4.1 [2.7-6.5] | 0/4416 | 4.2 [2.7-6.6] | 4.4 [3.1-6.9] | 75/4341 | 4.1 [2.7-6.4] |
| CGM-assessed standard deviation (mmol/L) | 0.84 [0.68-1.16] | 1860/5292 | 0.84 [0.69-1.23] | 0.86 [0.68-1.21] | 0/7152 | 0.79 [0.68-1.08] |
| Use of glucose-modifying medication, yes vs. no | 350 (14.5) | 0/6 | 1113 (19.9) | 88 (14.1) | 0/6 | 1375 (18.6) |
| Duration of diabetes | 2.0 [0.0-6.0] | 97/315 | 5.0 [1.0-11.0] | 1.0 [0.0-4.0] | 22/390 | 4.0 [1.0-10.0] |
| Waist circumference, men (cm) | 100.7 ± 11.3 | 0/2 | 101.1 ± 12.2 | 102.6 ± 11.8 | 0/2 | 100.8 ± 12.0 |
| Waist circumference, women (cm) | 89.4 ± 12.2 | 0/3 | 89.6 ± 13.2 | 90.8 ± 12.4 | 0/3 | 89.4 ± 12.9 |
| Total-to-HDL cholesterol ratio | 3.4 [2.8-4.2] | 0/4 | 3.4 [2.8-4.3] | 3.5 [2.8-4.4] | 0/4 | 3.4 [2.8-4.2] |

Table S7.3  (continued)

| | Incremental glucose peak | | | Continuous glucose monitoring-assessed standard deviation | | |
|---|---|---|---|---|---|---|
| | Included study Population (n=2,407) | Missing data in/excluded | Excluded study population (n=5,598) | Included study population (n=622) | Missing data in/excluded | Excluded study population (n=7383) |
| Use of lipid-modifying medication, yes vs. no | 735 (30.5) | 0/6 | 1816 (32.5) | 159 (25.6) | 0/6 | 2392 (32.4) |
| Office systolic blood pressure (mmHg) | 134.0 ± 17.8 | 0/3 | 133.5 ± 18.0 | 134.1 ± 18.3 | 0/3 | 133.6 ± 17.9 |
| Office diastolic blood pressure (mmHg) | 76.3 ± 10.0 | 0/4 | 75.2 ± 9.7 | 75.7 ± 10.2 | 0/4 | 75.5 ± 9.8 |
| Use of antihypertensive medication, yes vs. no | 878 (36.5) | 0/6 | 2144 (38.3) | 231 (37.1) | 0/6 | 2791 (37.8) |
| Smoking status | | 0/65 | | | 0/65 | |
| Never | 863 (35.9) | | 2096 (37.9) | 238 (38.3) | | 2721 (37.2) |
| Former | 1,238 (51.4) | | 2681 (48.5) | 308 (49.5) | | 3611 (49.3) |
| Current | 306 (12.7) | | 756 (13.7) | 76 (12.2) | | 986 (13.5) |
| Alcohol consumption | | 0/66 | | | 0/66 | |
| None | 401 (16.7) | | 1072 (19.4) | 104 (16.7) | | 1369 (18.7) |
| Moderate | 399 (58.1) | | 3,229 (58.4) | 400 (64.3) | | 4228 (57.8) |
| High | 607 (25.2) | | 1231 (22.3) | 118 (19.0) | | 1720 (23.5) |
| RNFL thickness (µm) | 94.5 ± 11.1 | 0/2454* | 95.1 ± 10.6 | 94.5 ± 10.8 | 0/2454* | 94.9 ± 10.9 |

Data are presented as mean ± standard deviation, median [interquartile range] or number (%). * Number of participants with missing data represents the number of participants that missed RNFL thickness assessment in both eyes or had a RNFL thickness assessment of insufficient quality for both eyes. CGM, continues glucose monitoring; HbA1c, hemoglobin A1c; HDL, high-density lipoprotein; NGM, normal glucose metabolism; RNFL, retinal nerve fiber layer; SAF, skin autofluorescence AU, arbitrary units; NA, not applicable.

7

**Table S7.4    P-values of interaction terms with glucose metabolism status and sex.**

| | Number of participants | Prediabetes p-value | Type 2 diabetes p-value | Sex p-value |
|---|---|---|---|---|
| **Determinant** | | | | |
| Prediabetes versus NGM | 5,455 | NA | NA | 0.07 |
| Type 2 diabetes versus NGM | 5,455 | NA | NA | 0.49 |
| Fasting plasma glucose | 5,454 | 0.81 | 0.85 | 0.69 |
| 2-hour post-load glucose | 5,180 | 0.61 | 0.77 | 0.26 |
| HbA$_{1c}$ | 5,449 | 0.27 | 0.77 | 0.98 |
| Skin autofluorescence | 5,132 | 0.97 | 0.90 | 0.35 |
| Incremental glucose peak | 2,407 | 0.30 | 0.76 | 0.76 |
| CGM-assessed standard deviation | 622 | 0.59 | 0.55 | 0.85 |

P-values represent the p-values for the interaction terms of sex, glucose metabolism status (i.e., prediabetes versus normal glucose metabolism status or type 2 diabetes versus normal glucose metabolism status) with determinants (e.g., sex*HbA$_{1c}$) in the associations of glucose metabolism status, measures of glycemia, incremental glucose peak, and CGM-assessed standard deviation with RNFL thickness. Variables in the model in addition to determinants and interaction term(s) with sex, glucose metabolism status and are: age, sex, educational status, office systolic blood pressure, total cholesterol/HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication, waist circumference, smoking status, and alcohol consumption status. In addition, for interaction analyses with glucose metabolism status, glucose metabolism status was also entered in the model. Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1. p-value <0.05 denotes statistically significant interaction. CGM, continues glucose monitoring; GMS, glucose metabolism status; HbA$_{1c}$, hemoglobin A$_{1c}$; NA, not applicable; NGM, normal glucose metabolism; RNFL, retinal nerve fiber layer.

**Table S7.5** Associations of GMS, measures of glycemia, and indices of daily glucose variability with RNFL thickness after additional adjustment for dietary intake and physical activity (model 4A), for spherical equivalent and intraocular pressure (model 4B), or for eGFR, urinary albumin excretion, history of CVD, plasma biomarkers of low-grade inflammation, retinopathy, and glaucoma (model 4C).

| | RNFL thickness, per SD | | | | | |
| | Model 4A* | | Model 4B | | Model 4C | |
| | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) |
|---|---|---|---|---|---|---|
| **Glucose metabolism status** | | | | | | |
| Prediabetes versus NGM | 4,353 | -0.07 (-0.16; 0.02) | 5,107 | -0.06 (-0.14; 0.02) | 2,148 | -0.08 (-0.20; 0.05) |
| Type 2 diabetes versus NGM | 4,353 | **-0.17 (-0.26; -0.08)** | 5,107 | **-0.15 (-0.23; -0.07)** | 2,148 | **-0.20 (-0.32; -0.07)** |
| **Measures of glycemia** | | | | | | |
| Fasting plasma glucose, per SD | 4,352 | **-0.04 (-0.07; -0.002)** | 5,106 | **-0.03 (-0.06; -0.002)** | 2,147 | -0.04 (-0.08; 0.01) |
| 2-hour post-load, per SD | 4,130 | **-0.06 (-0.10; -0.02)** | 4,873 | **-0.05 (-0.08; -0.02)** | 2,015 | **-0.07 (-0.12; -0.02)** |
| HbA1c, per SD | 4,348 | **-0.04 (-0.08; -0.01)** | 5,102 | **-0.04 (-0.07; -0.01)** | 2,145 | -0.04 (-0.09; 0.01) |
| Skin autofluorescence, per SD | 4,091 | **-0.04 (-0.08; -0.01)** | 4,805 | **-0.04 (-0.07; -0.01)** | 2,071 | **-0.07 (-0.12; -0.02)** |
| **Indices of daily glucose variability** | | | | | | |
| Incremental glucose peak, per SD | 1,187 | **-0.08 (-0.13; -0.02)** | 2,251 | -0.03 (-0.08; 0.02) | 1,637 | -0.05 (-0.10; 0.01) |
| Model 4 + HbA1c | 1,187 | **-0.08 (-0.15; -0.01)** | 2,251 | -0.03 (-0.09; 0.02) | 1,637 | -0.06 (-0.13; 0.01) |
| CGM- assessed standard deviation, per SD | 439 | -0.09 (-0.20; 0.02) | 615 | -0.06 (-0.15; 0.02) | 592 | -0.06 (-0.15; 0.03)† |
| Model 4 + MSG | 439 | -0.08 (-0.24; 0.08) | 615 | -0.07 (-0.19; 0.06) | 592 | -0.07 (-0.21; 0.07)† |

Standardized regression coefficient (stβ) represents the difference in RNFL thickness in SD for individuals with type 2 diabetes or prediabetes versus NGM or per SD greater measure of glycemia or daily glucose variability. In the GMS, fasting plasma glucose, 2-hour post-load glucose, HbA1c, and skin autofluorescence study populations 1 SD corresponds with 10.8 μm for RNFL thickness, and (respectively) 1.5 mmol/L for fasting plasma glucose, 4.0 mmol/L for 2-hour post-load glucose, 0.8% or 9.1 mmol/mol for HbA1c, and 0.5 AU for skin autofluorescence (model 4A). For incremental glucose peak, 1 SD corresponds with 11.1 μm for RNFL thickness and 2.9 mmol/L for incremental glucose peak (model 4A). For CGM-assessed standard deviation, 1 SD corresponds with 10.6 μm for RNFL thickness and 0.62 mmol/L for CGM-assessed standard deviation (model 4A). In models 4B and 4C values per SD were numerically comparable. Bold denotes p<0.05. Variables in model 3: age, sex, educational status, office systolic blood pressure, total cholesterol/HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication, waist circumference, smoking status, and alcohol consumption status. Incremental glucose peak was additionally adjusted for HbA1c. CGM-assessed standard deviation was additionally adjusted for MSG. Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1. *Diet intake was entered in the model as diet score minus the alcohol component to avoid multicollinearity. †Plasma biomarkers of low-grade inflammation were not available in the CGM study population, and were, thus, not included in the analyses. AU, arbitrary unit; CGM, continuous glucose monitoring; CI, confidence interval; CVD, cardiovascular disease; eGFR, estimated glomerular filtration rate; GMS, glucose metabolism status; HbA1c, hemoglobin A1c; MSG, mean sensor glucose; NGM, normal glucose metabolism; SD, standard deviation; RNFL, retinal nerve fiber layer.

7

**Table S7.6**   Associations of glucose metabolism status, measures of glycemia, and indices of daily glucose variability with RNFL thickness after exclusion of individuals with retinopathy (model 3A), glaucoma (model 3B), or other types of diabetes (model 3C).

| | RNFL thickness, per SD | | | | | |
| | Model 3A | | Model 3B | | Model 3C | |
| | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) |
| --- | --- | --- | --- | --- | --- | --- |
| **Glucose metabolism status** | | | | | | |
| Prediabetes versus NGM | 5,239 | -0.06 (-0.14; 0.02) | 4,902 | -0.02 (-0.11; 0.06) | 5,425 | -0.04 (-0.12; 0.04) |
| Type 2 diabetes versus NGM | 5,239 | **-0.17 (-0.26; -0.09)** | 4,902 | **-0.19 (-0.27; -0.10)** | 5,425 | **-0.16 (-0.24; -0.08)** |
| **Measures of glycemia** | | | | | | |
| Fasting plasma glucose, per SD | 5,238 | **-0.05 (-0.08; -0.02)** | 4,901 | **-0.05 (-0.08; -0.01)** | 5,424 | **-0.04 (-0.07; -0.01)** |
| 2-hour post-load glucose, per SD | 5,010 | **-0.06 (-0.09; -0.02)** | 4,689 | **-0.06 (-0.10; -0.03)** | 5,180 | **-0.06 (-0.09; -0.02)** |
| HbA1c, per SD | 5,233 | **-0.06 (-0.09; -0.03)** | 4,897 | **-0.05 (-0.08; -0.02)** | 5,419 | **-0.05 (-0.08; -0.02)** |
| Skin autofluorescence, per SD | 4,927 | **-0.05 (-0.08; -0.01)** | 4,616 | **-0.04 (-0.07; -0.01)** | 5,103 | **-0.04 (-0.07; -0.01)** |
| **Indices of daily glucose variability** | | | | | | |
| Incremental glucose peak, per SD | 2,322 | **-0.06 (-0.11; -0.01)** | 1,542 | **-0.05 (-0.11; -0.01)** | 2,407 | **-0.06 (-0.11; -0.01)** |
| Model 3 + HbA1c | 2,322 | **-0.06 (-0.12; -0.01)** | 1,542 | -0.05 (-0.12; 0.02) | 2,407 | **-0.06 (-0.12; -0.01)** |
| CGM- assessed standard deviation, per SD | 594 | -0.08 (-0.18; 0.01) | 564 | -0.07 (-0.17; 0.03) | 620 | -0.08 (-0.17; 0.01) |
| Model 3 + MSG | 594 | -0.06 (-0.20; 0.08) | 564 | -0.07 (-0.22; 0.07) | 620 | -0.06 (-0.19; 0.08) |

Standardized regression coefficient (stβ) represents the difference in RNFL thickness in SD for individuals with type 2 diabetes or prediabetes versus individuals with NGM or per SD greater measure of glycemia or daily glucose variability. In the GMS, fasting plasma glucose, 2-hour post-load glucose, HbA1c, and skin autofluorescence study populations, 1 SD corresponds with 10.8 μm for RNFL thickness,1.5 mmol/L for fasting plasma glucose, 4.0 mmol/L for 2-hour post-load glucose, 0.8% or 9.1 mmol/mol for HbA1c, and 0.5 AU for skin autofluorescence (model 3A). For incremental glucose peak, 1 SD corresponds with 11.1 μm for RNFL thickness and 2.9 mmol/L for incremental glucose peak (model 3A). For CGM-assessed standard deviation, 1 SD corresponds with 10.8 μm for RNFL thickness and 0.57 mmol/L for CGM-assessed standard deviation (model 3A). In models 3B and 3C values per SD were numerically comparable. Bold denotes p<0.05. Variables in model 3: age, sex, educational status, office systolic blood pressure, total cholesterol/HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication, waist circumference, smoking status, and alcohol consumption status. Incremental glucose peak was additionally adjusted for HbA1c (model 3 + HbA1c). CGM- assessed standard deviation was additionally adjusted for MSG (model 3 + MSG). Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1. AU, arbitrary unit; CGM, continues glucose monitoring; CI, confidence interval; GMS, glucose metabolism status; HbA1c, hemoglobin A1c; MSG, mean sensor glucose; NGM, normal glucose metabolism; SD, standard deviation; RNFL, retinal nerve fiber layer.

**Table S7.7  Associations or OGTT-based indices of daily glucose variability and CGM-based measures with RNFL thickness.**

| | Number of participants | RNFL thickness, per SD | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | Model 1 stβ (95% CI) | Model 2 stβ (95% CI) | Model 3 stβ (95% CI) | Model 3 + HbA1c stβ (95% CI) | Model 3 + SD stβ (95% CI) |
| **Other OGTT-based measures** | | | | | | |
| Maximum glucose peak, per SD | 2,407 | **-0.09 (-0.13; -0.05)** | **-0.07 (-0.11; -0.03)** | **-0.05 (-0.10; -0.01)** | -0.06 (-0.12; 0.004) | NA |
| 1-hour post-load glucose, per SD | 2,531 | **-0.07 (-0.11; -0.03)** | **-0.05 (-0.10; -0.01)** | -0.04 (-0.08; 0.01) | -0.03 (-0.09; 0.03) | NA |
| **CGM-based measures** | | | | | | |
| Mean sensor glucose, per SD | 622 | **-0.10 (-0.18; -0.01)** | -0.08 (-0.17; 0.004) | -0.07 (-0.17, 0.03) | NA | -0.02 (-0.16; 0.13) |
| Coefficient of variation, per SD | 622 | -0.07 (-0.15; 0.01) | -0.07 (-0.15; 0.02) | -0.06 (-0.15; 0.02) | NA | NA |
| Time in range, per SD | 622 | 0.06 (-0.03; 0.14) | 0.05 (-0.04; 0.13) | 0.03 (-0.06; 0.12) | NA | NA |

Results (β [95% confidence interval]) represent the difference in retinal nerve fiber layer thickness (in SD) for one SD greater exposure to a determinant. For incremental glucose peak and maximum glucose peak, 1 SD corresponds with 11.1 μm for RNFL thickness, 2.9 mmol/L for incremental glucose peak and 3.9 mmol/L for maximum glucose peak. For 1-hour post-load glucose, 1 SD corresponds with 11.0 μm for RNFL thickness and 4.0 mmol/L for 1-hour post-load glucose. For the CGM-based measures, 1 SD corresponds with 10.8 μm for RNFL thickness, 1.3 mmol/L for mean sensor glucose, 5.6% for coefficient of variation, and 12.9% for time in range. Model 1 was not adjusted for potential confounders (crude); model 2 was adjusted for age, sex, and education level [low, middle, high]; model 3 was additionally adjusted for waist circumference, office systolic blood pressure, antihypertensive medication, total cholesterol to HDL cholesterol ratio, use of lipid-modifying medication, smoking status [current, ever, never], and alcohol consumption status [none, low, high]; model 3 + HbA1c was additionally adjusted for HbA1c (in case of maximum glucose peak, 1-hour post-load glucose). Model 3 + SD was additionally adjusted for CGM-assessed standard deviation (only applicable for mean sensor glucose). Additionally, and only for CGM-assessed indices, we entered 'visit interval' in model 1. AU, arbitrary unit; CGM, continues glucose monitoring; CI, confidence interval; GMS, glucose metabolism status; HbA1c, hemoglobin A1c; NGM, normal glucose metabolism; SD, standard deviation; RNFL, retinal nerve fiber layer.

7

**Table S7.8** Associations of glucose metabolism status, measures of glycemia, and indices of daily glucose variability with RNFL thickness, where waist circumference was replaced with BMI (model 3A) or educational level was replaced with occupational status (model 3B) or income level (model 3C).

| | RNFL thickness, per SD | | | | | |
|---|---|---|---|---|---|---|
| | Model 3A | | Model 3B | | Model 3C | |
| | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) |
| Glucose metabolism status | | | | | | |
| Prediabetes versus NGM | 5,456 | -0.05 (-0.13; 0.03) | 1,847 | -0.10 (-0.23; 0.04) | 4,189 | -0.06 (-0.15; 0.04) |
| Type 2 diabetes versus NGM | 5,456 | **-0.17 (-0.25; -0.09)** | 1,847 | **-0.20 (-0.34; -0.06)** | 4,189 | **-0.16 (-0.26; -0.07)** |
| Measures of glycemia | | | | | | |
| Fasting plasma glucose, per SD | 5,455 | **-0.05 (-0.08; -0.02)** | 1,847 | -0.04 (-0.09; 0.01) | 4,188 | **-0.05 (-0.08; -0.01)** |
| 2-hour post-load glucose., per SD | 5,181 | **-0.06 (-0.09; -0.03)** | 1,727 | **-0.08 (-0.14; -0.03)** | 3,998 | **-0.05 (-0.09; -0.01)** |
| HbA1c, per SD | 5,450 | **-0.05 (-0.09; -0.02)** | 1,843 | -0.04 (-0.09; 0.02) | 4,184 | **-0.05 (-0.08; -0.01)** |
| Skin autofluorescence, per SD | 5,133 | **-0.04 (-0.07; -0.01)** | 1,781 | **-0.06 (-0.11; -0.01)** | 3,942 | **-0.05 (-0.08; -0.01)** |
| Indices of daily glucose variability | | | | | | |
| Incremental glucose peak, per SD | 2,407 | **-0.06 (-0.11; -0.02)** | 1,541 | -0.06 (-0.11; 0.003) | 1,860 | -0.05 (-0.10; 0.004) |
| Model 3 + HbA1c | 2,407 | **-0.06 (-0.12; -0.01)** | 1,541 | -0.06 (-0.13; 0.01) | 1,860 | -0.04 (-0.11; 0.02) |
| CGM- assessed standard deviation, per SD | 622 | -0.08 (-0.17; 0.01) | 0 | N/A | 483 | **-0.13 (-0.13; -0.02)** |
| Model 3 + MSG | 622 | -0.07 (-0.21; 0.07) | 0 | N/A | 483 | -0.11 (-0.27; 0.05) |

Standardized regression coefficient (stβ) represents the difference in RNFL thickness in SD for individuals with type 2 diabetes or prediabetes versus individuals with NGM or per SD greater measure of glycemia or daily glucose variability. In the GMS, fasting plasma glucose, 2-hour post-load glucose, HbA1c, and skin autofluorescence study populations, 1 SD corresponds with 10.8 μm for RNFL thickness, 1.5 mmol/L for fasting plasma glucose, 4.0 mmol/L for 2-hour post-load glucose, 0.8% or 9.1 mmol/mol for HbA1c, and 0.5 AU for skin autofluorescence (model 3A). For incremental glucose peak, 1 SD corresponds with 11.1 μm for RNFL thickness and 2.9 mmol/L for incremental glucose peak (model 3A). For CGM-assessed standard deviation, 1 SD corresponds with 10.8 μm for RNFL thickness and 0.58 mmol/L for CGM-assessed standard deviation (model 3A). In models 3B and 3C values per SD were numerically comparable. Variables in model 3: age, sex, educational status (where applicable), office systolic blood pressure, total cholesterol/HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication, waist circumference (where applicable), smoking status, and alcohol consumption status. Incremental glucose peak was additionally adjusted for HbA1c (model 3+ HbA1c). CGM-assessed standard deviation was additionally adjusted for MSG. Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1. AU, arbitrary unit; BMI, body mass index; CGM, continues glucose monitoring; CI, confidence interval; GMS, glucose metabolism status; HbA1c, hemoglobin A1c; MSG, mean sensor glucose; NGM, normal glucose metabolism; SD, standard deviation; RNFL, retinal nerve fiber layer.

