International Journal of
Diabetes and Clinical Research

ORIGINAL ARTICLE

The Use of Artifical Intelligence in Assessing Glucose Variability in Individuals with Diabetes Type 2 from Routine Primary Care Data

Ljiljana Trtica Majnarić, MD, PhD[1,2], František Babič, MSc, PhD[3], Zvonimir Bosnić[4], Marijana Zekić-Sušac, PhD[5] and Thomas Wittlinger, MD, PhD[6]*

[1]*Department of Internal Medicine, Family Medicine and the History of Medicine, Faculty of Medicine, JJ Strossmayer University Osijek, Croatia*

[2]*Department of Public Health, Faculty of Dental Medicine and Health, Croatia*

[3]*Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic*

[4]*Faculty of Dental Medicine and Health, JJ Strossmayer University Osijek, Croatia*

[5]*Faculty of Economics, JJ Strossmayer University Osijek, Croatia*

[6]*Department of Cardiology and Diabetology, Asklepios Harzkliniken, Germany*

*Corresponding author: *Thomas Wittlinger, MD, PhD, Asklepios Harzkliniken, Department of Cardiology and Diabetology, Köslinerstr 12, 38642 Goslar, Germany, Tel: +495321441729*

Abstract

Background: The continuous glucose monitoring technique is recommended for follow-up of individuals with diabetes type 1. For those with diabetes type 2, glucose variability measures, either performed automatically or by visit-to-visit method, can be used to complement glycosylated haemoglobin (HbA1c) in predicting long-term outcomes.

Methods: A-proof-of-concept study, conducted in primary care. A total of 63 variables were used from electronic health records to describe clinical characteristics of 110 individuals with diabetes type 2 of both gender, 40-86 years old (average 62.69), and on treatment with oral hypoglycaemic drugs. The artificial neural networks (ANN) of machine learning techniques was used to model inter-day glucose variability based on the estimation of variances (the square of the standard deviation) of sporadically recorded fasting and postprandial (2h after breakfast) glucose measurements as the outcome measures. Model of increased HbA1c (≥ 7%) was used as the benchmark. The number of variables for modelling was reduced by using the pre-processing method. Multiple linear regression (MLR) models were performed on the prepared subsets to compare to the predictive accuracy of ANN models.

Results: A higher glucose variability, for both fasting and postprandial glucose variances, was associated with higher HbA1c values (Q1-Q4 differences, p = 0.002 and 0.006, respectively). The two top-ranked variables in ANN models of glucose variability were the same, indicating HbA1c and glomerular filtration rate, a measure of chronic renal impairment. MLR models of glucose variability did not give the significant predictors.

Conclusions: For created models of glucose variability, to become practically useful, their outcome measures should be dichotomised and standardised according to the thresholds of HbA1c or some standardised measures of glucose variability, such as the coefficient of variation.

Keywords

Diabetes mellitus type 2, Glucose variability, Primary care, Routine data, Machine learning, Pilot study

Introduction

Diabetes mellitus type 2 (DM2) is a global public health issue due to its close associations with obesity, hypertension and other comorbidities, and a progression towards serious complications, such as heart attacks, heart failure, kidney failure, dementia, blindness

Majnarić et al. Int J Diabetes Clin Res 2020, 7:121

• Page 1 of 12 •

and lower limb amputations [1]. Besides the obligation to provide individuals with DM2 with counselling for healthy lifestyle, a responsibility of health care providers is also in making decisions for personalised pharmacological treatments. The guidelines recommend maintaining the blood glucose levels within the ranges of target values, the aim of which is a twofold: to avoid the medication-induced low blood glucose (hypoglycaemia) while reducing the deleterious effects of chronic hyperglycaemia on target organs and the vasculature [2].

The main constraint in making the choices of medication therapy is the lack of relevant information on diabetes control. For example, a tight diabetes control in older individuals with DM2 and comorbidities and target organs involvement is difficult to achieve, as the precise knowledge on factors influencing glycaemic control in these individuals is scarce [2]. Some of the mechanisms may include increased inflammation and oxidative stress, by which inflamed target organs tissues may worsen insulin resistance [3]. The recent reports suggest an inflammatory marker, neutrophil/lymphocyte ratio, as a new measure of diabetes control [4]. Some non-diabetic drugs, acting alone, or in the concert with other drugs, can interfere with the blood glucose levels [5].

Glycosylated haemoglobin (HbA1c), a measure of average blood glucose in the past three months, has been historically used as a measure of the overall glycaemic control, and as a surrogate marker for the development of long-term diabetes complications [6]. According to the current guidelines, HbA1c values below 7% indicate optimal glycaemic control [2]. A limitation of this measure is a lack of information on acute glycaemic variations, and an inability to predict acute hypoglycaemic events [7]. In addition, individuals with the same HbA1c value may have different mean glucose concentrations [8]. In the era of personalised medicine, HbA1c, if taken alone, is not sufficient to guide decisions on treatment [9].