**Table S7.9** Associations of glucose metabolism status, measures of glycemia, and indices of daily glucose variability with RNFL thickness where office systolic blood pressure was replaced with office diastolic blood pressure (model 3A), 24-hour ambulatory systolic blood pressure (model 3B), or 24-hour ambulatory diastolic blood pressure (model 3C).

| | RNFL thickness, per SD | | | | | |
| | Model 3A | | Model 3B | | Model 3C | |
| | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) | Number of participants | RNFL, per SD stβ (95% CI) |
| --- | --- | --- | --- | --- | --- | --- |
| **Glucose metabolism status** | | | | | | |
| Prediabetes versus NGM | 5,454 | -0.05 (-0.13; 0.03) | 4,921 | -0.04 (-0.13; 0.04) | 4,921 | -0.04 (-0.13; 0.04) |
| Type 2 diabetes versus NGM | 5,454 | **-0.17 (-0.25; -0.09)** | 4,921 | **-0.16 (-0.25; -0.08)** | 4,921 | **-0.17 (-0.25; -0.08)** |
| **Measures of glycemia** | | | | | | |
| Fasting plasma glucose, per SD | 5,453 | **-0.05 (-0.08; -0.02)** | 4,920 | **-0.05 (-0.08; -0.02)** | 4,920 | **-0.05 (-0.08; -0.02)** |
| 2-hour post-load glucose , per SD | 5,179 | **-0.06 (-0.09; -0.02)** | 4,683 | **-0.05 (-0.09; -0.02)** | 4,683 | **-0.05 (-0.09; -0.02)** |
| HbA1c, per SD | 5,448 | **-0.05 (-0.08; -0.02)** | 4,915 | **-0.05 (-0.08; -0.02)** | 4,915 | **-0.05 (-0.08; -0.02)** |
| Skin autofluorescence, per SD | 5,131 | **-0.04 (-0.08; -0.01)** | 4,627 | **-0.04 (-0.07; -0.01)** | 4,627 | **-0.04 (-0.07; -0.01)** |
| **Indices of daily glucose variability** | | | | | | |
| Incremental glucose peak, per SD | 2,407 | **-0.06 (-0.10; -0.01)** | 2,134 | -0.05 (-0.10; 0.002) | 2,134 | -0.05 (-0.10; 0.001) |
| Model 3 + HbA1c | 2,407 | **-0.06 (-0.12; -0.01)** | 2,134 | -0.06 (-0.12; 0.003) | 2,134 | -0.06 (-0.12; 0.001) |
| CGM- assessed standard deviation, per SD | 622 | -0.08 (-0.17; 0.01) | 562 | -0.08 (-0.17; 0.02) | 562 | -0.07 (-0.17; 0.02) |
| Model 3 + MSG | 622 | -0.07 (-0.21; 0.07) | 562 | -0.08 (-0.22; 0.06) | 562 | -0.08 (-0.22; 0.06) |

Standardized regression coefficient (stβ) represents the difference in RNFL thickness in SD for individuals with type 2 diabetes or prediabetes versus individuals with NGM or per SD greater measure of glycemia or daily glucose variability. In the GMS, fasting plasma glucose, 2-hour post-load glucose, HbA1c, and skin autofluorescence study populations, 1 SD corresponds with 10.8 μm for RNFL thickness, 1.5 mmol/L for fasting plasma glucose, 4.0 mmol/L for 2-hour post-load glucose, 0.8% or 9.1 mmol/mol for HbA1c, and 0.5 AU for skin autofluorescence (model 3A). For incremental glucose peak, 1 SD corresponds with 11.1 μm for RNFL thickness and 2.9 mmol/L for incremental glucose peak (model 3A). For CGM-assessed standard deviation, 1 SD corresponds with 10.8 μm for RNFL thickness and 0.58 mmol/L for CGM-assessed standard deviation (model 3A). In models 3B and 3C values per SD were numerically comparable. Variables in model 3: age, sex, educational status, office diastolic blood pressure (where applicable), total cholesterol/HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication, waist circumference (where applicable), smoking status, and alcohol consumption status. Incremental glucose peak was additionally adjusted for HbA1c (model 3+ HbA1c). CGM-assessed standard deviation was additionally adjusted for MSG. Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1. CGM, continues glucose monitoring; CI, confidence interval; GMS, glucose metabolism status; HbA1c, hemoglobin A1c; MSG, mean sensor glucose; NGM, normal glucose metabolism; SD, standard deviation; RNFL, retinal nerve fiber layer.

**Table S7.10** Standardized regression coefficients of skin autofluorescence with retinal nerve fiber thickness after additional adjustment for fasting plasma glucose, 2-hour post-load glucose, or HbA1c.

| | RNFL thickness, per SD | | | | |
|---|---|---|---|---|---|
| | Model 3 + fasting plasma glucose | | Model 3 + 2-hour post-load glucose | | Model 3 + HbA1c | |
| | Number of participants | stβ (95% CI) | Number of participants | stβ (95% CI) | Number of participants | stβ (95% CI) |
| SAF, per SD | 5,131 | -0.04 (-0.07; -0.01) | 4,872 | -0.04 (-0.07; -0.01) | 5,126 | -0.04 (-0.07; -0.01) |

Results (β [95% confidence interval]) represent the difference in retinal nerve fiber layer thickness (in SD) for one SD greater exposure to a determinant. One SD corresponds with 10.9 μm for the RNFL and 0.5 AU for SAF in model 3+fasting plasma glucose. In models 3B and 3C values per SD were numerically comparable. Bold denotes p<0.05. The associations were adjusted for age, sex, education level [low, middle, high], waist circumference, office systolic blood pressure, antihypertensive medication, total cholesterol to HDL cholesterol ratio, use of lipid-modifying medication, smoking status [current, ever, never], alcohol consumption status [none, low, high], and fasting plasma glucose (model 3 + fasting plasma glucose), 2-hour post load glucose (model 3 + 2-hour post load glucose), or HbA1c (model 3 + HbA1c). Abbreviations: SAF, skin autofluorescence; CI, confidence interval; HbA1c, hemoglobin A1c; SD, standard deviation; RNFL, retinal nerve fiber layer; SAF, skin autofluorescence.

**Table S7.11** Standardized regression coefficients of incremental glucose peak and CGM-assessed standard deviation with retinal nerve fiber thickness after replacement of HbA1c with fasting plasma glucose or skin autofluorescence, or mean sensor glucose with fasting plasma glucose, skin autofluorescence, or HbA1c.

| | Retinal nerve fiber layer thickness, per SD | | | | |
|---|---|---|---|---|---|
| | Model 3 + fasting plasma glucose | | Model 3 + SAF | | Model 3 + HbA1c | |
| | Number of participants | stβ (95% CI) | Number of participants | stβ (95% CI) | Number of participants | stβ (95% CI) |
| Incremental glucose peak, per SD | 2,407 | -0.07 (-0.13; -0.01) | 2,295 | -0.05 (-0.10; -0.01) | N/A | N/A |
| CGM-assessed standard deviation, per SD | 622 | -0.06 (-0.17; 0.04) | 573 | -0.07 (-0.17; 0.03) | 622 | -0.05 (-0.16; 0.07) |

Results (β [95% confidence interval]) represent the difference in retinal nerve fiber layer thickness (in SD) for 1 SD greater exposure to a determinant. For the incremental glucose peak results, 1 SD corresponds with 11.1 μm for the RNFL and 2.9 mmol/L for incremental glucose peak. For the CGM-assessed standard deviation results, 1 SD corresponds with 10.7 μm for the RNFL and 0.58 mmol/L for CGM-assessed standard deviation. The associations were adjusted for age, sex, education level [low, middle, high], waist circumference, office systolic blood pressure, antihypertensive medication, total cholesterol to HDL cholesterol ratio, use of lipid-modifying medication, smoking status [current, ever, never], alcohol consumption status [none, low, high], and fasting plasma glucose (model 3 + fasting plasma glucose), skin autofluorescence (model 3 + SAF), or HbA1c (model 3 + HbA1c). Additionally, and only for CGM-assessed SD, we entered 'visit interval' in model 1. Abbreviations: CGM, continues glucose monitoring; CI, confidence interval; HbA1c, hemoglobin A1c; SD, standard deviation; RNFL, retinal nerve fiber layer; SAF, skin autofluorescence.

**Table S7.12** Associations of continuous glucose monitoring-assessed standard deviation with retinal nerve fiber thickness after exclusion of individuals with less than 3 days of CGM data available (model 3A), individuals with CGM data gaps (model 3B), or with a 'visit interval' (Model 3C).

| | RNFL thickness, per SD | | | | | |
|---|---|---|---|---|---|---|
| | Model 3A | | Model 3B | | Model 3C* | |
| | Number of participants | stβ (95% CI) | Number of participants | stβ (95% CI) | Number of participants | stβ (95% CI) |
| CGM-assessed standard deviation, per SD | 596 | -0.09 (-0.18; 0.01) | 573 | **-0.10 (-0.19; -0.001)** | 409 | 0.01 (-0.10; 0.12) |
| Model 3 + MSG | 596 | -0.08 (-0.23; 0.06) | 573 | -0.08 (-0.23; 0.07) | 409 | 0.02 (-0.12; 0.17) |

Results (β [95% confidence interval]) represent the difference in retinal nerve fiber layer thickness (in SD) for 1 SD greater exposure to a determinant. For the results, 1 SD corresponds with 10.7μm for the RNFL and 0.58 mmol/L for CGM-assessed standard deviation (model 3A). In models 3B and 3C values per SD were numerically comparable. Bold denotes p<0.05. The associations were adjusted for visit interval, age, sex, education level [low, middle, high], waist circumference, office systolic blood pressure, antihypertensive medication, total cholesterol to HDL cholesterol ratio, use of lipid-modifying medication, smoking status [current, ever, never], alcohol consumption status [none, low, high], and mean sensor glucose. * Of note, model 3C was not adjusted for visit interval, as all individuals had an interval of 0 years. CGM, continues glucose monitoring; CI, confidence interval; GMS, glucose metabolism status; MSG, mean sensor glucose; SD, standard deviation; RNFL, retinal nerve fiber layer.

**Table S7.13** Associations of CGM-assessed standard deviation and mean sensor glucose with retinal nerve fiber thickness estimated with ridge regression and presented for different degrees of penalization.

| VIF | Number of participants | RNFL thickness, per SD | |
|---|---|---|---|
| | | CGM-assessed standard deviation, per SD (st.β, 95% CI) | CGM-assessed mean sensor glucose, per SD (st.β, 95% CI) |
| Model 3 + MSG (λ=0) | 622 | -0.063 (-0.183; 0.058) | -0.021 (-0.158; 0.121) |
| Halfway (λ=0.04) | 622 | -0.062 (-0.183; 0.056) | -0.023 (-0.156; 0.114) |
| Model 3 (λ=0.10) | 622 | -0.063 (-0.186; 0.053) | -0.020 (-0.152; 0.103) |

Standardized regression coefficients (st.β) indicate the median difference (95% confidence interval) associated with 1 SD higher SD or MSG. All coefficients were estimated by use of ridge regression. We pragmatically chose the level of penalization based on the λ required to reduce the variance inflation factor (VIF) of model 3 + MSG back to the VIF of model 3 (or halfway back). Point estimates and 95% confidence intervals were calculated by use of 1,000 bootstraps estimates. The associations were adjusted for visit interval, age, sex, education level [low, middle, high], waist circumference, office systolic blood pressure, antihypertensive medication, total cholesterol to HDL cholesterol ratio, use of lipid-modifying medication, smoking status [current, ever, never], alcohol consumption status [none, low, high], and mean sensor glucose. Abbreviations: CGM, continues glucose monitoring; CI, confidence interval; GMS, glucose metabolism status; MSG, mean sensor glucose; SD, standard deviation; RNFL, retinal nerve fiber layer.

7

**Table S7.14  Associations of duration of diabetes and age with retinal nerve fiber layer thickness.**

| | Number of participants | RNFL thickness, per SD | | |
| --- | --- | --- | --- | --- |
| | | Model 1<br>stβ (95% CI) | Model 2<br>stβ (95% CI) | Model 3<br>stβ (95% CI) |
| Duration of diabetes, per SD | 982 | **-0.07 (-0.13; -0.004)** | -0.06 (-0.13; 0.001) | **-0.07 (-0.13; -0.001)** |
| Age, per SD | 5,180 | **-0.06 (-0.08; -0.03)** | **-0.11 (-0.13; -0.08)** | **-0.09 (-0.12; -0.06)** |

Standardized regression coefficient (stβ) represents the difference in RNFL thickness in SD per SD longer duration of diabetes or age. One SD corresponds with 7.4 years for duration of diabetes or 8.7 years for age and  11.2 μm or 10.9 μm for RNFL thickness in respectively the duration of diabetes study population and the age study population. Bold denotes p<0.05. Variables entered in the models in addition to duration of diabetes or age: model 1: none (crude results); model 2: age (where applicable), sex, and educational status (low, medium, high); model 3: model 2 + office systolic blood pressure, total cholesterol to HDL cholesterol ratio, use of antihypertensive or lipid-modifying medication (yes/no), waist circumference, smoking status (current, ever, never), and alcohol consumption status (none, low, high). Additionally, and only for analyses with age as determinant, glucose metabolism status and spherical equivalent were entered in model 2. We compared the standardized betas of type 2 diabetes versus NGM  (i.e. -0.16SD) with the standardized beta for age to calculate that beta of type 2 diabetes [stβ, -0.16] corresponds with approximately 15 years of aging (calculated as stβtype 2 diabetes/ stβage * number of years per SD = [0.16/0.09]* 8.7 years = 15 years). CI, confidence interval; HDL, high-density lipoprotein; SD, standard deviation; RNFL, retinal nerve fiber layer; NGM, normal glucose metabolism; FPG, fasting plasma glucose; HbA1c, hemoglobin A1c.

# Chapter 8

## Higher daily glucose variability is associated with worse cognitive performance – The Maastricht Study

D. Foreman, Rutger J. Nijland, Martijn C.G.J. Brouwers, Sebastian Köhler, Martin P.J. van Boxtel, Marleen M.J. van Greevenbroek, Carla J.H. van der Kallen, Simone J.P.M. Eussen, Annemarie Koster, Anke Wesselius, Abraham A. Kroon, Pieter C. Dagnelie, Bastiaan E. de Galan, Nicolaas C. Schaper, Miranda T. Schram, and Coen D.A. Stehouwer

EMBARGOED

# Part III

Conclusions, perspectives and impact

# Chapter 9

Summary and general discussion

# A meaningful addition to the mean?

Diabetes mellitus is an increasingly prevalent global health concern that currently affects approximately 463 million individuals[1,2]. It is defined by elevated blood glucose concentrations (i.e., hyperglycemia)[3] and encompasses a number of distinct metabolic diseases, the most common of which are type 1 (5-10%) and type 2 diabetes (90-95%)[4]. Diabetes not only engenders a twofold higher risk of all-cause mortality[5-7], but also predisposes to many chronic morbidities. These include, amongst others, cardiovascular disease (CVD)[8], the classic microvascular complications (i.e., retinopathy, neuropathy, and nephropathy)[9,10], and dementia[11]. Notably, increasing evidence indicates that individuals with moderately elevated blood glucose concentration (i.e., prediabetes) already have a higher risk of such complications[10-13].

Diabetes and its co-morbidities have debilitating effects at an individual level[14] and also have a negative societal impact due to their large economic burden[15,16]. Diabetes treatment, which generally involves both lifestyle interventions and medication, is primarily aimed at the prevention of the development or progression of complications[17]. Lowering of glucose concentrations (i.e., improving glycemic control) is one of the main ways to achieve this[3]. For decades, $HbA_{1c}$ has been the gold-standard measure to quantify glycemic control[3]. However, this index of average glycaemia fails to provide sufficient information on daily glucose fluctuations (i.e., daily glucose variability; GV)[18].

Daily GV may be a modifiable risk factor for complications, regardless of mean glucose concentrations[19-25], and, hence, measuring it in addition to the mean glycemia may prove meaningful. However, since large-scale, accurate measurement of daily GV has been challenging up until recently, there is a paucity of large epidemiological studies on the potential consequences of daily glucose fluctuations[26]. This thesis, therefore, aims to extend the current literature by investigating these potential consequences and, in addition, strives to address specific methodological aspects of continuous glucose monitoring (CGM) and GV measurement in order to improve current clinical practice and future research. This chapter provides a summary of the main findings and discusses them in the context of existing scientific literature. In addition, the present chapter reflects on the methodology of the performed studies and suggests future directions for research on daily GV.

## *Key findings and their implications*

In this section, the key findings of this thesis as well as their scientific and clinical implications will be comprehensively discussed. First, I will review the foremost findings on the methodology of daily GV measurement and CGM (Part I). Next, I will discuss the main results on potential consequences of daily GV (Part II).

9

## Part I: Methodological aspects of daily glucose variability measurement and continuous glucose monitoring

While frequent venous blood glucose sampling (e.g., by venipuncture or fingerprick) was initially the only feasible way to assess daily GV[27,28], CGM has become the gold standard method to do so more than ten years ago[29,30]. CGM is increasingly being used in diabetes care[31], mostly by individuals with type 1 diabetes, either solitarily or incorporated in advanced closed-loop insulin delivery systems[32,33]. A considerable number of glycemic indices can be derived from frequently measured glucose values[18,34]. In this thesis, the main CGM-assessed indices of interest were:

- Mean sensor glucose (MSG), which is calculated by averaging all glucose measurements. Presently, MSG is one of the key CGM-derived metrics[29,30].
- Standard deviation (SD), which reflects the amount of variation in all glucose measurements. SD is one of the two currently recommended daily GV indices because it is easily interpretable, is widely used, has small random sampling error relative to other commonly used indices, and can be easily translated to a MSG-adjusted index (see below)[26,29,30]. It is justified to primarily use SD, since it strongly correlates with the other commonly used indices (e.g., interquartile range [IQR], mean amplitude of glycemic excursion [MAGE])[26].
- Coefficient of variation (CV), which is calculated as SD / MSG * 100%. Currently, CV is the main recommended index for the quantification of daily GV[29]. It has the advantage of being intrinsically MSG-adjusted, which is important because of the strong biological and mathematical correlation between MSG and SD[26]. In addition, this relativeness to MSG is thought to make CV more informative than SD with regard to hypoglycemia risk[29]. The current international consensus report on clinical CGM-based targets recommends a CV of ≤36%[30].
- Time in range (TIR), which reflects the percentage of glucose values between 3.9 and 10 mmol/L. TIR, which is partly determined by average glycemia and daily GV[35], is an emerging glycemic index[30,36].The specific target range of 3.9 to 10 mmol/L was decided by international consensus and applies to individuals with type 1 diabetes or type 2 diabetes; a TIR of ≥70% is the recommended clinical target[30].

In this thesis, the following oral glucose tolerance test-derived index was additionally used as a proxy of daily GV:

- Incremental glucose peak (IGP), which is calculated by subtracting the fasting plasma glucose (FPG) value from the highest glucose value during a complete seven-point OGTT. In this thesis, we establish that IGP is a valid OGTT-derived index of daily GV. Compared to CGM-derived indices of GV, IGP has the major advantage of being available in a substantially larger sample of The Maastricht Study population.

*Recording period reliability for continuous glucose monitoring-derived indices*

While the number of CGM recording days required for a reliable measurement of daily GV indices had been established for individuals with type 1 diabetes (i.e., 12-15 days)[37-39], it was insufficiently clear for individuals with other glucose metabolism status (GMS) categories. In *Chapter 2*, we showed that relatively few days are needed to reach sufficient reliability (R>0.8) for MSG quantification: 1 day for individuals with prediabetes or type 2 diabetes, and 3 days for individuals with normal glucose metabolism (NGM). The number of days required are slightly higher for SD (i.e., 2 days in case of prediabetes and type 2 diabetes; and 3 days in case of NGM) and CV (i.e., 3 days in case of NGM, prediabetes, and type 2 diabetes). The current international consensus report on clinical CGM-based targets recommends that the device is worn at least 14 days (with at least 70% of usable data)[30], an advice that is established on data in type 1 diabetes only[37-39]. Based on our findings, we propose reevaluation of this clinical recommendation with the intend of establishing a different target for individuals with type 2 diabetes. Six days of functional CGM would be appropriate to measure MSG, SD, and CV highly reliably (R>0.9) in these individuals. Additionally, from the specific viewpoint of scientific studies, it would be valid to further shorten the CGM period for individuals with NGM, prediabetes or type 2 diabetes to three functional days in order to alleviate participant burden without compromising reliability. Notably, however, we observed that reliable measurement of MAGE in individuals with type 2 diabetes required at least 6 days of CGM, which is in agreement with findings on the relatively large sampling error of this index[40]. This is an important observation because a substantial number of studies have calculated MAGE based on 72h of CGM or less (e.g., [41,42]), which may have caused biased results[43,44]. In addition, the reliability analyses were not performed for the increasingly important TIR, which might, like MAGE, require a longer recording period than three days.

*Machine learning-based glucose prediction as a method to improve continuous glucose monitoring and closed-loop insulin delivery systems*

While closed-loop insulin delivery systems are steadily being improved upon[45], certain obstacles to the optimal performance of these devices remain. With regard to CGM, these include ~10-minute sensor delay, which, in part, results from measuring interstitial glucose rather than blood glucose concentrations, and short periods of sensor malfunction[45-47]. In *Chapter 3*, we used 1.4 million glucose measurements of participants with NGM, prediabetes or type 2 diabetes (n=853) to show that machine learning-based glucose prediction can be used to address both. Namely, our models, which are based on a long-short term memory network architecture and only require 30 minutes of prior CGM data, are capable of accurately predicting glucose values at 15- and 60-minute intervals. Notably, in individuals with type 2 diabetes, more than 98% of the predicted glucose values were highly clinically safe, as assessed with

surveillance error grids. Our prediction model outperformed the model based on the largest study to date in individuals with type 2 diabetes(n=50)[48], which we expect to partly stem from the larger sample, which tends to yield better and more reliable prediction[49].

Since individuals with type 1 diabetes are, at present, the primary target population for closed-loop insulin delivery systems[50] and generally experience greater, less predictable daily glucose fluctuations[51], we explored whether the glucose prediction model that was trained on participants with NGM, prediabetes or type 2 diabetes would generalize to individuals with type 1 diabetes. For this, we employed the publically available OhioT1DM Dataset[52], which is commonly utilized for model performance evaluation (e.g.,[53,54]). Exploratory translation to individuals with type 1 diabetes showed that our glucose prediction models performed well in this group, albeit not to the extent that was observed in the original evaluation set, which underscores the importance of assessing prediction model generalizability. Nevertheless, the models had sufficient clinical safety and performed equally well as the most recent studies in individuals with type 1 diabetes[53-57]. We anticipate that further technical optimization of our models in combination with a larger population of individuals with type 1 diabetes will advance the accuracy and safety even more.

We also studied whether glucose prediction can be improved by incorporation of accelerometry data, since physical activity is known to impact glucose concentrations[58] and because it would be feasible to combine an accelerometer with the other devices of a closed-loop insulin delivery system. Somewhat unexpectedly, the additional information provided by accelerometry only slightly improved our CGM-based prediction models. This may be due to the already excellent performance of the CGM-based prediction. Namely, the better a prediction model, the more difficult it becomes to substantially improve it[59]. As such, the conclusion would be to not use accelerometry in order to improve closed-loop insulin delivery systems. However, it should be noted that the combined approach yielded larger improvements at the 60-minute interval, most notably during the day and in individuals with greater daily GV. Hence, accelerometer-derived data may be more important in case of longer prediction intervals or in individuals with more pronounced, less predictable daily glucose fluctuations[51]. This should be further explored in individuals with type 1 diabetes.

*Reference values of continuous glucose monitoring-derived indices*

Reference values for CGM-derived indices have, thus far, been sparsely investigated in individuals without diabetes[60, 61], especially in a setting without dietary restrictions[62]. In *Chapter 2*, we established—under normal living conditions—reference values for MSG (5.03 - 6.69 mmol/L), SD (0.44 - 1.37 mmol/L), and CV (7.74 - 22.45%) based on the 2.5[th]-97.5[th] percentiles of individuals with NGM[63]. Our SD and MSG reference values are slightly higher than those previously reported for Chinese individuals with NGM[60,61],

which could stem from ethnic or cultural differences, or from the fact that these participants were given dietary instructions as part of the study design. The previously mentioned international consensus report on clinical CGM-based targets recommends a CV of ≤36%[30], which lies substantially higher than our upper reference value of 22.5%. This target was chosen because it represented the highest CV value in diet- or metformin-treated individuals with type 2 diabetes and had been linked with substantially less hypoglycemic events compared to a CV of >36%[64]. However, hypoglycemic events still occurred in the CV≤36% group, which questions the legitimacy of this specific cut-off point. Interestingly, in a study on the association between CV and percentage in hypoglycemia, Rodbard showed that the pertinent regression line crossed the zero percent in hypoglycemia mark, both for a 4.4 and 2.8 mmol/L hypoglycemia threshold, for a CV of ~22.5%[65]. This implies that aiming for a more physiological degree of daily GV (i.e., a CV of 22.5% or lower), although not easily achieved in insulin-treated diabetes, would fully prevent hypoglycemic events.

*Novel daily glucose variability indices*

New and increasingly advanced techniques for the assessment of CGM-based daily GV are proposed on a regular basis[66-68]. One of these methods, the spectral clustering-based 'glucotype', has garnered considerable attention—especially in the context of precision medicine—from multiple high-impact journals[69-72] after it was promoted by the authors as a more comprehensive GV measure than those in current practice[66]. In *Chapter 4*, we evaluated the validity and generalizability of the glucotype as an index of daily GV by using CGM data of participants from The Maastricht Study[73] and PRE-D Trial[74]. We showed that the distribution of the three glucotypes (i.e., low, moderate, or severe variability) observed in the original study[66] generalizes poorly to other study populations. Additionally, by calculating relative entropy[75], we demonstrated that cluster separation is far from optimal (i.e., a relative entropy of 0.24), which means that all participants spend a substantial amount of time in other glucotypes than the main one assigned to them. Most importantly, we established, using data from the above mentioned cohorts as well as the original study, that the classification of glucotypes is almost completely driven by MSG. As such, we dispute the authors' claim that the glucotype is a 'more comprehensive measure of the pattern of glucose excursions than the standard laboratory tests in current use'. These findings again underscore the importance of using large study samples to develop valid data-driven models and of utilizing other populations to assess their generalizability.

We hypothesized that indices that quantify the glucose excursion during an OGTT could be used to assess daily GV, in particular when CGM is unavailable. Conceptually, we expected that the aforementioned IGP (i.e., highest OGTT glucose value minus FPG) would best reflect SD, as both indices represent a glucose excursion relative to the mean. In *Chapter 2*, we showed that CGM-assessed SD and CV were most strongly correlated with the maximum glucose peak (rho=0.60 and 0.49), IGP (rho=0.59 and

9

0.50), and 1-hour OGTT value (rho=0.61 and 0.49). Because of the strength of the correlations (i.e., moderate-to-strong), we posit that these OGTT-based indices can all be used to approximate daily GV when CGM is unavailable, particularly in a large epidemiological setting. However, because the 1-hour OGTT value often does not correspond with the actual glucose peak time point—especially in individuals with NGM (only in 15%) and type 2 diabetes (only in 20%)—and since IGP outperformed both the 1-hour OGTT value and maximum glucose peak in GMS-stratified correlations, we conclude that IGP is indeed the preferred OGTT-derived index to use as proxy of daily GV.

The independent associations of IGP with aortic stiffness (*Chapter 5*), neurodegeneration (*Chapter 7*), and cognitive performance (*Chapter 8*) further underscore the relevance of this OGTT-derived index. In agreement with the observed correlations of IGP with SD and CV (*Chapter 2*), the associations of IGP with outcome measures were generally consistent with those of CGM-assessed indices of daily GV (Part II). As such, we believe IGP to be a valid proxy of daily GV in a research setting, especially in case of such a large epidemiological cohort as The Maastricht Study. Nevertheless, we underscore the importance of verifying the generalizability to other studies[76].

Notwithstanding the above, we do not expect that IGP will have a place in clinical practice for several reasons. First, it is questionable whether IGP will reflect daily GV well on an individual basis, given that the correlation coefficients were at most 0.6 (*Chapter 2*). Second, CGM measures real-life physiology, whereas OGTT-based assessment of daily GV can be seen as a non-physiological stress test, which will, for example, disregard strict diet adherence in daily life. Third, CGM can provide more detailed information (e.g., MSG, TIR) than a sole OGTT can, and given its increasing availability[31] is more preferable. Last, an OGTT is not commonly performed clinically and, more importantly, is contraindicated for individuals who use insulin[77], in whom assessment of daily GV is arguably of most interest.

## Part II: The consequences of daily glucose variability

Fifteen years ago, daily GV was put forward as a potentially modifiable, mean glycemia-independent risk factor for diabetes-associated complications[19]. Since then, a relatively small number of studies have investigated whether this is the case for common complications of diabetes, which include CVD, microvascular complications, and cognitive decline[22, 78]. Moreover, the majority of prior research on the consequences of daily GV is limited by specific methodological concerns, which include small study populations, inadequate adjustment for confounding, and suboptimally precise quantification of daily GV. These limitations are understandable when one considers that CGM-based measurement of daily GV has been challenging on a large scale up until recently, in part, due to costliness and relative invasiveness of CGM[26]. Logically, there is an even greater paucity of CGM-based prospective studies.

In part II of this thesis, we studied the cross-sectional associations of daily GV, assessed with CGM and OGTT, with multiple precursors of the aforementioned diseases (**Figure 9.1**) using data from a large, well-characterized cohort study. This approach has certain advantages, which are particularly related to the use of precursors. First, the statistical power is higher when using continuous (e.g., carotid-femoral pulse wave velocity; cf-PWV) rather than dichotomous (e.g., CVD diagnosis) variables[79]. Second, it can provide insights into the underlying biological processes that lead up to a specific adverse outcome. Third, the greater biological proximity of daily GV to such processes is expected to result in stronger associations and, thereby, greater statistical power. However, it also has disadvantages. First, clinical relevance is much clearer for the association with an adverse outcome (e.g., CVD) than with one of its precursors (e.g., cf-PWV). Second, the cross-sectional design of our studies hampers insight into causality[80].

*Daily glucose variability as determinant of cardiovascular disease*

In the current thesis, we used several measures that reflect specific processes leading up to CVD. Stiffening of the aorta and carotid artery (i.e., arteriosclerosis) was measured as cf-PWV and carotid distensibility coefficient (carDC), respectively[81-83]. Higher cf-PWV indicates greater aortic stiffness, which independently increases the risk of incident CVD, cardiovascular mortality, and all-cause mortality[84-86]. Lower carDC reflects greater carotid stiffness, which has been linked to incident stroke[87]. In addition, carotid intima-media thickness (cIMT) and ankle-brachial index (ABI) were measured. Higher cIMT and lower ABI represent more pronounced atherosclerosis and, hence, are associated with a higher risk of CVD[8,88,89]. Last, mean and pulsatile circumferential wall stress of the carotid artery were measured with echography[90]. Higher circumferential wall stress is thought to reflect maladaptive arterial remodeling[91].

9
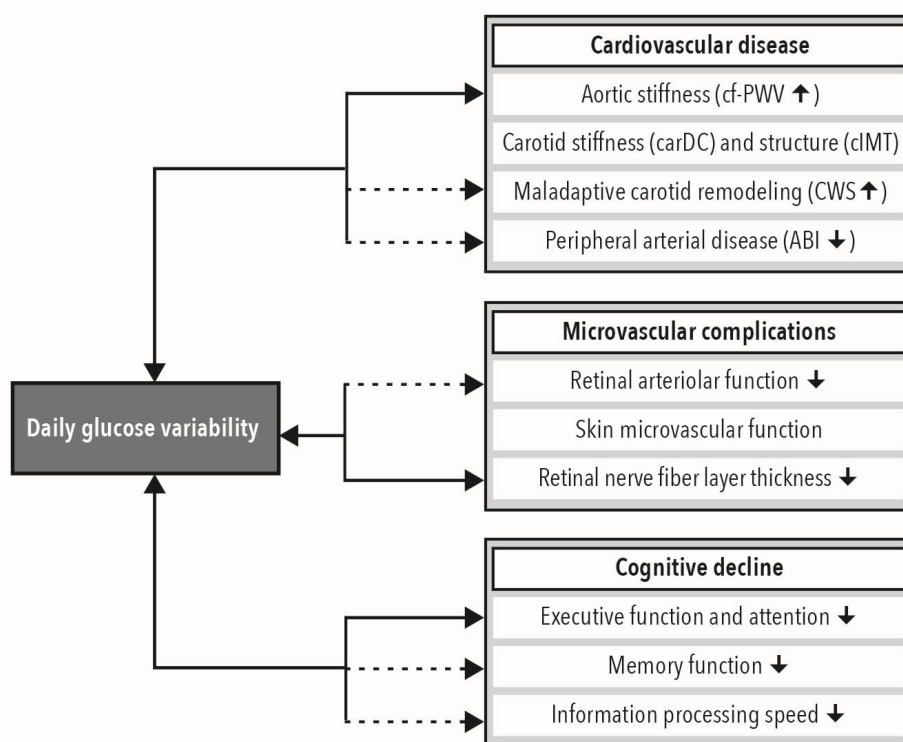
**Figure 9.1    Potential consequences of daily glucose variability.** Solid lines represent probable associations (i.e., statistically significant associations that were consistently observed for different indices of daily glucose variability). Dashed lines represent possible associations (i.e., non-consistent associations or findings that were indicative of an associations, but were not statistically significant). cf-PWV, carotid-femoral pulse wave velocity; carDC, carotid distensibility coefficient; cIMT, carotid intima-media thickness; CWS, circumferential wall stress; ABI, ankle-brachial index.

In *Chapter 6*, we showed that greater daily GV (i.e., SD and CV) and lower TIR (either as a continues variable or dichotomized [i.e., TIR≥70%]), all assessed with CGM, are linearly associated with greater aortic stiffness (i.e., cf-PWV), independent of demographics, cardiovascular risk factors, and lifestyle risk factors. Additional adjustment for MSG attenuated the association between SD and cf-PWV by a third and, thereby, undid statistical significance. This finding can be interpreted as evidence against an association between daily GV and aortic stiffness. Conversely, it can indicate that the sample size was too small to provide sufficient statistical power[92,93], which may be further undermined by the impact of multicollinearity on precision[94] (See *Overall methodological considerations*). Regardless of statistical significance, the associations of SD and MSG with cf-PWV were very similar, which may imply that both

have an equivalent pathophysiological relevance to aortic stiffness. The studies that have reported independent associations between measures of average glycemia (e.g., HbA$_{1c}$ and FPG) and cf-PWV[82] could, thereby, be viewed as support for daily GV as determinant of aortic stiffness. In agreement with this, we showed—with a threefold larger sample size—that higher IGP is associated with greater aortic stiffness, independent of all potential confounders including HbA$_{1c}$ (*Chapter 5*). The body of evidence that identifies cf-PWV as independent determinant of CVD, cardiovascular mortality, and all-cause mortality[84-86] underscores the relevance of our observations. Further, the difference in daily GV, either expressed as SD or IGP, between the highest and lowest population tertiles could both be translated to a substantial 2- to 3-year vascular aging difference[95].

In contrast with the findings on aortic stiffness, no independent associations were observed between OGTT-assessed (*Chapter 5*) or CGM-assessed (*Chapter 6*) daily GV and carotid stiffness (i.e., carDC). This difference could stem from structural differences between the aorta (i.e., mixed elastic and muscular) and carotid artery (i.e., predominately elastic)[84]. While no independent association was found between SD and cIMT, lower TIR was near statistically significantly associated with higher cIMT (*Chapter 6*). Both results contradict the findings in *Chapter 5*, where higher IGP was near statistically significantly associated with lower cIMT. Because of these contrasting results, and since higher HbA$_{1c}$ was associated with higher cIMT (*Chapter 5*), which is in agreement with prior studies on hyperglycemia[96], we presume the negative association between IGP and cIMT to be spurious. Further, CGM-based indices of daily GV were neither associated with mean or pulsatile circumferential wall stress (*Chapter 6*). By contrast, higher IGP was statistically significantly associated with higher mean circumferential wall stress (*Chapter 5*), which may indicate that higher daily GV is a determinant of maladaptive remodeling of the carotid artery, as has previously been reported for individuals with type 2 diabetes[91]. However, since cIMT is the denominator in the formula to calculate mean circumferential wall stress (i.e., lower cIMT mathematically results in higher circumferential wall stress)[90], this finding could also be spurious.

In *Chapter 6*, we observed that higher SD was associated with lower ankle-brachial index (ABI) values after adjustment for all potential confounders and MSG, albeit statistically non-significantly so (p=0.12). Although CV and TIR were not independently associated with ABI, a TIR≥70% was, even after additional adjustment for HbA$_{1c}$. This finding might indicate that daily GV is a determinant of atherosclerosis[88]. At present, the association between IGP and ABI has not yet been investigated. As sample size was observed to be crucial in case of cf-PWV and since ABI is a prognostic indicator of CVD[88], it would be of interest to do so in the future.

Several biological mechanisms may underlie the relationship between daily GV and CVD. Greater daily GV has been found to augment oxidative stress[97,98], which is known to promote the formation of advanced glycation endproducts (AGEs)[99]. AGEs have

been suggested to induce arterial stiffening by accumulating in the arterial wall and forming cross-links between elastin and collagen[81-83]. Thus far, an association of tissue and circulating AGEs has only been observed with aortic stiffness[100], which might provide an explanation for our contrasting findings regarding the structurally different aorta and carotid artery. Higher daily GV could also increase aortic stiffness by activating collagen-secreting fibroblasts[101] or by inducing large-artery endothelial dysfunction[83,102]. Last, not only higher glucose peaks, but also hypoglycemic nadirs, although they are expected to rarely occur in the study populations used, could contribute to the development of CVD[103,104].

Generally, our findings are in line with the few studies on the associations of CGM-based daily GV with arterial stiffness[105], cIMT[106], and ABI[107]. Since we observed a moderately strong correlation between CGM-assessed GV and the 1-hour OGTT value (*Chapter 2*), the studies that link the latter metric to arterial stiffness[108,109] and CVD[110] could be considered further support for daily GV as a determinant of CVD. However, whether lowering of daily GV is able to prevent CVD is still insufficiently clear. In this regard, two insulin treatment strategy arms with participants of similar HbA$_{1c}$ but different daily GV were compared in a post-hoc analysis of the HEART2D study. Notably, the arm with lower daily GV was not observed to have less CVD events[111]. However, the daily GV indices were measured with seven-point self-measured blood glucose (SMBG) profiles. Furthermore, only one of the three daily GV indices that had been calculated was statistically significantly lower, which questions whether the difference in daily GV was large enough to achieve lower CVD events.

*Daily glucose variability as determinant of microvascular complications*

In this thesis, we used several outcome measures that are pertinent to microvascular complications. In *Chapter 5*, we used retinal microvascular dilation in response to flicker light, as measured with the Dynamic Vessel Analyzer[112], and heat-induced skin hyperemic response, as assessed with a laser-Doppler system equipped with a thermostatic probe[113], as measures of retinal and skin microvascular function, respectively. Microvascular dysfunction is considered to precede microvascular complications[10]. In *Chapter 7*, we measured retinal nerve fiber layer (RNFL) thickness with optical coherence tomography (OCT)[114]. RNFL thinning reflects a gradual loss of retinal ganglion cell axons[115,116], and, as such, can be regarded as an index of neurodegeneration and precursor of retinopathy and neuropathy[114].

In *Chapter 7*, we found that more adverse GMS, greater glycemia (i.e., FPG, 2-hour post-load glucose, HbA1c, and skin autofluorescence), and greater daily GV (i.e., IGP and SD) are all linearly and—except for SD—statistically significantly associated with RNFL thickness. Notably, the strength of the associations for indices of daily GV was not materially different after additional adjustment for measures of average glycemia. The lack of statistical significance for the association of SD with RNFL thickness could result from insufficient statistical power[92], analogous to the associations with cf-PWV and

ABI (*Chapter 6*). This is supported by the observation that the statistically significant association between IGP and RNFL thickness, which was studied in a fourfold larger population, was numerically equivalent to the association of SD with RNFL thickness.

Contrarily, we did not observe independent associations between higher IGP and worse microvascular function in *Chapter 5*. Both the association with retinal arteriolar average dilatation and heat-induced skin hyperemia were strongly, but not completely, attenuated by adjustment for $HbA_{1c}$. This contrasts with the RNFL thickness findings, although comparison is hindered by a different regression model sequence. While the associations with microvascular function were adjusted for $HbA_{1c}$ directly after adjustment for demographics, the association with RNFL thickness was adjusted for $HbA_{1c}$ as a final step. Consequently, the cardiovascular risk and lifestyle factors may have already provided most of the information that $HbA_{1c}$ otherwise would have. It should also be noted that the sample sizes for retinal arteriolar average dilatation (n=1,591) and heat-induced skin hyperemia (n=1,134) were relatively small compared to the other IGP study populations (n=1,849 to 3,380). Hence, although a case against daily GV as independent determinant of microvascular function could certainly be made, it may also be argued—primarily for retinal arteriolar average dilatation—that there was insufficient statistical power to detect such an association[92,93]. Further, compared to RNFL thickness, retinal arteriolar average dilatation and heat-induced skin hyperemia are expected to be less precisely measured (See *Overall methodological considerations*), which is known to negatively impact association precision[44].

The main pathways through which hyperglycemia is thought to cause microvascular damage (i.e., AGE, polyol, hexosamine, and protein kinase C) have been extensively reviewed[99,117]. In case of retinal ganglion cells, these pathways can be directly neurotoxic, but may also be harmful indirectly through impairment of microvascular function[10,118]. Greater daily GV can entail both hyperglycemic peaks and hypoglycemic nadirs. Hyperglycemic peaks may be harmful on top of average glycemia as they have been found to augment oxidative stress[97,98], which is the unifying first step of the aforementioned pathways[99,117]. In case of retinal ganglion cells, hypoglycemic nadirs may be especially harmful because these cells are highly metabolically active and glucose is their key nutrient[118].

The findings on GMS, hyperglycemia and RNFL thickness (*Chapter 7*) are in line with current literature[119-125] and extend previous work of The Maastricht Study on the microvascular and neuronal branch of the "ticking clock hypothesis"[126-131], which postulates that hyperglycemia-induced damage is a continuous, gradual process that starts in prediabetes, progresses with the onset of type 2 diabetes, and continues during type 2 diabetes[13,132]. At present, no other studies have investigated the associations of daily GV with RNFL thickness or measures of microvascular function. Although numerous studies have reported that daily GV was associated with retinopathy[133,134], neuropathy[133,135,136], and nephropathy[133,137,138], their relevance is limited by small sample sizes, inadequate adjustment for confounders, or both. The findings of

9

these studies are also contradictory to several post-hoc analyses of the Diabetes Control and Complications Trial (DCCT) that were not able to independently associate daily GV, which had been suboptimally assessed with seven-point SMBG profiles, to microvascular complications in individuals with type 1 diabetes[28,139,140].

*Daily glucose variability as determinant of cognitive decline*

In *Chapter 8*, we used a concise neuropsychological test battery to assess performance on three individual cognitive domains (i.e., executive function and attention, memory function, and information processing speed)[73]. The individual domain scores were constructed in such a fashion that lower scores represent worse cognitive performance. We combined the scores on the three domains to reflect overall cognitive performance.

In *Chapter 8*, we showed that higher IGP was linearly associated with worse overall cognitive performance, independent of demographics, cardiovascular risk factors, lifestyle factors and average glycemia (i.e., $HbA_{1c}$ or FPG). For reference, the strength of the association translated to ~5 months of cognitive aging for every 1 mmol/L higher IGP. Worse performance on executive function and attention primarily explained the association between IGP and overall cognitive performance, which is in line with previous findings on diabetes and other glycemic measures[11]. While higher IGP was also associated with worse performance on memory function and information processing speed, these associations did not retain statistical significance after complete adjustment for potential confounders. Although the results for CGM-assessed daily GV are generally in agreement with the findings for IGP (i.e., the strongest associations were found with executive function and attention), the majority of the associations were not statistically significant. Of these, only TIR was independently associated with executive function and attention. Given the fourfold larger IGP sample size, this may again indicate concerns regarding statistical power[92] and effects of multicollinearity (See *Overall methodological considerations*)[94].

Our results are concordant with the few studies that have investigated the association between daily GV and cognitive outcomes[41,42,68,141]. Of these, the study that observed an independent association between greater GV[142], as assessed with 1,5-anhydroglucitol, and incident dementia in participants with diabetes and a $HbA_{1c}$ ≥53 mmol/mol only[141] specifically required corroboration because 1,5-anhydroglucitol has limited validity as GV proxy at elevated glucose concentrations[25,142].

Daily GV may impair cognitive performance in several ways. As cerebral glucose uptake is not insulin-dependent, hyperglycemic peaks will lead to elevated intraneuronal glucose concentrations[99], which can activate multiple neurotoxic pathways (e.g., polyol pathway, methylglyoxal formation, and protein glycation)[143-146]. A role for direct neurotoxicity is supported by the observed association between daily GV and RNFL thickness (*Chapter 7*), especially in the light of studies that have linked this retinal index of neurodegeneration to cognitive decline and incident dementia[147,148].

Daily GV could also indirectly affect cognitive performance through its effects on endothelial function[102,149] or aortic stiffness (*Chapters 5 and 6*), which have both been associated with cognitive performance[150,151]. Last, repeated periods of hypoglycemia can also provoke cognitive decline[152].

## Overall methodological considerations

In this section, various methodological aspects of *Chapters 2 through 8* will be comprehensively discussed. First, the three main threats to internal validity will be reviewed. Subsequently, causal inference, multiple testing adjustment, statistical power, and external validity will be addressed.

### Internal validity

The internal validity of a study refers to the correctness of the inferences drawn for the members of the source population upon part of whom they were based[43]. The internal validity of any epidemiological study is always at risk of being jeopardized by systematic error, also known as bias, which distorts the 'true' association between the main determinant and outcome. The three overarching forms of systematic error are information bias, confounding, and selection bias. In the coming paragraphs, I will elaborate on these types of bias and will discuss them in the context of the performed studies.

### Information bias and random measurement error

Information bias is caused by erroneous information—measurement error in case of continuous data and misclassification error in case of categorical data—on determinant, outcome, or both[43]. The impact of such error depends on whether it is systematic or random and whether it affects the main determinant or outcome[44]. Systematic error in a determinant or outcome variable can cause both overestimation and underestimation of an effect size. Random measurement error in a determinant biases results towards the null (i.e., regression dilution bias), whereas such error in an outcome variable widens the confidence intervals (i.e., reduces estimate precision)[44]. The Maastricht Study has implemented well-designed, standardized protocols in order to minimize misclassification and measurement error during data collection, handling, and analysis[73]. At the moment of analysis, for example, the first 24 hours of all CGM recordings were excluded because of insufficient calibration with SMBG values.

### Information bias related to continuous glucose monitoring-derived indices

CGM is the gold-standard method for daily GV quantification[29,30]. Nevertheless, information bias may have affected the analyses with CGM-based indices. First, in 9.3% of the participants, the CGM device either did not record the full 7-day period or it was not worn for this complete period of time. In *Chapter 2*, we evaluated the minimal

9

number of days to reliably assess daily GV, so that we could take this into consideration for the chapters of part II. Second, 7.8% of participants presented with data gaps in their CGM recording, which commonly lasted for several hours. Data gaps during the day would most probably result in underestimation of daily GV, while data gaps during the night would cause the opposite. Third, the 'catch-up' participants (29.8%) have undergone CGM, on average, 2 years later than all their other measurements. Although the participants' body weight and use of glucose-lowering medication, if applicable, had not materially changed in these 2 years, the CGM-derived indices of the 'catch-up' participants are more prone to random measurement error because the CGM-assessed indices are used as if they had been measured 2 years prior.

All the limitations outlined above are expected to have led to bias towards the null in the associations reported in part II[44]. We, therefore, excluded participants based on these three limitations as part of sensitivity analyses. Although exclusion due to data gaps or insufficient recording days generally increased the regression coefficients (e.g., associations with cf-PWV and ABI in *Chapter 6*), statistical significance was seldom altered, which may reflect a balanced trade-off between information bias and statistical power. The results after exclusion of 'catch-up' participants were less consistent. In addition to the bigger decrease in sample size, this could stem from the difference in GMS distribution after exclusion of 'catch-up' participants (i.e., smaller percentage of individuals with prediabetes and type 2 diabetes). In *Chapter 7*, which was written after *Chapter 8*, we instead opted to adjust the main analyses for the time interval between the 'catch-up visit' and all other measurements.

The high overall adherence to the CGM study protocol, which required SMBG four times daily in order to ensure adequate CGM calibration, is expected to have led to lower measurement error. However, the fact that participants were not blinded to the SMBG values may have led to altered behavior (e.g., healthier dietary choices or stricter glucose control) during the week of recording. This could have led to an underestimation of daily GV and, thus, cause bias, especially if individuals with worse GMS were more likely to alter their behavior. However, such bias is not very probable, since the CGM-derived indices of later days were not materially lower than those of earlier days, regardless of GMS (*Chapter 2*).

Last, there is larger relative error in the estimation of SD and CV than in that of MSG as a direct consequence of how they are calculated[26]. As higher random measurement error in the determinant is known to bias the regression estimate towards zero[44], one would expect to observe a stronger association with MSG, even if MSG and SD are of equal pathophysiological importance[26]. Since the standardized regression coefficients were relatively equal in magnitude for the fully adjusted associations of SD and MSG with cf-PWV (*Chapter 6*), this might suggest that the relevance of daily GV is greater in the development of aortic stiffness than that of

MSG. It would be interesting to explore how MSG and SD are related to different outcomes after the measurement error of MSG is artificially increased.