Based on the recent technology advances, continuous glucose monitoring (CGM) has increasingly been used in clinical practice [10]. This method was proved as to be preferable over the self-monitoring of blood glucose (SMBG) method in guiding decisions on treatment in individuals with type 1 diabetes [11]. In particular, CGM can provide information on current glucose levels and excursions, allowing individuals with type 1 diabetes to optimise insulin therapy and dietary intake, to prevent acute hypoglycaemic events. In individuals with DM2, in whom the hypoglycaemic events are less prominent, CGM has not yet been routinely used [12]. Preliminary evidence suggests that a use of CGM in these individuals, by quantifying time below, within, and above the established glycaemic targets, can promote adherence to lifestyle behaviour, and better glycaemic and weight control, when compared with traditional

SMBG method [13]. Recently, the International Consensus Group has established target percentages of "time in range" categories for specific diabetes populations to enable effective therapeutic decision making within the parameters of the established glycaemic goals [12]. Standardisation of this key CGM measure is expected to facilitate CGM use as a tool to guide personalised treatment of many individuals with diabetes, thus reducing the risk of long-term diabetes complications [14].

This is a proof-of-concept study. In a primary care (PC) setting, CGM technology is not widely available. The most of the diabetic population in PC are individuals with DM2, which are heterogeneous due to variations in the number and structure of comorbidities and prescribed medications. For this part of the diabetic population, to whom continuous monitoring of blood glucose is not necessary to perform, we suggest the use of simple measures of glucose variability, which can be estimated from fasting and postprandial (2 h after breakfast) blood glucose data, sporadically recorded in electronic health records (eHRs), as it is usually the case in PC. Although simple-to-estimate, these measures are able to capture the highest glucose excursions (both increments and decrements) over daytime [10]. We have also used a number of data from eHRs to model clinical profiles of individuals with increased glucose variability. We believe these models would help the PC providers in recognising individuals who are candidates for assessing the risk for negative outcomes.

Methods

Study subjects and design

This was a retrospective observational and analytical pilot study conducted from October 1, 2016, to January 31, 2017, in two PC practices, in eastern Croatia. Participants were individuals diagnosed with DM2. The number of participants included in the study exceeded 100 (N = 110), which is the minimal number of subjects recommended for data modelling [15]. They were of both gender, and 40-86 years old (average 62.69). Data was collected during encounters, in a four-month lasting recruitment process. This was a sufficiently long period that all individuals who adhere to treatment with oral hypoglycaemic drugs (OHDs) rotate. During this period, participants were invited several times for blood glucose measurements. They were informed about the purpose of the study and all signed their informed consent for participation in the study and for the possible publication of the study results.
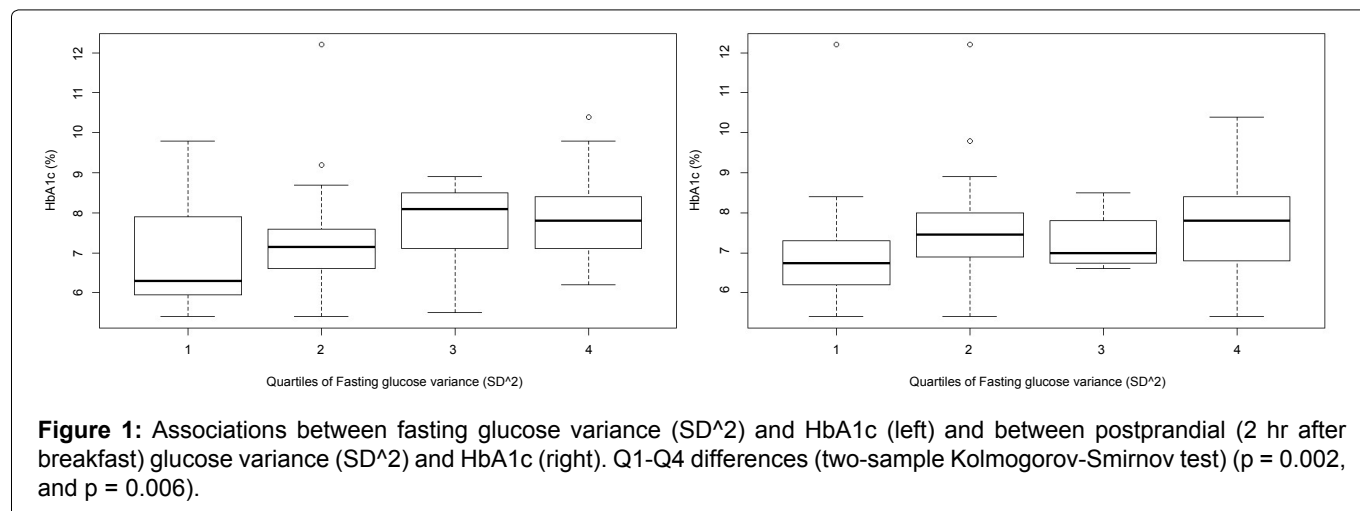
Data set