*Information bias related to incremental glucose peak*

The analyses in which IGP was used as main determinant may have also been negatively influenced by measurement error. First, while we only calculated IGP in individuals of whom all seven OGTT time points were available for analysis, the actual IGP value may have been underestimated in the individuals with a glucose peak not exactly at one of the sampled time points. The chance of underestimation would probably be larger in individuals with a glucose peak between 60 and 90 minutes or 90 and 120 minutes. This systematic error is expected to predominantly affect participants with prediabetes or type 2 diabetes, in whom the chance of a glucose peak after 1 hour is larger (*Chapter 2*). Underestimation of the main determinant in these generally less healthy participants may have led to overestimated associations. Second, IGP was calculated based on a single OGTT, which is known to have moderate reproducibility for GMS classification[153]. If the fasting and 2-hour post-load glucose values are moderately reproducible, the same could hold true for IGP. Due to this random measurement error, the regression results with IGP as main determinant may have been underestimated. Unfortunately, it was not possible to quantify bias resulting from the points addressed above with use of specific sensitivity analyses.

*Information bias related to glucose metabolism status and measures of hyperglycemia*

In The Maastricht Study, GMS was assessed with use of the gold-standard method (i.e., a standardized OGTT)[77]. Nevertheless, as described above, a single OGTT may misclassify GMS. Individuals classified with prediabetes based on their first OGTT are relatively more likely to receive a NGM classification based on their second OGTT[153]. In the prediabetes group, this may have led to underestimated results on, for example, the distribution of CGM-derived indices in *Chapter 2* or the association with RNFL thickness in *Chapter 7*.

In *Chapter 7*, we also used multiple continuous measures of glycemia as main determinants. FPG and 2-hour post-load glucose were assessed with a standard enzymatic hexokinase reference method and HbA1c was measured with ion-exchange high performance liquid chromatography[73]. All were measured only once and are, thus, susceptible to some degree of biological and analytical variability[154]. Skin autofluorescence was assessed with use of AGE Reader, a validated and reproducible device that estimates AGE accumulation in the skin of the forearm[155,156]. In case of skin autofluorescence, we used the mean of three consecutive measurements to reduce random measurement error.

9

*Information bias related to the investigated outcomes*

Measurement error in the outcome variables may have caused information bias as well, which, if random, would reduce estimate precision, and correspondingly would impede statistical significance[44].

In *Chapters 5 and 6*, we investigated the association of daily GV with several vascular measurements. Aortic stiffness was determined with use of the gold-standard method (i.e., cf-PWV)[84]. Local carotid characteristics (e.g., cIMT, lumen diameter, wall distension) were assessed with validated ultrasonography techniques[157,158]. Both cf-PWV and carotid properties were measured in triplicate and the median measurement was used to reduce random measurement error. Local carotid stiffness was evaluated with the carDC formula. Mean and pulsatile circumferential wall stress were based on the Lamé equation[90,91]. CarDC and measures of circumferential wall stress could be more prone to measurement error than cf-PWV and cIMT, as their calculation is based on formulas that incorporate multiple variables, which are all per se susceptible to measurement error. ABI was automatically determined based on simultaneous blood pressure measurements at both ankles and upper arms. We used the lowest ABI in our analyses and excluded individuals with an ABI above 1.4, as recommended by the American Heart Association[88].

In *Chapter 5*, microvascular function was assessed at two sites. First, retinal microvascular dilation in response to flicker light was measured with the Dynamic Vessel Analyzer and calculated automatically by the appurtenant software. Second, heat-induced skin hyperemic response was assessed with a laser-Doppler system equipped with a thermostatic probe. Although both are reproducible and valid measures of the microcirculation[112,113], they were only performed once due to time constraints, which prevented us from reducing the effect of biological variability by, for example, averaging repeated measurements.

In *Chapter 7*, RNFL thickness was assessed with use of OCT. The OCT software is known for its high intra- and interindividual reliability in the quantification of RNFL thickness (i.e., ICCs of 0.97 and 0.96, respectively)[159]. In approximately half of the participants, data on both eyes were available, which allowed us to average RNFL thickness in order to further reduce measurement error.

In *Chapter 8*, we calculated three cognitive domain composite scores based on several validated neuropsychological tests[160-163]. Moreover, these cognitive domains were used to construct an overall cognitive performance composite score. Composite scores are commonly employed to group variables with conceptual or statistical overlap[164]. By using a composite score outcome instead of multiple outcomes, type I error rate can be controlled. In addition, by grouping several variables into a composite score, the influence of biological variability and, thus, of random measurement error are reduced. Indeed, the estimate of the association between IGP and overall cognitive performance was slightly more precise compared to the estimates of the associations between IGP and individual cognitive domains.

*Confounding, overadjustment bias, and multicollinearity*

Confounding occurs when the 'true' association between main determinant and outcome is distorted by one or more extraneous factors (i.e., confounders)[43]. By definition, a confounder is a variable that is associated with the main determinant, can cause the outcome, and does not lie on the causal path between both. In the context of observational research, certain statistical methods (e.g., multiple linear regression) can be used to correct for confounding. By virtue of the extensive phenotyping approach of The Maastricht Study[73], we were able to correct for a large array of potential confounders, which generally included demographics (i.e., age, sex, and educational level), CVD risk factors (e.g., blood pressure, lipid profile, and smoking habits), lifestyle factors (e.g., alcohol use, diet, and physical activity), and an average glycemic index (e.g., $HbA_{1c}$, FPG, or MSG).

Notwithstanding our efforts to address this form of bias by including the aforementioned potential confounders in the regression analyses, we cannot fully eliminate the presence of residual confounding. At the moment of analysis, certain potential confounders (e.g., markers of inflammation, therapy adherence, environmental factors) were not available in the study populations and could, thus, not be included in the regression models. Residual confounding may also have occurred due to misclassified or imperfectly measured confounders (i.e., information bias). Dietary intake, assessed by food frequency questionnaire, and physical activity, quantified with the CHAMPS questionnaire, may be more prone to inaccurate recall or even recall bias[165-167]. This could, in part, explain why the regression coefficients were, in general, not materially altered by additional adjustment for these lifestyle factors.

Although adjustment for an extensive number of potential confounders is intended to reduce confounding bias, it may conversely cause overadjustment bias if the included variables are mediators or descendants of the outcome[168]. For example, urinary albumin excretion, as a proxy of generalized endothelial dysfunction, may both confound and mediate the association between daily GV and cognition[169]. Likewise, history of CVD and retinopathy might be viewed as confounders but should also be regarded as descendants of the outcome in the associations of daily GV with cf-PWV and RNFL thickness, respectively. To prevent overadjustment bias, we avoided inclusion of such variables in the main regression models and instead opted for exploring their effect in additional analyses.

In specific situations, multivariable linear regression, which is an established method to correct for confounding, can cause statistical concern. Namely, multicollinearity can arise if a strong correlation is present between the main determinant and one or more confounders[94,170,171]. The statistical phenomenon follows from the fact that strongly correlated variables provide little unique information. Consequently, the regression coefficients for these variables are based on limited information as well. This can lead to unreliable regression estimates with regard to association strength, direction and precision. Research on consequences of CGM-based

9

daily GV is susceptible to negative effects of multicollinearity due to the well-known strong correlation between the main determinant (i.e., SD) and key confounder (i.e., MSG)[26]. Multicollinearity can be especially concerning in analyses with a small sample size, which generally holds true for studies with CGM because of its costliness and relative invasiveness[26,172], or in case of a weak association between main determinant and outcome[171], which is increasingly common for modern epidemiological research on more subtle risk factors[173]. **Figure 9.2** illustrates the impact of these three conditions according to the conceptual depiction of the associations of daily GV indices with two cognitive domains (*Chapter 8*).
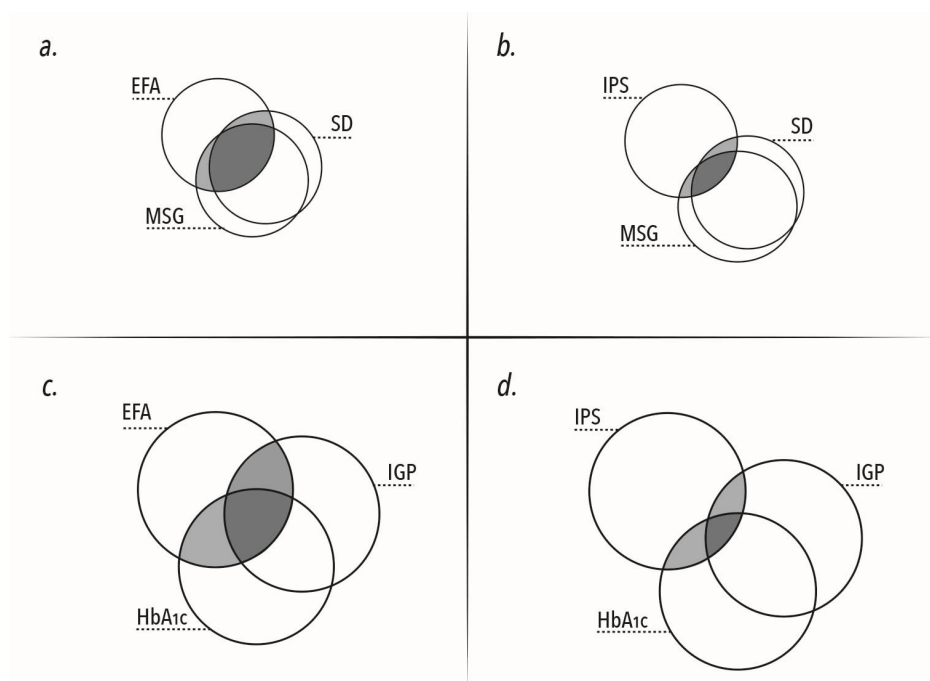


Figure 9.2    Conceptual depiction of the associations of SD and MSG or IGP and HbA_{1c} with executive function and attention (EFA, panels a and c) or information processing speed (IPS, panels b and d). More overlap between explanatory variables depicts a stronger correlation (panel ab vs. cd). More overlap of the explanatory variables with the outcome variable depicts a stronger association in the models before the other explanatory variable was added (panel ac vs. bd). Larger circles illustrate a larger sample size. Dark grey depicts the shared contribution of explanatory variables. Light grey represents the unique contribution of either explanatory variable upon which the regression coefficients are based. The larger the light grey areas are, the smaller the negative effects of multicollinearity. While in case of panel c, concern for multicollinearity would, thus, be the lowest (i.e., less strong correlation between explanatory variables, stronger association with outcome measure, and larger sample size), concern would be the greatest in case of panel b. Abbreviations: SD, standard deviation; MSG, mean sensor glucose; IGP, incremental glucose peak.

Although specific multicollinearity diagnostics, such as variance inflation factor (VIF), were not strikingly abnormal according to traditional cut-off values (i.e., VIF >5 or 10), previous studies have shown them to be insufficiently reliable[94]. In light of the aforementioned conditions, the opposite coefficient directions for SD and MSG with several outcome variables (e.g., ABI in *Chapter 6*, information processing speed in *Chapter 8*) indeed indicate the effects of multicollinearity. Hence, it is vital to report both the associations of SD and MSG with an outcome, as inspection of one only of the estimates can lead to very different conclusions. In *Chapter 8*, for example, prior to adjustment of the model with SD as main determinant for MSG, or vice versa, neither were associated with information processing speed. However, when added together, the associations of SD and MSG with information processing speed turned positive and negative, respectively. Furthermore, the regression estimates were, as could be predicted logically, almost exact opposites, which is biologically implausible. The larger measurement error of SD compared to MSG[26], and the consequently larger relative regression dilution bias[44], may explain why the regression estimate of SD turned negative in most instances.

Common solutions to multicollinearity (e.g., leaving out strongly correlated variables, constructing a composite score)[94,170] were not feasible in the context of our research questions. Although the construction of a ratio variable (i.e., coefficient of variation; CV) was, on the one hand, a viable strategy, it could, on the other hand, have intrinsically biased results[174,175]. CV was still moderately correlated with FPG and HbA$_{1c}$ (*Chapter 2*), which calls into question whether CV is fully MSG-adjusted. Hence, we investigated associations with CV as part of secondary or additional analyses. In addition, we repeated the main analyses using ridge regression. This L2-regularized form of linear regression is, amongst others, a valid statistical method to counter a degree of model instability caused by multicollinearity[176]. Ridge regression estimates are computed by combining the residual sum of squares and predefined coefficient penalization. By slightly biasing the regression coefficients, this method can strongly reduce inflated variances that arise when high multicollinearity exists. However, ridge regression was no panacea for the effects of multicollinearity either. The method is limited by the lack of consensus on the degree of penalization. Our levels of penalization were chosen pragmatically on the lambda required to reduce the VIF of the model with adjustment for MSG back to the VIF of the model without such adjustment (or halfway back). In some instances (e.g., ABI; *Chapter 6*), more pronounced penalization resulted in consistent association directions, which led us to believe that the effects of multicollinearity were indeed (partially) countered. However, in most situations the coefficients derived from ridge regression differed little from those derived from regular multiple linear regression. In these cases, it is unclear whether multicollinearity had little effect on the associations or whether it was insufficiently addressed, for example, because of too little penalization.

9

*Selection bias and range restriction*

Our results may have been biased as a result of the procedures used for participant selection or due to non-random factors that affect study participation (i.e., selection bias)[43]. All main analyses were performed according to the complete-case analysis principle, which entails that participants were excluded from analysis if at least one variable (determinant, outcome, or confounder) is missing[177]. This method has the advantage of making the comparison of consecutive regression models clearer, since their sample sizes are the same. However, it, in addition to reducing statistical power of earlier models, may introduce selection bias in these models if participant exclusion due to missing confounders is non-random. We, therefore, consistently compared characteristics of in- and excluded individuals in order to establish whether missing values were random or systematic.

Based on the aforementioned comparisons, missing values on most key variables were regarded as random. However, selection bias is probable in certain instances. First, several OGTT-derived indices –most importantly IGP– were systematically missing in participants with an OGTT contraindication (i.e., elevated plasma glucose, insulin use, or both). Use of IGP, thus, inherently caused range restriction because it could not be assessed in participants who are expected to have the greatest daily GV (i.e., those with an OGTT contraindication)[178]. These excluded individuals are also generally less healthy, which implicates that an important part of the outermost spectrum of determinant and outcome data is missing. This form of range restriction is expected to lead to underestimated associations between IGP and all investigated outcome variables[178]. The use of IGP as determinant also causes range restriction for $HbA_{1c}$ or FPG, which were used to adjust the associations of IGP with outcome measures. If the range restriction would be relatively larger for $HbA_{1c}$ and FPG, this could be to the advantage of IGP.

The effects of range restriction were also observed in Part I of this thesis. In *Chapter 2*, correlation coefficients of CGM-assessed glycemic indices with OGTT-derived indices and HbA1c were investigated for the total CGM study population and for different GMS strata. The stratified correlations were much weaker than the overall correlation for the total study population. In *Chapter 3*, we presented correlation coefficients between the actual and predicted glucose concentrations for the total study population and the individuals with type 2 diabetes of the validation dataset. As expected, the correlations were stronger in the total study population.

In addition to causing a degree of measurement error (see above), the 'catch-up' visit procedure could have caused selection bias if, for example, less healthy individuals were less likely to return to The Maastricht Study to undergo CGM. Indeed, overall cognitive performance, memory function, and information processing speed were slightly worse in participants who were invited for but did not attend a 'catch-up' visit (*Chapter 8*).

*Causal inference*

A cause can be defined as an antecedent event, condition or characteristic that was required for the disease to occur at the moment it did, given that all other conditions are fixed[80]. The temporality criterion of this definition can aid in establishing whether the determinant can indeed be the cause of the outcome or whether, in fact, the opposite (or both) is true. Since all studies in this thesis were cross-sectional by design, we were unable to rule out reverse causality for our findings on the potential consequences of daily GV (Part II). In *Chapters 5 and 6*, we theorize that greater daily GV causes aortic stiffening, which is consistent with prospective findings on type 2 diabetes and measures of glycemia[81,82]. However, as greater arterial stiffness has been associated with a higher incidence of type 2 diabetes[179], it could increase daily GV as well. In *Chapters 7*, we posit that both greater average glycaemia and daily GV provoke RNFL thinning. In addition, we believe that these associations may be bidirectional, as neurodegeneration can impair microvascular function and thus can lead to or aggravate hyperglycemia[10]. In *Chapter 8*, we presume that daily GV impairs cognitive performance. Conversely, worse cognitive performance may influence behavior (e.g., therapy adherence) and, thereby, could increase daily GV. Prospective studies on these potential consequences are needed to gain more insight into the causality of the associations of daily GV with aortic stiffness, neurodegeneration, and cognitive performance.

*Adjustment for multiple testing*

Adjustment for multiple testing refers to correction of the statistical significance threshold (i.e., alpha) based on the number of statistical comparisons made[180]. Because multiple main (primary and secondary) determinants were used per chapter (Part II), one could argue that adjustment for multiple testing would be required[180]. However, since the different indices of daily GV are conceptually and statistically related[34,35], it would be overly strict to enforce a multiple testing adjustment based on this[180]. As multiple outcomes were used per chapter (Part II), it could again be argued that multiple testing adjustment is needed. However, the reflex to always correct for multiple comparisons is not per definition desirable[180,181]. Namely, it can lead to more mistakes in interpretation if data are actual observations on nature, which commonly holds true, rather than random numbers[181]. Furthermore, adjustment for multiple testing will increase the chance that possibly important findings will be missed (i.e., type 2 error). Especially in a more explorative setting, it is more preferable to risk examining leads that turn out to be untrue if that ensures that possibly important findings are not overlooked[181]. Because we have not adjusted for multiple comparisons in any of the chapters, we consistently underscore the importance of confirming our results (e.g., in prospective studies) in order to discriminate between bona fide and spurious findings.

*Statistical power*

Statistical power, which refers to the number of participants that are required to avoid falsely rejecting the null hypothesis (i.e., type II error)[92], may have been an issue in part II of this thesis. Namely, in case of relatively modest effect sizes, which are particularly common in modern epidemiology[173], large samples sizes are needed to provide sufficient statistical power[92]. However, due to costliness of CGM devices[26] our study was limited to nearly 900 inclusions. Although this presently makes this subpopulation of The Maastricht Study one of the largest observational cohorts on CGM, statistical power may still have been inadequate. Namely, in multiple chapters, regression estimates were indicative of an association but failed to reach statistical significance (e.g., between SD and ABI in *Chapter 6* and SD and RNFL thickness in *Chapter 7*). Although a power calculation is based on several simplifying assumptions, it can grant more insight into whether statistical power was sufficient[92]. For the purpose of this *Summary and general discussion*, I performed a post-hoc power calculation for the association between SD and RNFL thickness. Given the modest effect size (Cohen's $f^2$=0.0046), alpha of 0.05, and beta of 0.20, a sample size of just over 1,700 participants would be required to reach sufficient statistical power. As such, the current sample size of 622 participants only provided 39% power to avoid type II error. The CGM-based findings of this thesis are nevertheless useful, as regression estimates and confidence intervals provide valuable information, regardless of statistical significance[182].

*External validity*

External validity reflects the applicability or generalizability of our findings to individuals who are not part of the study or source population[43]. All our findings were based on a relatively large, well-characterized population of 40- to 75-year-old, mostly Caucasian individuals who all had access to high-quality health care and were predominately classified with either NGM, prediabetes, or type 2 diabetes. Strictly speaking, the generalizability of our results to other ethnicities, age groups, or other types of diabetes –most importantly type 1 diabetes– would require further study. However, based on the general pathophysiology of hyperglycemia and its complications[99], we expect our findings on the potential consequences of daily GV to be largely applicable to other populations[43]. Indeed, daily GV was also observed to be associated with, for example, cognitive performance in Chinese elderly individuals[42].

## Conclusions and future directions

In conclusion, this thesis provides means to improve daily GV measurement and CGM, both in clinical practice and scientific studies, by assessing recording period reliability, exploring machine-learning based glucose prediction, establishing reference values for CGM-derived indices, and evaluating the legitimacy of novel indices. This thesis also

grants more insight into whether daily GV should be regarded as an independent determinant of diabetes-associated complications. Such a role is most convincing for greater aortic stiffness (i.e., higher cf-PWV), neurodegeneration (i.e., lower RNFL thickness), and worse overall cognitive performance (specifically executive function and attention), as the associations with these measures persisted after adjustment for all potential confounders, including average glycemia, and were consistent for OGTT- and CGM-based indices of daily GV. The impact of this thesis' findings is more extensively discussed in the next chapter (See *Scientific and societal impact*).

While this thesis aimed to fully address the methodological limitations of prior studies in the field, certain aspects could still be improved in future studies. As the performed studies were all cross-sectional by design, prospective studies are required in order to rule out reverse causality. Further, although the study populations used in this thesis were relatively large compared to most previous studies, insufficient statistical power may still have affected the associations. Especially in light of the observed effects of multicollinearity on estimate strength, direction and precision, larger sample sizes of approximately 2,000 participants are needed to draw more definitive conclusions on whether daily GV could also be an independent determinant of microvascular dysfunction, atherosclerosis, and worse performance on other cognitive domains.

Prospective confirmation of our findings would further justify studying whether targeted lowering of daily GV can improve clinical outcomes. In that regard, it would be important to comprehensively examine the determinants of daily GV as well as their relative contribution. Beta cell function, and—to a lesser extent—diet and physical activity are expected to be key determinants[78,183]. Further, certain types of glucose-lowering drugs or specific therapeutic regimens may have different effects on daily GV. For example, medication that predominantly reduces postprandial glucose excursions (e.g., incretin-modulating drugs) may improve daily GV more than those that mainly improve insulin sensitivity[78]. In addition, increased usage of CGM, in particular of flash glucose monitoring devices, in combination with reporting of intuitive metrics such as TIR are expected to further optimize therapy and adherence[31,36]. In case of near or complete loss of beta cell function, closed-loop insulin delivery systems are probably the most effective way to safely reduce daily GV[184].

9

# References

1.  Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol. 2018;14(2):88-98.
2.  Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. Diabetes Res Clin Pract. 2019;157:107843.
3.  American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S66-S76.
4.  American Diabetes A. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S14-S31.
5.  Preis SR, Hwang SJ, Coady S, Pencina MJ, D'Agostino RB, Sr., Savage PJ, et al. Trends in all-cause and cardiovascular disease mortality among women and men with and without diabetes mellitus in the Framingham Heart Study, 1950 to 2005. Circulation. 2009;119(13):1728-35.
6.  Taylor KS, Heneghan CJ, Farmer AJ, Fuller AM, Adler AI, Aronson JK, et al. All-cause and cardiovascular mortality in middle-aged people with type 2 diabetes compared with people without diabetes in a large U.K. primary care database. Diabetes Care. 2013;36(8):2366-71.
7.  Bragg F, Holmes MV, Iona A, Guo Y, Du H, Chen Y, et al. Association Between Diabetes and Cause-Specific Mortality in Rural and Urban Areas of China. JAMA. 2017;317(3):280-9.
8.  Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical Update: Cardiovascular Disease in Diabetes Mellitus: Atherosclerotic Cardiovascular Disease and Heart Failure in Type 2 Diabetes Mellitus - Mechanisms, Management, and Clinical Considerations. Circulation. 2016;133(24):2459-502.
9.  American Diabetes A. 11. Microvascular Complications and Foot Care: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S135-S51.
10. Stehouwer CDA. Microvascular Dysfunction and Hyperglycemia: A Vicious Cycle With Widespread Consequences. Diabetes. 2018;67(9):1729-41.
11. Xue M, Xu W, Ou YN, Cao XP, Tan MS, Tan L, et al. Diabetes mellitus and risks of cognitive impairment and dementia: A systematic review and meta-analysis of 144 prospective studies. Ageing Res Rev. 2019;55:100944.
12. Cai X, Zhang Y, Li M, Wu JH, Mai L, Li J, et al. Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis. BMJ. 2020;370:m2297.
13. Wong MS, Gu K, Heng D, Chew SK, Chew LS, Tai ES. The Singapore impaired glucose tolerance follow-up study: does the ticking clock go backward as well as forward? Diabetes Care. 2003;26(11):3024-30.
14. Jing X, Chen J, Dong Y, Han D, Zhao H, Wang X, et al. Related factors of quality of life of type 2 diabetes patients: a systematic review and meta-analysis. Health Qual Life Outcomes. 2018;16(1):189.
15. Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Barnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. Lancet Diabetes Endocrinol. 2017;5(6):423-30.
16. Bommer C, Sagalova V, Heesemann E, Manne-Goehler J, Atun R, Barnighausen T, et al. Global Economic Burden of Diabetes in Adults: Projections From 2015 to 2030. Diabetes Care. 2018;41(5):963-70.
17. American Diabetes A. 4. Comprehensive Medical Evaluation and Assessment of Comorbidities: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S37-S47.
18. Kovatchev B, Cobelli C. Glucose Variability: Timing, Risk Analysis, and Relationship to Hypoglycemia in Diabetes. Diabetes Care. 2016;39(4):502-10.
19. Brownlee M, Hirsch IB. Glycemic variability: a hemoglobin A1c-independent risk factor for diabetic complications. JAMA. 2006;295(14):1707-8.
20. Hirsch IB. Glycemic Variability and Diabetes Complications: Does It Matter? Of Course It Does! Diabetes Care. 2015;38(8):1610-4.

21. Ceriello A. Glucose Variability and Diabetic Complications: Is It Time to Treat? Diabetes Care. 2020;43(6):1169-71.
22. Ceriello A, Monnier L, Owens D. Glycaemic variability in diabetes: clinical and therapeutic implications. Lancet Diabetes Endocrinol. 2019;7(3):221-30.
23. Bergenstal RM. Glycemic Variability and Diabetes Complications: Does It Matter? Simply Put, There Are Better Glycemic Markers! Diabetes Care. 2015;38(8):1615-21.
24. Cavalot F. Do data in the literature indicate that glycaemic variability is a clinical problem? Glycaemic variability and vascular complications of diabetes. Diabetes Obes Metab. 2013;15 Suppl 2:3-8.
25. Siegelaar SE, Holleman F, Hoekstra JB, DeVries JH. Glucose variability; does it matter? Endocr Rev. 2010;31(2):171-82.
26. Rodbard D. The challenges of measuring glycemic variability. J Diabetes Sci Technol. 2012;6(3): 712-5.
27. Service FJ, Molnar GD, Rosevear JW, Ackerman E, Gatewood LC, Taylor WF. Mean amplitude of glycemic excursions, a measure of diabetic instability. Diabetes. 1970;19(9):644-55.
28. Kilpatrick ES, Rigby AS, Atkin SL. The effect of glucose variability on the risk of microvascular complications in type 1 diabetes. Diabetes Care. 2006;29(7):1486-90.
29. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International Consensus on Use of Continuous Glucose Monitoring. Diabetes Care. 2017;40(12):1631-40.
30. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care. 2019;42(8):1593-603.
31. Lin R, Brown F, James S, Jones J, Ekinci E. Continuous glucose monitoring: A review of the evidence in type 1 and 2 diabetes mellitus. Diabet Med. 2021:e14528.
32. Klonoff DC, Ahn D, Drincic A. Continuous glucose monitoring: A review of the technology and clinical use. Diabetes Res Clin Pract. 2017;133:178-92.
33. van den Boom L, Karges B, Auzanneau M, Rami-Merhar B, Lilienthal E, von Sengbusch S, et al. Temporal Trends and Contemporary Use of Insulin Pump Therapy and Glucose Monitoring Among Children, Adolescents, and Adults With Type 1 Diabetes Between 1995 and 2017. Diabetes Care. 2019;42(11):2050-6.
34. Rodbard D. Glucose Variability: A Review of Clinical Applications and Research Developments. Diabetes Technol Ther. 2018;20(S2):S25-S215.
35. Rodbard D. Glucose Time In Range, Time Above Range, and Time Below Range Depend on Mean or Median Glucose or HbA1c, Glucose Coefficient of Variation, and Shape of the Glucose Distribution. Diabetes Technol Ther. 2020.
36. Advani A. Positioning time in range in diabetes management. Diabetologia. 2020;63(2):242-52.
37. Xing D, Kollman C, Beck RW, Tamborlane WV, Laffel L, Buckingham BA, et al. Optimal sampling intervals to assess long-term glycemic control using continuous glucose monitoring. Diabetes Technol Ther. 2011;13(3):351-8.
38. Neylon OM, Baghurst PA, Cameron FJ. The Minimum Duration of Sensor Data From Which Glycemic Variability Can Be Consistently Assessed. J Diabetes Sci Technol. 2014;8(2):273-6.
39. Riddlesworth TD, Beck RW, Gal RL, Connor CG, Bergenstal RM, Lee S, et al. Optimal Sampling Duration for Continuous Glucose Monitoring to Determine Long-Term Glycemic Control. Diabetes Technol Ther. 2018;20(4):314-6.
40. Baghurst PA, Rodbard D, Cameron FJ. The minimum frequency of glucose measurements from which glycemic variation can be consistently assessed. J Diabetes Sci Technol. 2010;4(6):1382-5.
41. Rizzo MR, Marfella R, Barbieri M, Boccardi V, Vestini F, Lettieri B, et al. Relationships between daily acute glucose fluctuations and cognitive performance among aged type 2 diabetic patients. Diabetes Care. 2010;33(10):2169-74.
42. Zhong Y, Zhang XY, Miao Y, Zhu JH, Yan H, Wang BY, et al. The relationship between glucose excursion and cognitive function in aged type 2 diabetes patients. Biomed Environ Sci. 2012;25(1):1-7.

**9**

43. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. 3rd ed. Philadelphia, USA: Lippincott Williams & Wilkins; 2008.

44. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. BMJ. 2010;340:c2289.

45. Rodbard D. Continuous Glucose Monitoring: A Review of Successes, Challenges, and Opportunities. Diabetes Technol Ther. 2016;18 Suppl 2:S3-S13.

46. Cobelli C, Renard E, Kovatchev B. Artificial pancreas: past, present, future. Diabetes. 2011;60(11):2672-82.

47. Blauw H, Keith-Hynes P, Koops R, DeVries JH. A Review of Safety and Design Requirements of the Artificial Pancreas. Ann Biomed Eng. 2016;44(11):3158-72.

48. Mohebbi A, Johansen AR, Hansen N, Christensen PE, Tarp JM, Jensen ML, et al. Short Term Blood Glucose Prediction based on Continuous Glucose Monitoring Data. arXiv e-prints [Internet]. 2020 February 01, 2020:[arXiv:2002.02805 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2020arXiv200202805M.

49. Alwosheel A, van Cranenburgh S, Chorus CG. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. J Choice Model. 2018;28:167-82.

50. Weisman A, Bai JW, Cardinez M, Kramer CK, Perkins BA. Effect of artificial pancreas systems on glycaemic control in patients with type 1 diabetes: a systematic review and meta-analysis of outpatient randomised controlled trials. Lancet Diabetes Endocrinol. 2017;5(7):501-12.

51. Bode BW, Schwartz S, Stubbs HA, Block JE. Glycemic characteristics in continuously monitored patients with type 1 and type 2 diabetes: normative values. Diabetes Care. 2005;28(10):2361-6.

52. Marling C, Bunescu RC, editors. The OhioT1DM Dataset For Blood Glucose Level Prediction. KHD@IJCAI; 2018.

53. Martinsson J, Schliep A, Eliasson B, Mogren O. Blood Glucose Prediction with Variance Estimation Using Recurrent Neural Networks. Journal of Healthcare Informatics Research. 2020;4(1):1-18.

54. Kriventsov S, Lindsey A, Hayeri A. The Diabits App for Smartphone-Assisted Predictive Monitoring of Glycemia in Patients With Diabetes: Retrospective Observational Study. JMIR Diabetes. 2020;5(3):e18660.

55. Woldaregay AZ, Arsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, et al. Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes. Artif Intell Med. 2019;98:109-34.

56. Li K, Liu C, Zhu T, Herrero P, Georgiou P. GluNet: A Deep Learning Framework for Accurate Glucose Forecasting. IEEE J Biomed Health Inform. 2020;24(2):414-23.

57. Chen J, Li K, Herrero P, Zhu T, Georgiou P, editors. Dilated Recurrent Neural Network for Short-time Prediction of Glucose Concentration. KHD@IJCAI; 2018.

58. Stanford KI, Goodyear LJ. Exercise and type 2 diabetes: molecular mechanisms regulating glucose uptake in skeletal muscle. Adv Physiol Educ. 2014;38(4):308-14.

59. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. Am J Epidemiol. 2012;176(6):473-81.

60. Zhou J, Li H, Ran X, Yang W, Li Q, Peng Y, et al. Establishment of normal reference ranges for glycemic variability in Chinese subjects using continuous glucose monitoring. Med Sci Monit. 2011;17(1):CR9-13.

61. Zhou J, Li H, Ran X, Yang W, Li Q, Peng Y, et al. Reference values for continuous glucose monitoring in Chinese subjects. Diabetes Care. 2009;32(7):1188-93.

62. Gude F, Diaz-Vidal P, Rua-Perez C, Alonso-Sampedro M, Fernandez-Merino C, Rey-Garcia J, et al. Glycemic Variability and Its Association With Demographics and Lifestyles in a General Adult Population. J Diabetes Sci Technol. 2017;11(4):780-90.

63. Horn PS, Pesce AJ. Reference intervals: an update. Clin Chim Acta. 2003;334(1-2):5-23.

64. Monnier L, Colette C, Wojtusciszyn A, Dejager S, Renard E, Molinari N, et al. Toward Defining the Threshold Between Low and High Glucose Variability in Diabetes. Diabetes Care. 2017;40(7):832-8.

65. Rodbard D. Hypo- and hyperglycemia in relation to the mean, standard deviation, coefficient of variation, and nature of the glucose distribution. Diabetes Technol Ther. 2012;14(10):868-76.

66.  Hall H, Perelman D, Breschi A, Limcaoco P, Kellogg R, McLaughlin T, et al. Glucotypes reveal new patterns of glucose dysregulation. PLoS Biol. 2018;16(7):e2005143.

67.  Peyser TA, Balo AK, Buckingham BA, Hirsch IB, Garcia A. Glycemic Variability Percentage: A Novel Method for Assessing Glycemic Variability from Continuous Glucose Monitor Data. Diabetes Technol Ther. 2018;20(1):6-16.

68.  Cui X, Abduljalil A, Manor BD, Peng C-K, Novak V. Multi-scale glycemic variability: a link to gray matter atrophy and cognitive decline in type 2 diabetes. PloS one. 2014;9(1).

69.  Zaccardi F, Khunti K. Glucose dysregulation phenotypes - time to improve outcomes. Nat Rev Endocrinol. 2018;14(11):632-3.

70.  Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.

71.  Hamideh D, Arellano B, Topol EJ, Steinhubl SR. Your digital nutritionist. Lancet. 2019;393(10166):19.

72.  Li J, Li X, Zhang S, Snyder M. Gene-Environment Interaction in the Era of Precision Medicine. Cell. 2019;177(1):38-44.

73.  Schram MT, Sep SJ, van der Kallen CJ, Dagnelie PC, Koster A, Schaper N, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol. 2014;29(6):439-51.

74.  Faerch K, Amadid H, Nielsen LB, Ried-Larsen M, Karstoft K, Persson F, et al. Protocol for a randomised controlled trial of the effect of dapagliflozin, metformin and exercise on glycaemic variability, body composition and cardiovascular risk in prediabetes (the PRE-D Trial). BMJ Open. 2017;7(5):e013802.

75.  Lennon H, Kelly S, Sperrin M, Buchan I, Cross AJ, Leitzmann M, et al. Framework to construct and interpret latent class trajectory modelling. BMJ Open. 2018;8(7):e020683.

76.  Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical research. Circ Res. 2015;116(1):116-26.

77.  WHO. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. WHO. 2006.

78.  Zhou Z, Sun B, Huang S, Zhu C, Bian M. Glycemic variability: adverse clinical outcomes and how to improve it? Cardiovasc Diabetol. 2020;19(1):102.

79.  Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080.

80.  Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health. 2005;95 Suppl 1:S144-50.

81.  Stehouwer CD, Henry RM, Ferreira I. Arterial stiffness in diabetes and the metabolic syndrome: a pathway to cardiovascular disease. Diabetologia. 2008;51(4):527-39.

82.  Prenner SB, Chirinos JA. Arterial stiffness in diabetes mellitus. Atherosclerosis. 2015;238(2):370-9.

83.  Zieman SJ, Melenovsky V, Kass DA. Mechanisms, pathophysiology, and therapy of arterial stiffness. Arterioscler Thromb Vasc Biol. 2005;25(5):932-43.

84.  Laurent S, Cockcroft J, Van Bortel L, Boutouyrie P, Giannattasio C, Hayoz D, et al. Expert consensus document on arterial stiffness: methodological issues and clinical applications. Eur Heart J. 2006;27(21):2588-605.

85.  Vlachopoulos C, Aznaouridis K, Stefanadis C. Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. J Am Coll Cardiol. 2010;55(13):1318-27.

86.  Sequi-Dominguez I, Cavero-Redondo I, Alvarez-Bueno C, Pozuelo-Carrascosa DP, Nunez de Arenas-Arroyo S, Martinez-Vizcaino V. Accuracy of Pulse Wave Velocity Predicting Cardiovascular and All-Cause Mortality. A Systematic Review and Meta-Analysis. J Clin Med. 2020;9(7).

87.  van Sloten TT, Sedaghat S, Laurent S, London GM, Pannier B, Ikram MA, et al. Carotid stiffness is associated with incident stroke: a systematic review and individual participant data meta-analysis. J Am Coll Cardiol. 2015;66(19):2116-25.

88.  Aboyans V, Criqui MH, Abraham P, Allison MA, Creager MA, Diehm C, et al. Measurement and interpretation of the ankle-brachial index: a scientific statement from the American Heart Association. Circulation. 2012;126(24):2890-909.

89.  Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev. 2013;93(1):137-88.

9

90.  Geijselaers SL, Sep SJ, Schram MT, van Boxtel MP, van Sloten TT, Op Het Roodt J, et al. Carotid circumferential wall stress is not associated with cognitive performance among individuals in late middle age: The Maastricht Study. Atherosclerosis. 2018;276:15-22.

91.  Henry RM, Kostense PJ, Dekker JM, Nijpels G, Heine RJ, Kamp O, et al. Carotid arterial remodeling: a maladaptive phenomenon in type 2 diabetes but not in impaired glucose metabolism: the Hoorn study. Stroke. 2004;35(3):671-6.

92.  Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J. 2003;20(5):453-8.

93.  Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ. 1995;311(7003):485.

94.  Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. Epidemiology (Sunnyvale). 2016;6(2).

95.  McEniery CM, Yasmin, Hall IR, Qasem A, Wilkinson IB, Cockcroft JR, et al. Normal vascular aging: differential effects on wave reflection and aortic pulse wave velocity: the Anglo-Cardiff Collaborative Trial (ACCT). J Am Coll Cardiol. 2005;46(9):1753-60.

96.  Brohall G, Oden A, Fagerberg B. Carotid artery intima-media thickness in patients with Type 2 diabetes mellitus and impaired glucose tolerance: a systematic review. Diabet Med. 2006;23(6):609-16.

97.  Ceriello A, Kilpatrick ES. Glycemic variability: both sides of the story. Diabetes Care. 2013;36 Suppl 2:S272-5.

98.  Monnier L, Mas E, Ginet C, Michel F, Villon L, Cristol JP, et al. Activation of oxidative stress by acute glucose fluctuations compared with sustained chronic hyperglycemia in patients with type 2 diabetes. JAMA. 2006;295(14):1681-7.

99.  Brownlee M. The pathobiology of diabetic complications: a unifying mechanism. Diabetes. 2005;54(6):1615-25.

100.  van Eupen MG, Schram MT, van Sloten TT, Scheijen J, Sep SJ, van der Kallen CJ, et al. Skin Autofluorescence and Pentosidine Are Associated With Aortic Stiffening: The Maastricht Study. Hypertension. 2016;68(4):956-63.

101.  Jones SC, Saunders HJ, Qi W, Pollock CA. Intermittent high glucose enhances cell growth and collagen synthesis in cultured human tubulointerstitial cells. Diabetologia. 1999;42(9):1113-9.

102.  Ceriello A, Esposito K, Piconi L, Ihnat MA, Thorpe JE, Testa R, et al. Oscillating glucose is more deleterious to endothelial function and oxidative stress than mean glucose in normal and type 2 diabetic patients. Diabetes. 2008;57(5):1349-54.

103.  International Hypoglycaemia Study G. Hypoglycaemia, cardiovascular disease, and mortality in diabetes: epidemiology, pathogenesis, and management. Lancet Diabetes Endocrinol. 2019;7(5):385-96.

104.  Gimenez M, Gilabert R, Monteagudo J, Alonso A, Casamitjana R, Pare C, et al. Repeated episodes of hypoglycemia as a potential aggravating factor for preclinical atherosclerosis in subjects with type 1 diabetes. Diabetes Care. 2011;34(1):198-203.

105.  Wakasugi S, Mita T, Katakami N, Okada Y, Yoshii H, Osonoi T, et al. Associations between continuous glucose monitoring-derived metrics and arterial stiffness in Japanese patients with type 2 diabetes. Cardiovasc Diabetol. 2021;20(1):15.

106.  Lu J, Ma X, Shen Y, Wu Q, Wang R, Zhang L, et al. Time in Range Is Associated with Carotid Intima-Media Thickness in Type 2 Diabetes. Diabetes Technol Ther. 2019.

107.  Zhang C, Tang M, Lu X, Zhou Y, Zhao W, Liu Y, et al. Relationship of ankle-brachial index, vibration perception threshold, and current perception threshold to glycemic variability in type 2 diabetes. Medicine (Baltimore). 2020;99(12):e19374.

108.  Nakagomi A, Sunami Y, Okada S, Ohno Y, Shoji T, Fujisawa T, et al. Association between 1-h post-load plasma glucose levels and arterial stiffness in normotensive subjects with normal glucose tolerance. Diab Vasc Dis Res. 2018;15(1):39-45.

109.  Sciacqua A, Maio R, Miceli S, Pascale A, Carullo G, Grillo N, et al. Association between one-hour post-load plasma glucose levels and vascular stiffness in essential hypertension. PLoS One. 2012;7(9):e44470.

110. Bergman M, Jagannathan R, Buysschaert M, Pareek M, Olsen MH, Nilsson PM, et al. Lessons learned from the 1-hour post-load glucose level during OGTT: Current screening recommendations for dysglycaemia should be revised. Diabetes Metab Res Rev. 2018;34(5):e2992.

111. Siegelaar SE, Kerr L, Jacober SJ, Devries JH. A decrease in glucose variability does not reduce cardiovascular event rates in type 2 diabetic patients after acute myocardial infarction: a reanalysis of the HEART2D study. Diabetes Care. 2011;34(4):855-7.

112. Nguyen TT, Kreis AJ, Kawasaki R, Wang JJ, Seifert BU, Vilser W, et al. Reproducibility of the retinal vascular response to flicker light in Asians. Curr Eye Res. 2009;34(12):1082-8.

113. Agarwal SC, Allen J, Murray A, Purcell IF. Comparative reproducibility of dermal microvascular blood flow changes in response to acetylcholine iontophoresis, hyperthermia and reactive hyperaemia. Physiol Meas. 2010;31(1):1-11.

114. De Clerck EE, Schouten JS, Berendschot TT, Kessels AG, Nuijts RM, Beckers HJ, et al. New ophthalmologic imaging techniques for detection and monitoring of neurodegenerative changes in diabetes: a systematic review. Lancet Diabetes Endocrinol. 2015;3(8):653-63.

115. Zakin E, Abrams R, Simpson DM. Diabetic Neuropathy. Semin Neurol. 2019;39(5):560-9.

116. Srinivasan S, Dehghani C, Pritchard N, Edwards K, Russell AW, Malik RA, et al. Optical coherence tomography predicts 4-year incident diabetic neuropathy. Ophthalmic Physiol Opt. 2017;37(4):451-9.

117. Brownlee M. Biochemistry and molecular cell biology of diabetic complications. Nature. 2001;414(6865):813-20.

118. Yu DY, Cringle SJ, Balaratnasingam C, Morgan WH, Yu PK, Su EN. Retinal ganglion cells: Energetics, compartmentation, axonal transport, cytoskeletons and vulnerability. Prog Retin Eye Res. 2013;36:217-46.

119. Khawaja AP, Chua S, Hysi PG, Georgoulas S, Currant H, Fitzgerald TW, et al. Comparison of Associations with Different Macular Inner Retinal Thickness Parameters in a Large Cohort: The UK Biobank. Ophthalmology. 2020;127(1):62-71.

120. Huru J, Leiviska I, Saarela V, Liinamaa MJ. Prediabetes influences the structure of the macula: thinning of the macula in the Northern Finland Birth Cohort. Br J Ophthalmol. 2020.

121. Ho H, Tham YC, Chee ML, Shi Y, Tan NYQ, Wong KH, et al. Retinal Nerve Fiber Layer Thickness in a Multiethnic Normal Asian Population: The Singapore Epidemiology of Eye Diseases Study. Ophthalmology. 2019;126(5):702-11.

122. Fujiwara K, Yasuda M, Hata J, Hirakawa Y, Hashimoto S, Ueda E, et al. Glucose Tolerance Levels and Circumpapillary Retinal Nerve Fiber Layer Thickness in a General Japanese Population: The Hisayama Study. Am J Ophthalmol. 2019;205:140-6.

123. Duan XR, Liang YB, Friedman DS, Sun LP, Wong TY, Tao QS, et al. Normal macular thickness measurements using optical coherence tomography in healthy eyes of adult Chinese persons: the Handan Eye Study. Ophthalmology. 2010;117(8):1585-94.

124. De Clerck EEB, Schouten J, Berendschot T, Goezinne F, Dagnelie PC, Schaper NC, et al. Macular thinning in prediabetes or type 2 diabetes without diabetic retinopathy: the Maastricht Study. Acta Ophthalmol. 2018;96(2):174-82.

125. De Clerck EEB, Schouten JSAG, Berendschot TTJM, Beckers HJM, Schaper NC, Schram MT, et al. Loss of Temporal Peripapillary Retinal Nerve Fibers in Prediabetes or Type 2 Diabetes Without Diabetic Retinopathy: The Maastricht Study. Invest Ophth Vis Sci. 2017;58(2).

126. Coopmans C, Zhou TL, Henry RMA, Heijman J, Schaper NC, Koster A, et al. Both Prediabetes and Type 2 Diabetes Are Associated With Lower Heart Rate Variability: The Maastricht Study. Diabetes Care. 2020;43(5):1126-33.

127. Vergoossen LW, Schram MT, de Jong JJ, Stehouwer CD, Schaper NC, Henry RM, et al. White Matter Connectivity Abnormalities in Prediabetes and Type 2 Diabetes: The Maastricht Study. Diabetes Care. 2020;43(1):201-8.

128. van der Velde J, Koster A, Strotmeyer ES, Mess WH, Hilkman D, Reulen JPH, et al. Cardiometabolic risk factors as determinants of peripheral nerve function: the Maastricht Study. Diabetologia. 2020;63(8):1648-58.

9

129. van Agtmaal MJM, Houben A, de Wit V, Henry RMA, Schaper NC, Dagnelie PC, et al. Prediabetes Is Associated With Structural Brain Abnormalities: The Maastricht Study. Diabetes Care. 2018;41(12):2535-43.

130. Sorensen BM, Houben AJ, Berendschot TT, Schouten JS, Kroon AA, van der Kallen CJ, et al. Prediabetes and Type 2 Diabetes Are Associated With Generalized Microvascular Dysfunction: The Maastricht Study. Circulation. 2016;134(18):1339-52.

131. Li W, Schram MT, Berendschot T, Webers CAB, Kroon AA, van der Kallen CJH, et al. Type 2 diabetes and HbA1c are independently associated with wider retinal arterioles: the Maastricht study. Diabetologia. 2020;63(7):1408-17.

132. Haffner SM, Stern MP, Hazuda HP, Mitchell BD, Patterson JK. Cardiovascular risk factors in confirmed prediabetic individuals. Does the clock for coronary heart disease start ticking before the onset of clinical diabetes? JAMA. 1990;263(21):2893-8.

133. Šoupal J, Škrha Jr J, Fajmon M, Horová E, Mráz M, Škrha J, et al. Glycemic variability is higher in type 1 diabetes patients with microvascular complications irrespective of glycemic control. Diabetes technology & therapeutics. 2014;16(4):198-203.

134. Sartore G, Chilelli NC, Burlina S, Lapolla A. Association between glucose variability as assessed by continuous glucose monitoring (CGM) and diabetic retinopathy in type 1 and type 2 diabetes. Acta diabetologica. 2013;50(3):437-42.

135. Hu Y-m, Zhao L-h, Zhang X-l, Cai H-l, Huang H-y, Xu F, et al. Association of glycaemic variability evaluated by continuous glucose monitoring with diabetic peripheral neuropathy in type 2 diabetic patients. Endocrine. 2018;60(2):292-300.

136. Xu F, Zhao L-h, Su J-b, Chen T, Wang X-q, Chen J-f, et al. The relationship between glycemic variability and diabetic peripheral neuropathy in type 2 diabetes with well-controlled HbA1c. Diabetology & metabolic syndrome. 2014;6(1):139.

137. Jin SM, Kim TH, Oh S, Baek J, Joung J, Park SM, et al. Association between the extent of urinary albumin excretion and glycaemic variability indices measured by continuous glucose monitoring. Diabetic Medicine. 2015;32(2):274-9.

138. Hou X, Wang C, Wang S, Yang W, Ma Z, Wang Y, et al. Fluctuation between fasting and 2-H postload glucose state is associated with glomerular hyperfiltration in newly diagnosed diabetes patients with HbA1c< 7%. PloS one. 2014;9(10).

139. Siegelaar SE, Kilpatrick ES, Rigby AS, Atkin SL, Hoekstra JB, Devries JH. Glucose variability does not contribute to the development of peripheral and autonomic neuropathy in type 1 diabetes: data from the DCCT. Diabetologia. 2009;52(10):2229-32.

140. Kilpatrick ES, Rigby AS, Atkin SL. Effect of glucose variability on the long-term risk of microvascular complications in type 1 diabetes. Diabetes Care. 2009;32(10):1901-3.

141. Rawlings AM, Sharrett AR, Mosley TH, Ballew SH, Deal JA, Selvin E. Glucose peaks and the risk of dementia and 20-year cognitive decline. Diabetes Care. 2017;40(7):879-86.

142. Service FJ. Glucose variability. Diabetes. 2013;62(5):1398-404.

143. Tomlinson DR, Gardiner NJ. Glucose neurotoxicity. Nat Rev Neurosci. 2008;9(1):36-45.

144. Schalkwijk CG, Stehouwer CDA. Methylglyoxal, a Highly Reactive Dicarbonyl Compound, in Diabetes, Its Vascular Complications, and Other Age-Related Diseases. Physiol Rev. 2020;100(1):407-61.

145. Russo VC, Higgins S, Werther GA, Cameron FJ. Effects of fluctuating glucose levels on neuronal cells in vitro. Neurochemical research. 2012;37(8):1768-82.

146. Quincozes-Santos A, Bobermin LD, de Assis AM, Goncalves CA, Souza DO. Fluctuations in glucose levels induce glial toxicity with glutamatergic, oxidative and inflammatory implications. Biochim Biophys Acta Mol Basis Dis. 2017;1863(1):1-14.

147. Ko F, Muthy ZA, Gallacher J, Sudlow C, Rees G, Yang Q, et al. Association of Retinal Nerve Fiber Layer Thinning With Current and Future Cognitive Decline: A Study Using Optical Coherence Tomography. JAMA Neurol. 2018;75(10):1198-205.

148. Mutlu U, Colijn JM, Ikram MA, Bonnemaijer PWM, Licher S, Wolters FJ, et al. Association of Retinal Neurodegeneration on Optical Coherence Tomography With Dementia: A Population-Based Study. JAMA Neurol. 2018;75(10):1256-63.

149. Piconi L, Quagliaro L, Assaloni R, Da Ros R, Maier A, Zuodar G, et al. Constant and intermittent high glucose enhances endothelial cell apoptosis through mitochondrial superoxide overproduction. Diabetes Metab Res Rev. 2006;22(3):198-203.

150. van Sloten TT, Protogerou AD, Henry RM, Schram MT, Launer LJ, Stehouwer CD. Association between arterial stiffness, cerebral small vessel disease and cognitive impairment: A systematic review and meta-analysis. Neurosci Biobehav Rev. 2015;53:121-30.

151. van Sloten TT, Sedaghat S, Carnethon MR, Launer LJ, Stehouwer CDA. Cerebral microvascular complications of type 2 diabetes: stroke, cognitive dysfunction, and depression. Lancet Diabetes Endocrinol. 2020;8(4):325-36.

152. Biessels GJ, Despa F. Cognitive decline and dementia in diabetes mellitus: mechanisms and clinical implications. Nature Reviews Endocrinology. 2018;14(10):591-604.

153. Mooy JM, Grootenhuis PA, de Vries H, Kostense PJ, Popp-Snijders C, Bouter LM, et al. Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. Diabetologia. 1996;39(3):298-305.

154. Ucar F, Erden G, Ginis Z, Ozturk G, Sezer S, Gurler M, et al. Estimation of biological variation and reference change value of glycated hemoglobin (HbA(1c)) when two analytical methods are used. Clin Biochem. 2013;46(15):1548-53.

155. Meerwaldt R, Graaff R, Oomen PHN, Links TP, Jager JJ, Alderson NL, et al. Simple non-invasive assessment of advanced glycation endproduct accumulation. Diabetologia. 2004;47(7):1324-30.

156. Spauwen PJ, van Eupen MG, Kohler S, Stehouwer CD, Verhey FR, van der Kallen CJ, et al. Associations of advanced glycation end-products with cognitive functions in individuals with and without type 2 diabetes: the maastricht study. J Clin Endocrinol Metab. 2015;100(3):951-60.

157. Willekes C, Hoeks AP, Bots ML, Brands PJ, Willigers JM, Reneman RS. Evaluation of off-line automated intima-media thickness detection of the common carotid artery based on M-line signal processing. Ultrasound Med Biol. 1999;25(1):57-64.

158. Reneman RS, Meinders JM, Hoeks AP. Non-invasive ultrasound in arterial wall dynamics in humans: what have we learned and what remains to be solved. Eur Heart J. 2005;26(10):960-6.

159. Hong JT, Sung KR, Cho JW, Yun SC, Kang SY, Kook MS. Retinal nerve fiber layer measurement variability with spectral domain optical coherence tomography. Korean J Ophthalmol. 2012;26(1):32-8.

160. Van der Elst W, van Boxtel MP, van Breukelen GJ, Jolles J. Rey's verbal learning test: normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. J Int Neuropsychol Soc. 2005;11(3):290-302.

161. Van der Elst W, Van Boxtel MP, Van Breukelen GJ, Jolles J. The Stroop color-word test: influence of age, sex, and education; and normative data for a large sample across the adult age range. Assessment. 2006;13(1):62-79.

162. Van der Elst W, Van Boxtel MP, Van Breukelen GJ, Jolles J. The Concept Shifting Test: adult normative data. Psychol Assess. 2006;18(4):424-32.

163. Van der Elst W, Van Boxtel MP, Van Breukelen GJ, Jolles J. The Letter Digit Substitution Test: normative data for 1,858 healthy participants aged 24-81 from the Maastricht Aging Study (MAAS): influence of age, education, and sex. J Clin Exp Neuropsychol. 2006;28(6):998-1009.

164. Song MK, Lin FC, Ward SE, Fine JP. Composite variables: when and how. Nurs Res. 2013;62(1):45-9.

165. Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Res. 2017;6:926.

166. Hekler EB, Buman MP, Haskell WL, Conway TL, Cain KL, Sallis JF, et al. Reliability and validity of CHAMPS self-reported sedentary-to-vigorous intensity physical activity in older adults. J Phys Act Health. 2012;9(2):225-36.

167. Coughlin SS. Recall bias in epidemiologic studies. J Clin Epidemiol. 1990;43(1):87-91.

168. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009;20(4):488-95.

9

169. Martens RJ, Kooman JP, Stehouwer CD, Dagnelie PC, van der Kallen CJ, Koster A, et al. Estimated GFR, Albuminuria, and Cognitive Performance: The Maastricht Study. Am J Kidney Dis. 2017;69(2):179-91.

170. Hutcheson GD, Sofroniou N. The Multivariate Social Scientist. London, UK: SAG E Publications Ltd; 1999.

171. Freund RJ, Wilson WJ, Sa P. Regression Analysis: Statistical Modeling of a Response Variable. 2nd ed. Amsterdam, the Netherlands: Academic Press; 2006.

172. Wood A, O'Neal D, Furler J, Ekinci EI. Continuous glucose monitoring: a review of the evidence, opportunities for future use and ongoing challenges. Intern Med J. 2018;48(5):499-508.

173. Taubes G. Epidemiology faces its limits. Science. 1995;269(5221):164-9.

174. Curran-Everett D. Explorations in statistics: the analysis of ratios and normalized data. Adv Physiol Educ. 2013;37(3):213-9.

175. Rodbard D. Glycemic variability: challenges in interpretation. Diabetes Technol Ther. 2015;17(6):370-2.

176. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carre G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography. 2013;36(1):27-46.

177. Altman DG, Bland JM. Missing data. BMJ. 2007;334(7590):424.

178. Bland JM, Altman DG. Correlation in restricted ranges of data. BMJ. 2011;342:d556.

179. Muhammad IF, Borne Y, Ostling G, Kennback C, Gottsater M, Persson M, et al. Arterial Stiffness and Incidence of Diabetes: A Population-Based Cohort Study. Diabetes Care. 2017;40(12):1739-45.

180. Althouse AD. Adjust for Multiple Comparisons? It's Not That Simple. Ann Thorac Surg. 2016;101(5):1644-5.

181. Rothman KJ. No Adjustments Are Needed for Multiple Comparisons. Epidemiology. 1990;1(1):43-6.

182. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016;31(4):337-50.

183. Kramer CK, Choi H, Zinman B, Retnakaran R. Glycemic variability in patients with early type 2 diabetes: the impact of improvement in beta-cell function. Diabetes Care. 2014;37(4):1116-23.

184. Blauw H, Onvlee AJ, Klaassen M, van Bon AC, DeVries JH. Fully Closed Loop Glucose Control With a Bihormonal Artificial Pancreas in Adults With Type 1 Diabetes: An Outpatient, Randomized, Crossover Trial. Diabetes Care. 2021.

# Chapter 10

Scientific and societal impact

This chapter reflects on the scientific and societal impact of this thesis by identifying— in the context of the research field and current clinical practice—the different short- and long-term contributions of its findings.

# Research field and main objectives of this thesis

Diabetes mellitus is a metabolic disease that is characterized by elevated blood glucose values[1]. At present, approximately 463 million individuals are affected by diabetes worldwide; a figure that is expected to rise to 700 million by 2045[2,3]. This is problematic because diabetes lowers life expectancy[4-6] and strongly increases the chances of several diseases, including cardiovascular disease[7], eye, kidney, and nerve disease[8,9], and dementia[10]. More and more studies show that having prediabetes, which is the stage between normal glucose metabolism and diabetes, is also unhealthy[9-12]. This is concerning because the number of individuals with prediabetes is increasing worldwide as well[13,14].

Diabetes and its related diseases have a negative effect on quality of life[15] and represent a large share of the global health spending[16,17]. Prevention of these diseases will, thus, benefit the individual as well as society. However, this requires knowledge of the ways through which diabetes and prediabetes can cause disease. While decades of stellar research have led to great insights, and consequently have improved treatment and prevention[18], much is still unclear. Relatively recently, daily glucose fluctuations, also known as daily glucose variability, were suggested to be harmful[19]. Hence, a main objective of this thesis was to study whether daily glucose variability is indeed related to cardiovascular disease, eye, kidney, and nerve disease, and cognitive decline.

At the start of this century, a new device was introduced to improve diabetes care[20]. This device enables continuous glucose monitoring (CGM) and, thus, is able to record hundreds of glucose values per day. This allows for detailed study of daily glucose patterns. Hence, it is a device that is particularly suited to assess the degree of daily glucose variability[21]. Partly because of this, it is increasingly being used in scientific research and clinical practice. Studies have shown that use of CGM can improve diabetes care[20], and that the specific values calculated with CGM, in particular the intuitive index time in range, are important to individuals with diabetes[22]. However, several important aspects of CGM and the measurement of daily glucose variability have been insufficiently studied. Therefore, a main objective of this thesis was to do so in order to improve both future studies and clinical practice.

# Relevance of the key findings

*Continuous glucose monitoring and daily glucose variability measurement*

In this thesis, we studied how many days are needed to get a reliable CGM recording. This is relevant because CGM is quite expensive and undergoing it can be a burden[23],

**10**

especially when daily fingerpricks are required to calibrate the device[20]. Regarding short-term scientific value, knowledge of the minimum number of days needed can prevent both too short or too long recording periods, which would enhance measurement quality and limit participant strain in future studies, respectively. This information may also be used to reduce such burden in the clinical care of individuals with diabetes who intermittently wear a CGM. We believe that these findings can be used to improve the next *International Consensus Report* on *Clinical Targets for CGM Data Interpretation*[21].

CGM is also used in closed-loop insulin delivery systems, which can dose insulin based on the glucose values measured with CGM[24,25]. These systems are a very promising option to improve diabetes care in the near future[26-29]. These devices may be improved by addressing certain inherent shortcomings related to CGM, such as sensor delay and brief periods of sensor malfunction[24, 30, 31]. In this thesis, we developed a model to predict future glucose values in order to improve the CGM part of closed-loop insulin delivery systems. Our model was able to accurately and safely predict glucose values at 15- and 60-minute intervals, which could be useful in case of short periods of CGM malfunction. Our aim is to explore the possibility of further optimizing these prediction models in cooperation with companies that are specialized in diabetes care.

A large number of values can be measured with CGM, such as the average glucose and degree of daily glucose variability[23]. In this thesis, we studied, using data of participants with normal glucose metabolism, what the normal ranges of these values are. Interestingly, the daily glucose variability target that is currently recommended by the *International Consensus Report* lies much higher than the normal values found[21]. Future research should assess whether the target values should be lowered to more reflect the normal values by establishing whether daily glucose variability is harmful, and if so, at what point it becomes too damaging of unsafe.

In this thesis, we also investigated new ways to measure the amount of daily glucose variability. First, we found that the 'glucotype'[32], which garnered substantial scientific attention[33-36], did not live up to the expectation when used in other studies than the original one. Critical assessment of reproducibility is an important but often overlooked aspect of science[37]. Second, we showed that the incremental glucose peak during an oral glucose tolerance test (OGTT) as well as the 1-hour OGTT values can be used as measures of daily glucose variability, albeit predominantly in a research setting. This has short-term scientific value because studies that have such data available can immediately start contributing to the study of daily glucose variability. In addition, it puts the results of previous studies on the 1-hour OGTT value in a new light[38].

In this thesis, we also extensively reflected on the methods and statistics used in all of its chapters with the aim of providing insights that can help improve future research. For example, we showed that researchers should be especially on their guard for the statistical phenomenon called multicollinearity, which occurs due to a strong

relationship between two or more measures[39], when studying the consequences of daily glucose variability.

### The consequences of daily glucose variability

In this thesis, we have studied whether daily glucose variability is associated with several measures that precede actual diseases. We show that higher glucose variability is related to higher stiffness of the aorta, which reflects cardiovascular disease and mortality, to lower thickness of the nerves in the eye, which reflects neurodegeneration, and to worse cognitive performance, which reflects cognitive decline. Our findings thereby add to growing evidence that lowering of daily glucose variability (e.g., through lifestyle interventions or specific therapeutic regimens) could be important to prevent disease[40,41]. However, it is too early to reach such definite conclusions based on our findings. Namely, we assessed precursors of disease rather than actual diseases. In addition, we studied the relationships at one point in time, which makes it difficult to establish cause and effect[42]. As such, the short-term merit of our research lies in the fact that we provide directions for future research on the consequences of daily glucose variability. In the long run—especially if studies convincingly show that daily glucose variability causes future disease and that its treatment can prevent such disease—we hope to be able to state that we have made a meaningful contribution to diabetes care with our initial studies on the effects of daily glucose variability.

### Other key findings

In this thesis, we additionally found that prediabetes, diabetes, and measures of average glucose are also related to thickness of the nerves in the eye. This can have several implications for clinical practice. Measurement of the nerves in the eye could help select individuals who are at risk for developing eye or nerve disease. This would be feasible clinically because such measurement is non-invasive, inexpensive, and easy[43]. In addition, it further indicates early lowering of blood glucose values, possibly already in prediabetes, is key in the early prevention of eye and nerve disease.

## Knowledge dissemination to target groups

The impact of this thesis is primarily scientific. Namely, we report multiple ways to improve the quality of scientific research into CGM and daily glucose variability. As such, the scientific community is the main target group of this thesis. Publication of the results in scientific journals is one of the main ways to inform this community. In addition, the results of this thesis have been presented at several national and international conferences (see *Curriculum vitae*). As a result, it seems that these findings are starting to get noticed. For example, the incremental glucose peak has

10

recently received attention from one of the experts in the glucose variability field[40]. Further, our findings on CGM recording period reliability have been used by other authors to justify their recording period length[44]. In the end, the aim of scientific research is to improve society. If this thesis' suggestions on research quality continue to be applied in the coming years, this thesis will indirectly, positively impact society in the near future.

To a lesser extent, this thesis also has a direct societal impact, predominantly related to healthcare. Namely, certain key findings can have short- and long-term impact on diabetes care, such as reducing CGM burden and altering future therapy. As such, individuals with diabetes and prediabetes, as well as their relatives, are a target group of this thesis. While most of the findings of the current thesis are insufficient to alter the diabetes care of tomorrow, they do make a contribution to a field that can potentially improve the lives of a very large group of people (approximately 700 million in 2045)[3]. At present, it is debatable whether this target group would profit from being extensively updated on the current results. Regardless, the findings of this thesis have been presented in layman's terms on two separate occasions as part of the popular public symposium that is organized annually for the interested participants of The Maastricht Study.

# References

1.   American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S66-S76.

2.   Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol. 2018;14(2):88-98.

3.   Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. Diabetes Res Clin Pract. 2019;157:107843.

4.   Preis SR, Hwang SJ, Coady S, Pencina MJ, D'Agostino RB, Sr., Savage PJ, et al. Trends in all-cause and cardiovascular disease mortality among women and men with and without diabetes mellitus in the Framingham Heart Study, 1950 to 2005. Circulation. 2009;119(13):1728-35.

5.   Taylor KS, Heneghan CJ, Farmer AJ, Fuller AM, Adler AI, Aronson JK, et al. All-cause and cardiovascular mortality in middle-aged people with type 2 diabetes compared with people without diabetes in a large U.K. primary care database. Diabetes Care. 2013;36(8):2366-71.

6.   Bragg F, Holmes MV, Iona A, Guo Y, Du H, Chen Y, et al. Association Between Diabetes and Cause-Specific Mortality in Rural and Urban Areas of China. JAMA. 2017;317(3):280-9.

7.   Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical Update: Cardiovascular Disease in Diabetes Mellitus: Atherosclerotic Cardiovascular Disease and Heart Failure in Type 2 Diabetes Mellitus - Mechanisms, Management, and Clinical Considerations. Circulation. 2016;133(24): 2459-502.

8.   American Diabetes A. 11. Microvascular Complications and Foot Care: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Suppl 1):S135-S51.

9.   Stehouwer CDA. Microvascular Dysfunction and Hyperglycemia: A Vicious Cycle With Widespread Consequences. Diabetes. 2018;67(9):1729-41.

10.  Xue M, Xu W, Ou YN, Cao XP, Tan MS, Tan L, et al. Diabetes mellitus and risks of cognitive impairment and dementia: A systematic review and meta-analysis of 144 prospective studies. Ageing Res Rev. 2019;55:100944.

11.  Cai X, Zhang Y, Li M, Wu JH, Mai L, Li J, et al. Association between prediabetes and risk of all cause mortality and cardiovascular disease: updated meta-analysis. BMJ. 2020;370:m2297.

12.  Wong MS, Gu K, Heng D, Chew SK, Chew LS, Tai ES. The Singapore impaired glucose tolerance follow-up study: does the ticking clock go backward as well as forward? Diabetes Care. 2003;26(11):3024-30.

13.  Bullard KM, Saydah SH, Imperatore G, Cowie CC, Gregg EW, Geiss LS, et al. Secular changes in U.S. Prediabetes prevalence defined by hemoglobin A1c and fasting plasma glucose: National Health and Nutrition Examination Surveys, 1999-2010. Diabetes Care. 2013;36(8):2286-93.

14.  Mainous AG, 3rd, Tanner RJ, Baker R, Zayas CE, Harle CA. Prevalence of prediabetes in England from 2003 to 2011: population-based, cross-sectional study. BMJ Open. 2014;4(6):e005002.

15.  Jing X, Chen J, Dong Y, Han D, Zhao H, Wang X, et al. Related factors of quality of life of type 2 diabetes patients: a systematic review and meta-analysis. Health Qual Life Outcomes. 2018;16(1):189.

16.  Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Barnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. Lancet Diabetes Endocrinol. 2017;5(6):423-30.

17.  Bommer C, Sagalova V, Heesemann E, Manne-Goehler J, Atun R, Barnighausen T, et al. Global Economic Burden of Diabetes in Adults: Projections From 2015 to 2030. Diabetes Care. 2018;41(5):963-70.

18.  Zinman B, Skyler JS, Riddle MC, Ferrannini E. Diabetes Research and Care Through the Ages. Diabetes Care. 2017;40(10):1302-13.

19.  Brownlee M, Hirsch IB. Glycemic variability: a hemoglobin A1c-independent risk factor for diabetic complications. JAMA. 2006;295(14):1707-8.

10

20. Lin R, Brown F, James S, Jones J, Ekinci E. Continuous glucose monitoring: A review of the evidence in type 1 and 2 diabetes mellitus. Diabet Med. 2021:e14528.

21. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care. 2019;42(8):1593-603.

22. Advani A. Positioning time in range in diabetes management. Diabetologia. 2020;63(2):242-52.

23. Rodbard D. The challenges of measuring glycemic variability. J Diabetes Sci Technol. 2012;6(3):712-5.

24. Cobelli C, Renard E, Kovatchev B. Artificial pancreas: past, present, future. Diabetes. 2011;60(11):2672-82.

25. Bruttomesso D. Toward Automated Insulin Delivery. N Engl J Med. 2019;381(18):1774-5.

26. Kovatchev BP, Kollar L, Anderson SM, Barnett C, Breton MD, Carr K, et al. Evening and overnight closed-loop control versus 24/7 continuous closed-loop control for type 1 diabetes: a randomised crossover trial. Lancet Digit Health. 2020;2(2):e64-e73.

27. Blauw H, Onvlee AJ, Klaassen M, van Bon AC, DeVries JH. Fully Closed Loop Glucose Control With a Bihormonal Artificial Pancreas in Adults With Type 1 Diabetes: An Outpatient, Randomized, Crossover Trial. Diabetes Care. 2021.

28. Bergenstal RM, Nimri R, Beck RW, Criego A, Laffel L, Schatz D, et al. A comparison of two hybrid closed-loop systems in adolescents and young adults with type 1 diabetes (FLAIR): a multicentre, randomised, crossover trial. Lancet. 2021;397(10270):208-19.

29. Amadou C, Franc S, Benhamou PY, Lablanche S, Huneker E, Charpentier G, et al. Diabeloop DBLG1 Closed-Loop System Enables Patients With Type 1 Diabetes to Significantly Improve Their Glycemic Control in Real-Life Situations Without Serious Adverse Events: 6-Month Follow-up. Diabetes Care. 2021.

30. Rodbard D. Continuous Glucose Monitoring: A Review of Successes, Challenges, and Opportunities. Diabetes Technol Ther. 2016;18 Suppl 2:S3-S13.

31. Blauw H, Keith-Hynes P, Koops R, DeVries JH. A Review of Safety and Design Requirements of the Artificial Pancreas. Ann Biomed Eng. 2016;44(11):3158-72.

32. Hall H, Perelman D, Breschi A, Limcaoco P, Kellogg R, McLaughlin T, et al. Glucotypes reveal new patterns of glucose dysregulation. PLoS Biol. 2018;16(7):e2005143.

33. Zaccardi F, Khunti K. Glucose dysregulation phenotypes - time to improve outcomes. Nat Rev Endocrinol. 2018;14(11):632-3.

34. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.

35. Hamideh D, Arellano B, Topol EJ, Steinhubl SR. Your digital nutritionist. Lancet. 2019;393(10166):19.

36. Li J, Li X, Zhang S, Snyder M. Gene-Environment Interaction in the Era of Precision Medicine. Cell. 2019;177(1):38-44.

37. Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical research. Circ Res. 2015;116(1):116-26.

38. Bergman M, Jagannathan R, Buysschaert M, Pareek M, Olsen MH, Nilsson PM, et al. Lessons learned from the 1-hour post-load glucose level during OGTT: Current screening recommendations for dysglycaemia should be revised. Diabetes Metab Res Rev. 2018;34(5):e2992.

39. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. Epidemiology (Sunnyvale). 2016;6(2).

40. Ceriello A. Glucose Variability and Diabetic Complications: Is It Time to Treat? Diabetes Care. 2020;43(6):1169-71.

41. Zhou Z, Sun B, Huang S, Zhu C, Bian M. Glycemic variability: adverse clinical outcomes and how to improve it? Cardiovasc Diabetol. 2020;19(1):102.

42. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health. 2005;95 Suppl 1:S144-50.

43. Srinivasan S, Dehghani C, Pritchard N, Edwards K, Russell AW, Malik RA, et al. Optical coherence tomography predicts 4-year incident diabetic neuropathy. Ophthalmic Physiol Opt. 2017;37(4):451-9.

44. Mo Y, Ma X, Lu J, Shen Y, Wang Y, Zhang L, et al. Defining the target value of the coefficient of variation by continuous glucose monitoring in Chinese people with diabetes. J Diabetes Investig. 2020.

10

# Addendum

Nederlandse samenvatting
Dankwoord
Scientific output
Curriculum vitae

*Addendum*

# De methodologie en gevolgen van dagelijkse glucosevariabiliteit

### *Een zinvolle aanvulling op het gemiddelde?*

## Introductie

Diabetes mellitus, ook wel bekend als suikerziekte, treft op dit moment wereldwijd circa 463 miljoen mensen. Aangezien voorspeld wordt dat dit aantal de komende jaren nog sterk zal toenemen, gaat men ervan uit dat suikerziekte een nog groter gezondheidsprobleem zal gaan vormen. Onder diabetes worden verschillende stofwisselingsziekten geschaard die allen gekenmerkt worden door verhoogde suikerwaarden in het bloed, waarvan diabetes type 1 (5-10%) en diabetes type 2 (90-95%) het vaakst voorkomen. Diabetes verhoogt niet alleen het risico op vroegtijdig overlijden, maar maakt aangedane personen ook vatbaarder voor verscheidene chronische aandoeningen, waaronder hart- en vaatziekten, dementie en de klassieke complicaties aan het netvlies (retinopathie), de zenuwen (neuropathie) en nieren (nefropathie). Toenemend bewijs laat zien dat zelfs al het hebben van licht verhoogde bloedsuikerconcentraties (prediabetes) een verhoogd risico op dergelijke ziekten geeft.

Diabetes an sich heeft, evenals de daaraan gerelateerde aandoeningen, duidelijke negatieve gevolgen op individueel niveau door verlies van kwaliteit van leven. Bovendien ondervindt de maatschappij schade, onder meer door de grote economische effecten van diabetes. De behandeling van diabetes, die doorgaans zowel leefstijlinterventies als medicatie omvat, is in het bijzonderder gericht op preventie van het optreden of verergeren van complicaties. Het verlagen van bloedsuikerwaarden is een van de belangrijkste manieren om dit te bereiken. Het HbA$_{1c}$-gehalte in het bloed is al tientallen jaren de eerste keus om een indruk te krijgen van de gemiddelde bloedsuikerspiegel van de afgelopen twaalf weken. Het geeft echter onvoldoende informatie over de schommelingen in bloedsuiker (glucose) over de dag.

Deze bloedsuikerschommelingen, ook wel dagelijkse glucosevariabiliteit genoemd, zouden een beïnvloedbare risicofactor kunnen zijn voor bovengenoemde complicaties, onafhankelijk van de gemiddelde bloedsuikerwaarde. Aangezien het nauwkeurig meten van dagelijkse glucosevariabiliteit tot voor kort op grote schaal zeer beperkt mogelijk was, is er op dit moment een gebrek aan grote en degelijk uitgevoerde studies die de mogelijke gevolgen van dagelijkse glucosevariabiliteit onderzocht hebben. Met dit proefschrift streef ik ernaar om het onderzoek naar de mogelijke gevolgen van glucosevariabiliteit uit te breiden en aspireer ik om specifieke methodologische aspecten van het meten van dagelijkse glucosevariabiliteit en het middel dat daar normaliter voor gebruikt wordt, continue glucosemonitoring (CGM), onder de loep te nemen. Dit met het doel om op het gebied van diabeteszorg zowel de huidige klinische praktijk als toekomstig onderzoek te verbeteren.

A

## Belangrijkste bevindingen

### *Methodologie van continue glucosemonitoring en het meten van dagelijkse glucosevariabiliteit*

In de diabeteszorg wordt CGM in toenemende mate toegepast. De meeste van zulke monitors zijn in staat om minstens één week lang om de vijf minuten glucosewaardes te meten, waardoor een grote hoeveelheid gegevens over dagelijkse bloedsuikerprofielen verzameld kan worden. Deze bloedsuikergegevens kunnen vervolgens gebruikt worden om een aanzienlijk aantal indices te berekenen. In dit proefschrift is vooral gewerkt met de volgende indices: gemiddelde sensorglucose, standaarddeviatie, variatiecoëfficiënt en 'time in range' (TIR). Van deze indices zijn standaarddeviatie en variatiecoëfficiënt de meest gebruikte maten om dagelijkse glucosevariabiliteit uit de drukken. TIR is een relatief nieuwe index die het percentage van glucosewaarden binnen het streefgebied van 3,9 en 10 mmol/l weergeeft. TIR wordt bepaald door de gemiddelde glucosewaardes en de mate van dagelijkse glucosevariabiliteit.

In *hoofdstuk 2* hebben we in De Maastricht Studie onderzocht hoeveel dagen er nodig zijn om met CGM voldoende betrouwbaar de gemiddelde sensorglucose, standaarddeviatie en variatiecoëfficiënt te meten bij mensen met een normale suikerstofwisseling, prediabetes of diabetes type 2. Voorheen was dit namelijk alleen voor personen met diabetes type 1 onderzocht. Wij hebben gevonden dat afhankelijk van de gekozen index en de suikerstofwisselingstatus het benodigde aantal dagen neerkomt op één tot drie. Dit is een relevante bevinding vanuit het specifieke oogpunt van wetenschappelijke studies, omdat het laat zien dat het valide is om de meetperiode voor personen met een normale suikerstofwisseling, prediabetes of diabetes type 2 te verkorten tot drie functionele dagen om zo de last van studiedeelname te verlichten zonder afbreuk te doen aan de betrouwbaarheid van de meting.

In *hoofdstuk 2* hebben we tevens referentiewaarden vastgesteld voor de gemiddelde sensorglucose (5,03 - 6,69 mmol/l), standaarddeviatie (0,44 - 1,37 mmol/l) en variatiecoëfficiënt (7,74 - 22,45%) aan de hand van een groep deelnemers van De Maastricht Studie die allen een normale suikerstofwisseling hadden. De referentiewaarden voor deze indices waren tot nu toe zeer beperkt onderzocht bij personen zonder diabetes.

In *hoofdstuk 2* hebben we verder onderzocht of indices die het bloedsuikerprofiel tijdens een orale glucosetolerantietest (OGTT) beschrijven kunnen worden gebruikt om dagelijkse glucosevariabiliteit te kwantificeren. We hebben laten zien dat de incrementele glucosepiek (IGP), die wordt berekend door de nuchtere glucosewaarde af te trekken van de hoogste bloedsuikerwaarde tijdens de OGTT, een valide index van dagelijkse glucosevariabiliteit is en tevens de voorkeur verdient boven de absolute glucosepiekwaarde en de 1-uursglucosewaarde. Deze bevinding is met name relevant

voor onderzoeken die wel over OGTT- maar niet over CGM-data beschikken. Hoewel er in het geval van De Maastricht Studie wel CGM-gegevens verzameld zijn, heeft IGP als duidelijk voordeel dat deze waarden beschikbaar zijn voor een aanzienlijk groter aantal deelnemers (zie *De gevolgen van dagelijkse glucosevariabiliteit*).

In *hoofdstuk 3* hebben we onderzocht of op machine learning gebaseerde glucosewaardevoorspellingsmodellen gebruikt kunnen worden om bepaalde aspecten van de kunstmatige alvleesklier te verbeteren. De kunstmatige alvleesklier is een medisch hulpmiddel dat CGM, insulinepomp en controlealgoritme combineert om automatisch de bloedglucosespiegels te reguleren. De verbeterpunten die wij streefden aan te pakken zijn de circa 10 minuten die de glucosesensor op de suikerwaarden in het bloed achterloopt en de korte perioden van sensorstoring. We hebben 1,4 miljoen glucosedatapunten van deelnemers van De Maastricht Studie benut om te laten zien dat een met machine learning getraind model dat slechts 30 minuten aan voorafgaande CGM-gegevens gebruikt in staat is om nauwkeurig 15 en 60 minuten vooruit te voorspellen. Zo was meer dan 98% van de voorspellingen bij personen met diabetes type 2 klinisch zeer veilig. Ons voorspellingsmodel presteerde bovendien beter dan het model op basis van de grootste studie bij personen met diabetes type 2 tot nu toe.

In *hoofdstuk 3* hebben we verder onderzocht of ons voorspellingsmodel tevens bruikbaar was voor personen met diabetes type 1, aangezien zij op dit moment de voornaamste doelgroep voor de kunstmatige alvleesklier zijn en in het algemeen grotere, minder voorspelbare dagelijkse bloedsuikerschommelingen hebben. Hiervoor hebben we de openbaar beschikbare OhioT1DM-dataset gebruikt, die vaak wordt aangewend voor de evaluatie van modelprestaties. Hoewel we zagen dat onze modellen goed presteerden in mensen met diabetes type 1, was de prestatie wel minder goed dan voor de studiepopulatie van De Maastricht Studie. Desalniettemin hadden de modellen voldoende klinische veiligheid en presteerden ze net zo goed als de meest recente onderzoeken in het veld. We verwachten dat verdere technische optimalisatie van onze modellen in combinatie met een grotere populatie van individuen met diabetes type 1 de nauwkeurigheid en veiligheid nog verder zal verbeteren.

In *hoofdstuk 3* hebben we aanvullend onderzocht of de bloedsuikervoorspelling kan worden verbeterd door bewegingssensorgegevens aan de CGM-data toe te voegen. Het is namelijk bekend dat fysieke activiteit de bloedsuikerwaarden kan beïnvloeden. Onze modellen werden slechts in zeer geringe mate verbeterd door de informatie die de bewegingssensor bijdroeg, wat voor een deel verklaard kan worden door de reeds hoge accuratesse van de basale modellen.

In *hoofdstuk 4* hebben we, gebruik makende van gegevens van deelnemers uit De Maastricht Studie en de Deense PRE-D Trial, geëvalueerd wat de validiteit en generaliseerbaarheid zijn van het recent gepropageerde 'glucotype' als een CGM-index van dagelijkse glucosevariabiliteit. We toonden aan dat de verdeling van de drie

A

glucotypes (lage, matige of ernstige variabiliteit) waargenomen in de originele studie slecht generaliseert naar andere studiepopulaties. Bovendien hebben we aangetoond dat de meeste deelnemers een groot deel van de tijd doorbrengen in andere glucotypes dan het glucotype dat hun primair is toegewezen. Het belangrijkste is dat we, met behulp van gegevens van de bovengenoemde cohorten en de originele studie, hebben vastgesteld dat de classificatie van glucotypes bijna volledig wordt bepaald door gemiddelde sensorglucose. Daarmee betwisten we de bewering van de originele auteurs dat het glucotype een 'allesomvattendere maat is voor het patroon van bloedsuikerfluctuaties dan de standaard methoden die momenteel worden gebruikt'.

## *De mogelijke gevolgen van dagelijkse glucosevariabiliteit*

Dagelijkse glucosevariabiliteit werd vijftien jaar geleden geopperd als een onafhankelijke en potentieel beïnvloedbare risicofactor voor met diabetes samenhangende complicaties. Sindsdien heeft een relatief klein aantal onderzoeken onderzocht of dit het geval is voor hart- en vaatziekten, cognitieve achteruitgang en schade aan de allerkleinste bloedvaten en zenuwen. Het merendeel van deze onderzoeken heeft qua studieopzet duidelijk verbeterpunten. Deze betreffen onder meer de grootte van de onderzoekspopulaties (kleinere studies geven grotere kans op toevalstreffers, maar ook op het missen van relevante verbanden), het meenemen van specifieke risicofactoren in de analyses (het weglaten van belangrijke risicofactoren kan het onderzochte verband verstoren) en het meten van dagelijkse glucosevariabiliteit (weinig precies meten geeft zwakkere verbanden). De hiervoor genoemde beperkingen zijn overigens wel begrijpelijk als men bedenkt dat het tot voor kort een uitdaging was om op grote schaal CGM-metingen uit te voeren, gedeeltelijk vanwege de apparatuurskosten en deelnemerbelasting. Logischerwijs volgt hieruit dat studies die deelnemers niet op één vast moment bestudeerd hebben (cross-sectioneel onderzoek), maar over langere tijd opgevolgd hebben (prospectief onderzoek) nog schaarser zijn.

In dit proefschrift hebben we, in een grote, goed gekarakteriseerde onderzoeks-populatie (De Maastricht Studie), de cross-sectionele verbanden van dagelijkse glucosevariabiliteit, gemeten met CGM en OGTT, met meerdere 'voorlopers' van de bovengenoemde ziekten bestudeerd. Stijfheid van de grote lichaamsslagader (aorta) wordt bijvoorbeeld als een relevante voorloper van hart- en vaatziekten gezien en eerdere onderzoeken hebben laten zien dat de mate ervan ook samenhangt met de kans op overlijden. De hiervoor beschreven benadering heeft bepaalde voordelen. Zo worden verbanden eerder aangetoond als gebruik wordt gemaakt van continue (bijv. mate van aortastijfheid) in plaats van dichotome (bijv. hartinfarct) uitkomstmaten. Ook kan de benadering inzicht verschaffen in de onderliggende biologische processen. Het heeft echter ook bepaalde nadelen. Zo is het klinische belang veel duidelijker voor de verbanden met een ongunstige uitkomstmaat (bijv. hartinfarct) dan met voorlopers

ervan (bijv. mate van aortastijfheid). Aanvullend belemmert het cross-sectionele ontwerp van onze studies het inzicht in de oorzakelijkheid van de verbanden.

In *hoofdstukken 5 en 6* hebben wij dit onderzocht voor verschillende uitkomstmaten die als voorlopers van hart- en vaatziekten worden beschouwd. Wij hebben laten zien dat grotere dagelijkse glucosevariabiliteit (uitgedrukt als standaarddeviatie, variatiecoëfficiënt en IGP) en lagere TIR consistent samenhangen zijn met grotere aortastijfheid, onafhankelijk van demografische aspecten, cardiovasculaire risicofactoren en levensstijlfactoren. Dit kan erop wijzen dat dagelijkse glucosevariabiliteit een onafhankelijke risicofactor is voor hart- en vaatziekten. In tegenstelling tot de aortastijfheidsbevindingen, werden er geen consistente verbanden gevonden tussen dagelijkse glucosevariabiliteit en stijfheid, wanddikte en wandstress van de halsslagader. In *hoofdstuk 6* worden mogelijk aanwijzingen gezien voor een verband tussen hogere dagelijkse glucosevariabiliteit en een lagere enkel-arm-index (slechtere slagaderlijke doorbloeding van de benen ten gevolge van slagaderverkalking). Het deelnemersaantal van ca. 800 personen was echter niet hoog genoeg om een toevallig verband te verwerpen.

In *hoofdstukken 5 en 7* hebben wij de verbanden onderzocht tussen dagelijkse glucosevariabiliteit en de allerkleinste bloedvaten en zenuwen. In *hoofdstuk 5* hebben wij geen verbanden gevonden tussen dagelijkse glucosevariabiliteit en de functie van de allerkleinste bloedvaten in het oog en in de huid. In *hoofdstuk 7* hebben wij wel een consistent en onafhankelijk verband gevonden tussen hogere dagelijkse glucosevariabiliteit en lagere zenuwlaagdikte in het netvlies. Dit laatste kan erop wijzen dat dagelijkse glucosevariabiliteit een risicofactor is voor neurodegeneratie en daarmee ook voor retinopathie en neuropathie. In hoofdstuk 7 is tevens gevonden dat suikerstofwisselingsstatus en maten van gemiddelde glykemie samenhangen met de zenuwlaagdikte in het netvlies.

In *hoofdstuk 8* hebben wij het verband onderzocht tussen dagelijkse glucosevariabiliteit en het denkvermogen, beoordeeld met verschillende neuropsychologische testen. In dit hoofdstuk hebben wij in een groep van ruim 3000 deelnemers consistente verbanden gevonden tussen hogere IGP-waarden en lagere score op het algehele denkvermogen en het specifieke domein 'executief functioneren en aandacht' (de hogere controlefuncties van de hersenen), maar niet op de domeinen 'geheugen' en 'informatieverwerkingssnelheid'. Dit kan erop wijzen dat dagelijkse glucosevariabiliteit een onafhankelijke risicofactor is voor dementie. In de kleinere CGM-onderzoeksgroep (ca. 800 deelnemers) zijn er echter geen duidelijk verbanden gevonden tussen dagelijkse glucosevariabiliteit en denkvermogen, hoewel er wel een onafhankelijk verband werd gevonden voor lagere TIR en slechtere scores op 'executief functioneren en aandacht'.

A

## Conclusie

Concluderend biedt dit proefschrift verschillende aangrijpingspunten om het meten van dagelijkse glucosevariabiliteit (bij voorkeur met CGM) te verbeteren, niet alleen in wetenschappelijke studies, maar ook in de klinische praktijk. Dit door de betrouwbaarheid van registratieperiodes te evalueren, door op machinaal leren gebaseerde glucosevoorspellingsmodellen te onderzoeken, referentiewaarden voor CGM-afgeleide indices vast te stellen en de legitimiteit van nieuwe glucose-variabiliteitsmaten te beoordelen. Dit proefschrift geeft ook meer inzicht in de vraag of dagelijkse glucosevariabiliteit als onafhankelijke risicofactor van diabetescomplicaties kan worden beschouwd. Een dergelijke rol is het meest overtuigend voor grotere aortastijfheid, neurodegeneratie en slechtere cognitieve prestaties, aangezien de verbanden met deze uitkomsten overeind bleven na het meenemen van verscheidene potentieel verstorende factoren, waaronder leeftijd, geslacht, cardiovasculaire risicofactoren en gemiddelde glykemie, en consistent waren voor verschillende indices van dagelijkse glucosevariabiliteit.

Hoewel dit proefschrift erop gericht was om methodologische beperkingen van eerdere studies in het veld te ondervangen, kunnen bepaalde aspecten in de toekomst nog verder verbeterd worden. Aangezien de uitgevoerde onderzoeken allemaal cross-sectioneel waren, zijn prospectieve onderzoeken nodig om meer inzicht te krijgen in oorzaak en gevolg. Hoewel de deelnemersaantallen relatief groot waren in vergelijking met de meeste eerdere onderzoeken, zijn er toch nog grotere aantallen nodig om definitievere conclusies te trekken over de vraag of dagelijkse glucosevariabiliteit ook een onafhankelijke determinant zou kunnen zijn van onder andere slagaderverkalking en slechtere prestaties op andere domeinen van het denkvermogen.

Als onze bevindingen met prospectief onderzoek bevestigd worden, zou dit studies die zich richten op de gunstige effecten van gerichte verlaging van dagelijkse glucosevariabiliteit verder rechtvaardigen. In dat opzicht zou het relevant zijn om te onderzoeken welke factoren dagelijkse glucosevariabiliteit bepalen en wat hun relatieve bijdrage is. Alvleesklierfunctie, voeding en fysieke activiteit zullen naar verwachting belangrijke factoren zijn. Verder kunnen bepaalde soorten bloedsuikerverlagende medicijnen of specifieke therapeutische regimes verschillende effecten hebben op dagelijkse glucosevariabiliteit. Bovendien wordt verwacht dat een toenemend gebruik van CGM, met name van apparaten die de patiënt in een oogopslag de glucosewaarde tonen (flash glucose monitoring), en rapportage van intuïtieve indices als TIR, de behandeling en therapietrouw verder zullen optimaliseren. In het geval van bijna of volledig verlies van alvleesklierfunctie, is de kunstmatige alvleesklier waarschijnlijk de effectiefste manier om dagelijkse glucosevariabiliteit veilig te verminderen.

# Dankwoord

Het is volbracht, mijn proefschrift is klaar! Alle hoofdstukken staan in inkt op papier, omgeven door een strak ontworpen cover. Mijn promotietraject was een leerzame, vormende en veelbewogen periode die gekenmerkt werd door hard werken en erg fijne collegialiteit. Als allerlaatste heb ik dit dankwoord, misschien wel het vitaalste onderdeel van het proefschrift, geschreven. Met het zetten van deze afsluitende punten en komma's wil ik graag mijn collega's, vrienden en familie hartelijk bedanken voor hun belangrijke bijdrages aan de totstandkoming van dit proefschrift.

Om te beginnen wil ik mijn grote dank uitspreken aan de duizenden deelnemers van De Maastricht Studie, in het bijzonder aan de 900 proefpersonen die de continue glucosemonitor hebben willen dragen. Zonder de wetenschappelijke inzet en interesse van de studieparticipanten had dit proefschrift niet kunnen bestaan.

Vervolgens wil ik mijn promotieteam hartelijk bedanken voor het van begin af aan in mij gestelde vertrouwen en voor de zeer fijne en persoonlijke begeleiding van mijn onderzoek.

Mijn eerste promotor, prof.dr. C.D.A. Stehouwer, beste Coen. Er is al veel geschreven over je kritische blik, scherpe analytische vaardigheden en streven naar perfectie, karakterkenmerken die ik alleen maar beamen kan. Ook ik ben zeer onder de indruk van de snelheid waarmee je op mails en manuscripten reageert. Tijdens voortgangsgesprekken en schrijfsessies, maar zeker ook op basis van je geschreven commentaar, heb ik veel van je geleerd, zowel over de algehele structuur van een manuscript alsook over de specifieke grammaticale regels en taalnuances die het geheel naar een hoger niveau kunnen tillen. Ik wil afsluitend een paar aspecten benoemen die ik in andere dankwoorden onderbelicht zie: je gevoel voor humor en het grote verantwoordelijkheidsgevoel voor je werknemers.

Mijn tweede promotor, prof.dr. M.C.G.J. Brouwers, beste Martijn. Mijn eerste onderzoekservaringen heb ik bij jou opgedaan toen ik met mijn wetenschapsstage aansloot bij de HFI-studie van Nynke. Tijdens deze stage bood je me de gouden kans om te solliciteren op een nieuwe promotieplek bij De Maastricht Studie. Toen ik hiermee in 2016 begon, ontbrak de welverdiende 'prof.' nog in je rij met titels en was je dus nog 'slechts' mijn copromotor. Onze wekelijkse werkbesprekingen vormden de eerste instructieve gelegenheden om nieuwe projecten, data-aanvragen, analyses, resultaten en manuscripten te bediscussiëren. Ik heb veel respect voor je ambitie, inhoudelijke kennis en wetenschappelijke nieuwsgierigheid. Ondanks dat mijn onderzoek minder in jouw onderzoekslijn paste dan dat van je andere promovendi, heb ik daar qua begeleiding nooit iets van gemerkt. Ik kijk zeer uit naar je inauguratie.

Mijn derde promotor, prof.dr. N.C. Schaper, beste Nicolaas. Tijdens onze vaste overleggen over mijn onderzoek, maar zeker ook gedurende de vergaderingen van het

A

managementteam van De Maastricht Studie, heb ik je sterke creatieve, communicatieve en managementvaardigheden van dichtbij mee mogen maken. Je bent een verbinder, maar zeker ook iemand die kritisch en bekwaam doorpakt. Dit heb ik ook duidelijk meegekregen bij mijn participatiestage op afdeling B5 terwijl we de rondes liepen langs patiënten met een diabetische voet. Het is heel fijn dat je na je emeritaat als voorheen bij mijn onderzoek betrokken bent gebleven.

Geachte leden van de beoordelingscommissie, prof.dr. M.H. Prins, prof.dr.ir. E.E. Blaak, prof.dr. C.J.J. Tack, prof.dr. J.H. de Vries en dr. T.T. van Sloten, veel dank voor het lezen en beoordelen van mijn proefschrift, alsook voor uw aanwezigheid tijdens mijn verdediging.

De afgelopen jaren heb ik het genoegen gehad om onderdeel te zijn van een groot en hecht team. Een team dat ervoor zorgde dat ik iedere dag met veel plezier naar mijn werk ging. Graag wil ik jullie—in de volgorde vanaf mijn vaste bureau—bedanken.

Om te beginnen mijn eerste kamergenoten, Sytze en Tan Lai, allebei 'PhD extraordinaire'. Het is uitermate fijn dat ik altijd een beroep heb mogen doen op jullie extra jaar aan onderzoekservaring ten aanzien van data-aanvragen, statistiek en alle andere zaken die nodig zijn om uiteindelijk tot een fatsoenlijk manuscript te komen. Maar toegegeven, op onze werkkamer werd niet louter hard aan onze proefschriften gewerkt. Op scherpe taalgrappen, obscure verwijzingen en allerhande inside jokes deden we minstens net zo hard onze best. Ik wil jullie enorm bedanken voor alle gezellige momenten binnen en buiten De Maastricht Studie en niet in de minste plaats voor het feit dat jullie mij weer sportief hebben weten te activeren.

Sytze, dankzij onze sterk overlappende muzieksmaken hebben we, onder het noeste schrijf- en analysewerk, vele uren naar elkaars favoriete artiesten geluisterd, wat culmineerde in het gezamenlijk bezoek aan een explosief Rammsteinconcert. Je sportiviteit en fanatisme maken je een geduchte squashtegenstander. Ik heb er het volste vertrouwen in dat je dankzij je rustige, weloverwogen karakter en doelgerichte instelling een uitstekende radioloog zult worden.

Tan Lai, achter je ontspannen voorkomen gaat een zeer harde en slimme werker schuil. Desondanks ben je absoluut geen solist, je zet je in voor het collectief en bent altijd bereid om iemand te helpen. Ik heb ontzag voor wat je bereikt en vind het uitermate fijn om ook klinisch weer met je samen te kunnen werken. En dat is zeker niet alleen maar zo, omdat ik daarmee wederom een beroep kan doen op je extra jaar aan wijsheid en ervaring.

Tijdens het laatste jaar onderzoek deelden Sara en ik de werkkamer. Helaas moest er noodgedwongen een aanzienlijk deel van de tijd thuis worden gewerkt.

Sara, ik heb veel respect voor je arbeidsethos, optimisme en vermogen om werk en privé te combineren. Ik vond het heel prettig om met je samen te werken en hoop een vraagbaak voor je te hebben kunnen zijn, zoals mijn vorige kamergenoten dat voor mij waren.

Vervolgens is de aangrenzende promovendikamer aan de beurt.

Frank, je tomeloze energie en enthousiasme werken aanstekelijk. Ik vind het indrukwekkend dat het je gelukt is om, ondanks je vele activiteiten buiten De Maastricht Studie, een enorm en gedegen boekwerk af te leveren. Eén van onze hoofdstukken is hetzelfde. Ik heb onwijs genoten en veel geleerd van onze synergistische samenwerking en ben uiterst tevreden over het eindresultaat. Met nog meer plezier denk ik terug aan onze muzikale repetities en optredens.

Rianneke, als ik dit dankwoord na mijn eerste jaar had geschreven, was je pas in de volgende alinea aan bod gekomen. Echter heb je later de excellente keuze gemaakt om ook promotieonderzoek te gaan doen. Het was erg gezellig om op het werk vele koppen thee met je te delen en om buiten De Maastricht Studie samen met onze wederhelften danslessen bij Bernaards te volgen en tijdens spelletjesavonden elkaar te slim af te zijn.

April, hoewel je pas later onderdeel van De Maastricht Studie werd, voelt het alsof we gedurende mijn hele onderzoek samen hebben werkt. Je bent een collega op wie je kunt bouwen en het lijkt me een genoegen om in de kliniek ook met je samen te werken.

Het was altijd aangenaam om naar rechts te kijken, richting de kamers van de fameuze LOC ('leiding onderzoekscentrum'; een alternatieve uitleg van het acroniem is ook beschreven).

Carla, bedankt voor je betrokkenheid bij mijn manuscripten, maar bovenal voor de prettige en persoonlijke manier waarop je leidinggeeft aan het onderzoekscentrum. Als ik ergens vragen over had, klopte ik nooit tevergeefs bij je aan. Verder moet ik terugdenken aan de vele autoritten naar Brabant, de gezamenlijke wandelingen van en naar de MT-overleggen en het verzamelen van balen hooi (eerst maaien, schudden en duinen uiteraard).

Yvette, zonder jou zou De Maastricht Studie De Maastricht Studie niet zijn. Enerzijds heeft dit te maken met de uitstekende wijze waarop je het onderzoekscentrum draaiende houdt. Anderzijds ligt dit natuurlijk aan onze goede persoonlijke klik. Ik heb enorm genoten van je directe humor, onze gedeelde voorliefde voor schoenen en de vele donderdagmiddagborrels.

Nadine, bedankt voor je snelle en altijd degelijke regel- en planwerk.

De omvangrijkste werkkamer in de kantoorruimte van De Maastricht Studie huisvest niet voor niets de datamanagers.

A

Marion, Lisanne, Michiel en Anouk, zonder jullie zorgvuldige, ijverige en professionele werk zouden wij simpele promovendi niks kunnen beginnen. Het is enorm fijn dat jullie 'dataset iPro' meegenomen hebben bij het opschonen en opleveren van 'dataset 3'.

Tot slot, kom ik uit bij de vele onderzoekskamers, waar de onderzoeksassistenten het overgrote deel van het meetwerk verrichten.

Mijn V1-collega's, Ank, Brigitte, Carine, Else, Hannah, Ingeborg, Jos, Joséphine, Kayleigh, Laura, Niels, Regien en Roel, bedankt voor jullie hulp bij het aansluiten van de continue glucosemonitoren en voor de aangename gesprekken tussen de vele honderden bloedafnames door.

Niki, dank voor alle hulp tijdens mijn promotieonderzoek en voor je grote positivisme en oprechte interesse.

Ramona, enorm veel dank voor je hulp met het uploaden van de iPro-gegevens. Het is ons samen toch maar mooi gelukt om 900 sensoren onder de kraan schoon te schrobben.

Barbara, Brenda, Chantalle, Daniel, Eline, Evelien, Ineke, Iris, Jeroen, Lieke, Lina, Manon, Manuela, Mitch, Myrthe, Nikki, Paul, Paula, Robin en Veerle, dank voor de fijne collegialiteit.

Demi, Lisanne, Emma, Sadé, Rutger en Anne, bedankt dat jullie ervoor hebben gekozen om tijdens jullie studie bij te dragen aan mijn promotieonderzoek. Ik heb er veel van geleerd om jullie te begeleiden.

Aan de andere promovendi van De Maastricht Studie, Anouk, Ben, Charlotte, Cindy, Ellis, Jennifer, Jeroen, Jerremy, Laura, Marja, Marnix, Nathan, Shunxin, Steffi, Veronica en Wenjie, en de collega-promovendi van de zogeheten 'overkant', Amée, Armand, Evelien, Kim, Mathias, Mitchell, Myrthe, Nynke, Pomme en Ying, veel dank voor de gezelligheid en interessante discussies, zij het gewoon in Maastricht of tijdens één van de vele geslaagde congresbezoeken. Omar, many thanks for your enthusiastic and fruitful research stay at The Maastricht Study.

William, bedankt voor de fijne samenwerking bij ons gedeelde manuscript. Ook veel dank voor je hulp bij het vinden van oplossingen voor het multicollineariteitsprobleem. Ik heb enorm veel van je geleerd en zal immer onder de indruk zijn van je brede kennis en kunde op het gebied van verscheidene analysetechnieken, ingewikkelde scripts en machine learning. Helaas is ons stuk minder hoog geland dan we van tevoren ingeschat of gehoopt hadden. Desondanks ben ik zeer trots op wat we samen hebben weten neer te zetten.

Graag wil ik de andere coauteurs die geholpen hebben bij het schrijven van mijn manuscripten bedanken voor het uitgebreide commentaar en de fijne samenwerking.

Dr. M.M.J. van Greevenbroek, beste Marleen. Bedankt voor je grote methodologische expertise en immer scherpe feedback op mijn stukken, waarbij ik nadrukkelijk heb genoten van onze gedeelde affiniteit voor taal.

Dr. R.M.A. Henry, beste Ronald. In het bijzonder dank voor je rol bij de prille onderhandelingen met Medtronic. Zonder jouw standvastigheid was er waarschijnlijk geen geld voor mijn aanstelling, en daarmee mogelijk dus ook geen proefschrift, geweest.

Prof.dr. B.E. de Galan, beste Bastiaan. Je nauwe betrokkenheid bij twee van de hoofdstukken uit mijn thesis was inhoudelijk en tekstueel zeer waardevol. Ik kijk ernaar uit om de collaboratie met Medtronic samen met jou voort te zetten.

Dr. A. Hulman, dear Adam. We first met at the EASD in Berlin when you stylishly introduced yourself by pointing at the first reference under my poster. Thank you for the great collaboration that followed from your research visit to The Maastricht Study. Although the glycotype, unfortunately, turned out to be just a glucohype, it still resulted in an excellent formal comment in PLoS Biology. It was also great to meet up again during the EASD in Barcelona.

Prof.dr. D.R. Witte, beste Daniel. Dank voor de prettige samenwerking bij hoofdstuk 4 en de geboden mogelijkheid om een postdoc in Denemarken te doen.

Dr. S.J.R. Meex, beste Steven. Het project 'machine learning' was er echt een van de lange adem. Echter kijk ik vooral terug op de grote leerzaamheid van deze samenwerking, waarbij speciaal het grote belang van het 'storyboard' naar boven komt.

Dr. A. Koster, beste Annemarie. Dank voor je rustige en heldere uitleg bij data-opschoonprojecten en je constructieve commentaar op mijn manuscripten.

Dr. M.T. Schram, beste Miranda. Bedankt voor je zorgvuldige reacties op mijn stukken, specifiek bij onze samenwerking tijdens hoofdstuk 8.

Dr. S.J.P.M. Eussen, best Simone. Dank voor je inhoudelijke feedback en voor je belangrijke rol in het beschikbaar komen van de OGTT- en voedingsdagboekjesdata.

Dr. A. Wesselius, best Anke. Dank voor je geëngageerdheid bij mijn stukken. Het was bijzonder om te ontdekken dat we 'aangetrouwde' familie zijn.

De overige coauteurs, prof.dr. O. Bekers, dr. T.T.J.M. Berendschot, dr. M.P.J. van Boxtel, prof.dr. P.C. Dagnelie, dr. M.C.J.M. van Dongen, dr. A.J.H.M. Houben, dr. S. Köhler, prof.dr. A.A. Kroon, dr. K.D. Reesink, prof.dr. H.H.C.M. Savelberg, prof.dr. C.G. Schalkwijk, dr. J.S.A.G. Schouten en prof.dr. C.A.B. Webers, bedankt voor uw commentaar op en de participatie in onze gedeelde projecten.

I also would like to thank all foreign coauthors, Prof. K. Færch, Prof. M.E. Jørgensen, and Dr. D. Vistisen, for their collaboration on Chapter 4.

All Medtronic employees with whom I have had the pleasure of working, Odile O'Sullivan, Robert Vigerski, Ohad Cohen, Julien Da Silva, Javier Castenada, Raphael Schneider, Arcelia Arrieta, Eva Korthagen, Lisette van der Vight, John Shin, Remco

A

Lieve Granny, bedankt voor de oprechte betrokkenheid die ik vanaf mijn geboorte altijd heb mogen voelen.

Mitch, Grote Neef en collega-internist in opleiding, ik kijk ernaar uit om over enkele jaren samen met jou in het MUMC+ te werken.

Oom Peter, hartstikke bedankt voor het prachtige ontwerp van de kaft.

Lieve oma Mies, een heel aantal van mijn betere karaktereigenschappen heb ik aan jou te danken. Daarbij moet ik specifiek denken aan je grote zorgzaamheid, scherpe gevoel voor humor en kritische blik. Bedankt voor alle warmte die jij en opa mij altijd geboden hebben.

Lieve opa Jacques, tot mijn grote verdriet is het niet meer mogelijk om je op 20 oktober stralend van trots in de zaal van de aula te zien zitten. Het valt 'je Gouwe' zwaar om zijn grootste fan al bijna een jaar te moeten missen. Bedankt voor alles wat je voor mij betekent.

Lieve Meryl, ik vind het mooi dat we inmiddels niet louter broer en zus, maar ook collega-dokters zijn (al beperk ik me tot de enige diersoort die jij slechts indirect behandelt). Ik heb bewondering voor je positivisme, de manier waarop je je talenten benut en de sterke keuzes die je in je leven hebt gemaakt. Veel dank aan jou en Ferry voor de warme betrokkenheid, gezelligheid en—niet te vergeten—de muziek.

Mijn lieve ouders, jullie hebben me een geweldige, warme en liefdevolle opvoeding gegeven en zijn in velerlei opzichten een voortreffelijk voorbeeld voor me. Jullie bieden de solide basis die mij het vertrouwen geeft om mijn doelen in het leven na te streven. Bedankt dat jullie altijd voor me klaarstaan.

Allerliefste Ilse, iedere keer als ik tegen je zeg dat ik zonder jou niet had kunnen bereiken wat mij allemaal gelukt is, heb je de neiging dit direct te ontkrachten. Toch is het echt zo. Je bent voor mij het toonbeeld van een harde werker, vaardige dokter en uitstekende wetenschapper en sterkt mij buitengewoon met de geborgenheid, warmte en liefde die je me geeft. Dank dat je voor mij hebt gekozen. Dank voor al je hulp, steun en gezelligheid. Dank dat je het leven zo veel mooier maakt. Ik kijk enorm uit naar wat de toekomst ons nog allemaal gaat brengen.

A

# Scientific output

## *List of publications*

**Foreman YD**, van Doorn WPTM, Schaper NC, van Greevenbroek MMJ, van der Kallen CJH, Henry RMA, Koster A, Eussen SJPM, Wesselius A, Reesink KD, Schram MT, Dagnelie PC, Kroon AA, Brouwers MCGJ, Stehouwer CDA. *Greater daily glucose variability and lower time in range assessed with continuous glucose monitoring are associated with greater aortic stiffness: The Maastricht Study*. Diabetologia. 2021 Aug;64(8):1880-1892.

van Doorn WPTM, **Foreman YD**, Schaper NC, Savelberg HHCM, Koster A, van der Kallen CJH, Wesselius A, Schram MT, Henry RMA, Dagnelie PC, de Galan BE, Bekers O, Stehouwer CDA, Meex SJR, Brouwers MCGJ. *Machine learning-based glucose prediction with use of continuous glucose and physical activity monitoring data: The Maastricht Study*. PLoS One. 2021 Jun;16(6):e0253125.

Hulman A, **Foreman YD**, Brouwers MCGJ, Kroon AA, Reesink KD, Dagnelie PC, van der Kallen CJH, Greevenbroek MMJV, Færch K, Vistisen D, Jørgensen ME, Stehouwer CDA, Witte DR. *Towards precision medicine in diabetes? A critical review of glucotypes*. PLoS Biol. 2021 Mar;19(3):e3000890.

**Foreman YD**, Brouwers MCGJ, van der Kallen CJH, Pagen DME, van Greevenbroek MMJ, Henry RMA, Koster A, Wesselius A, Schaper NC, Stehouwer CDA. *Glucose Variability Assessed with Continuous Glucose Monitoring: Reliability, Reference Values, and Correlations with Established Glycemic Indices-The Maastricht Study*. Diabetes Technol Ther. 2020 May;22(5):395-403.

**Foreman YD**, Brouwers MCGJ, Berendschot TTJM, van Dongen MCJM, Eussen SJPM, van Greevenbroek MMJ, Henry RMA, Houben AJHM, van der Kallen CJH, Kroon AA, Reesink KD, Schram MT, Schaper NC, Stehouwer CDA. *The oral glucose tolerance test-derived incremental glucose peak is associated with greater arterial stiffness and maladaptive arterial remodeling: The Maastricht Study*. Cardiovasc Diabetol. 2019 Nov;18(1):152.

## *Submitted*

van der Heide FCT, **Foreman YF**, Franken IWM, Henry RMA, Houben AJHM, Kroon AA, Dagnelie PC, Eussen SJPM, Berendschot TTJM, Schouten JSAG, Webers CAB, Schram MT, van der Kallen CJH, van Greevenbroek MMJ, Wesselius A, Schalkwijk CG, Schaper NC, Brouwers MCGJ, Stehouwer CDA. *(Pre)diabetes, greater glycemia, and greater daily glucose variability are associated with lower retinal nerve fiber layer thickness, an index of neurodegeneration – The Maastricht Study*.

A

**Foreman YD**, Nijland RJ, Brouwers MCGJ, Köhler S, van Boxtel MPJ, van Greevenbroek MMJ, van der Kallen CJH, Eussen SJPM, Koster A, Wesselius A, Kroon AA, Dagnelie PC, de Galan BE, Schaper NC, Schram MT, Stehouwer CDA. *Higher daily glucose variability is associated with worse cognitive performance – The Maastricht Study.*

van der Heide FCT, Steens ILM, Geraets AFJ, **Foreman YD**, Henry RMA, Kroon AA, van der Kallen CJH, van Sloten TT, Dagnelie PC, van Dongen MCJM, Eussen SJPM, Berendschot TTJM, Schouten JSAG, Webers CAB MD, Schram MT, van Greevenbroek MMJ, Wesselius A, Schalkwijk CG, Koster A, Savelberg HHCM, Schaper NC, Stehouwer CDA. *Lower retinal nerve fiber layer thickness, an index of neurodegeneration, is associated with higher incidence of clinically relevant depressive symptoms and more depressive symptoms over time – The Maastricht Study.*

## Oral presentations

Annual Dutch Diabetes Research Meeting (ADDRM) 2019 (Wageningen, the Netherlands): *Glucose variability assessed with continuous glucose monitoring: reliability, reference values and correlations with established glycaemic indices - The Maastricht Study*

Public Symposium The Maastricht Study 2019 (Maastricht, the Netherlands): *Continu gemeten bloedsuikers; wie schommelt er meer?*

Annual Dutch Diabetes Research Meeting (ADDRM) 2018 (Oosterbeek, the Netherlands): *The incremental glucose peak during an oral glucose tolerance test is associated with greater arterial stiffness – The Maastricht Study*

Public Symposium The Maastricht Study 2018 (Maastricht, the Netherlands): *Zijn suikerschommelingen slecht?*

## Poster presentations

Annual Meeting of the European Association for the Study of Diabetes (EASD) 2019 (Barcelona, Spain): *Glucose variability in relation to glucose metabolism status, sex, age and established glycaemic indices: the Maastricht study*

Annual Meeting of the European Association for the Study of Diabetes (EASD) 2018 (Berlin, Germany): *The maximum glucose peak during an oral glucose tolerance test is associated with greater arterial stiffness: the Maastricht Study*

## Online presentations

Annual Dutch Diabetes Research Meeting (ADDRM) 2020: *Higher daily glucose variability is associated with worse cognitive performance – The Maastricht Study*

Annual Scientific Meeting of The Maastricht Study 2020: *Metabolism and Continuous Glucose Monitoring in The Maastricht Study* (co-presentation with Prof. Martijn Brouwers)

Annual Meeting of the European Association for the Study of Diabetes (EASD) 2020: *Machine learning-based glucose prediction with use of continuous glucose and physical activity data: The Maastricht Study*

A

*Addendum*

# Curriculum vitae

### Dutch

Yuri Declan Foreman werd geboren op 12 maart 1992 in Geleen en groeide op in Eindhoven, Nederland. Zijn gezin bestaat uit zijn ouders, Tim en Margret, en zus, Meryl. Op zijn middelbare school (Pleincollege Bisschop Bekkers, Eindhoven), waaraan hij in 2010 afstudeerde met een gymnasiumdiploma, ontmoette Yuri zijn wederhelft Ilse. In 2010 begon Yuri aan zijn medische opleiding aan de Universiteit Maastricht, wat resulteerde in een behaalde bul in 2016 (cum laude). Van september 2016 tot en met februari 2021 deed Yuri promotieonderzoek onder begeleiding van prof.dr. Coen Stehouwer, prof.dr. Martijn Brouwers en prof.dr. Nicolaas Schaper bij De Maastricht Studie, afdeling Interne Geneeskunde, binnen CARIM School for Cardiovascular Diseases, van Maastricht Universitair Medisch Centrum+. Gedurende deze periode was Yuri betrokken bij het includeren van deelnemers, het opschonen van onderzoeksgegevens en de implementatie van continue glucosemonitoring. Daarnaast had hij verschillende onderwijstaken en was hij vertegenwoordiger van zijn medepromovendi bij De Maastricht Studie. Vanaf maart 2021 is Yuri werkzaam als aios Interne Geneeskunde bij Zuyderland Medisch Centrum, locaties Heerlen en Sittard-Geleen.

### English

Yuri Declan Foreman was born on March 12th, 1992, in Geleen, and was raised in Eindhoven, the Netherlands. His family consists of his parents, Tim and Margret, and sister, Meryl. At his secondary school (Pleincollege Bisschop Bekkers, Eindhoven), from which he graduated with a gymnasium degree in 2010, Yuri first met his significant other Ilse. In 2010, Yuri started his medical training at Maastricht University, which resulted in the attainment of his medical degree in 2016 (*cum laude*). From September 2016 until February 2021, Yuri conducted his PhD research under supervision of Prof. Coen Stehouwer, Prof. Martijn Brouwers, and Prof. Nicolaas Schaper at The Maastricht Study, department of Internal Medicine, within CARIM School for Cardiovascular Diseases, of Maastricht University Medical Center+. During this period, Yuri was involved in participant inclusion, data cleaning, and the implementation of continuous glucose monitoring. In addition, he had several teaching responsibilities and served as representative of his fellow PhD students at The Maastricht Study. From March 2021 onward, Yuri has been working as a resident Internal Medicine at Zuyderland Medical Center, locations Heerlen and Sittard-Geleen.

A