Overall, 63 variables were used to describe characteristics of participants (Part I of Appendix Table 1) and regimen treatments (Part II of Appendix Table 1). A major part of the data was used from eHRs, including general data, information on diabetes duration and age

Table 1: Multiple linear regression models predicting fasting and postprandial glucose variability and HbA1c.

Fasting glucose variability					
An explanatory variable	**t-value**	**Pr(> |t|)**	**Coefficient**	**CI**	
				5%	**95%**
Glomerular filtration rate (mL/min/1.73 m²)	0.838	0.405	0.02	-0.02	0.06
Haematocrit (%)	0.492	0.654	2.137	-5.11	9.384
HbA1c (%)	0.574	0.568	0.125	-0.239	0.49
Albuminuria (mg/24h) = No	0.735	0.465	0.652	-0.829	2.133
Comorbidities: Dg of musculoskeletal disease = No	0.674	0.503	0.457	-0.676	1.591
Comorbidities: Dg of the upper gastrointestinal tract disorders = Dyspepsio	-0.357	0.723	-0.231	-1.315	0.853
Comorbidities: Dg of the upper gastrointestinal tract disorders = Gastritis (confirmed by gastroscopy	0.761	0.449	0.617	-0.737	1.972
A type of oral antidiabetic drugs (metformin) = Yes	1.098	0.277	1.216	-0.633	3.064
A type of oral antidiabetic drugs (sulfonlurea group) = Yes	1.461	0.149	1.083	-0.155	2.322
An antidiabetic drug in use > 6 months = Yes 1	0.368	0.714	0.315	-1.113	1.742
Antidiabetic drugs in use > 6 months = 2 or more	0.602	0.549	0.526	-0.932	1.923
Hypolipidemic drug statins = Yes	0.924	0.359	0.5	-0.404	1.402
Physical activity every day (self-reported) = Yes	1	0.321	0.675	-0.452	1.801
Physical activity level - general (self-reported) = good	1.213	0.23	0.886	-0.333	2.105

Postprandial glucose variability					
An explanatory variable	**t-value**	**Pr(> |t|)**	**Coefficient**	**CI**	
				5%	**95%**
Waist circumference (cm)	0.569	0.571	0.017	-0.033	0.068
Glomerular filtration rate (mL/min/1.73 m²)	1.198	0.235	0.032	-0.013	0.078
Lymphocyte %	1.716	0.091	0.099	0.003	0.194
HbA1c (%)	1.677	0.098	0.534	0.003	1.065
Dg of autoimmune disease = Yes	1.136	0.26	1.507	-0.705	3.718
A type of oral antidiabetic drug (repaglinid or acarbose) = Yes	-0.001	0.999	-0.002	-2.188	2.185
Hypolipidemic drug statins = Yes	0.643	0.523	0.506	-0.807	1.818
Antidepressant drugs = Yes	1.195	0.236	1.348	-0.533	3.23
Meat meal every day (self-reported) = Yes	1.211	0.23	0.929	-0.531	2.21
Physical activity level - general (self-reported) = good	1.113	0.27	0.947	-0.473	2.366

Increased HbA1c					
Explanatory variables	**t-value**	**Pr(> |t|)**	**Coefficient**	**CI**	
				5%	**95%**
Mid-arm circumference (cm)	0.879	0.383	0.012	-0.01	0.034
Glomerular filtration rate (mL/min/1.73 m²)	1.27	0.209	0.014	-0.004	0.032
C-reactive protein (mg/L)	1.637	0.107	0.109	-0.002	0.219
Neutrophil %	3.315	0.002	0.038	0.019	0.057
Haematocrit (%)	1.687	0.097	3.95	0.037	7.864
LDL-cholesterol (mmol/L)	1.51	0.136	0.269	-0.029	0.567
Triglycerides (mmol/L)	1.525	0.132	0.082	-0.008	0.173
Gender (M,F) = Female	2.114	0.039	0.65	0.136	1.164
Sleep disorders (self-reported) = Yes	1.413	0.163	0.496	-0.09	1.083
Albuminuria (mg/24h) = Yes	1.828	0.072	0.75	0.065	1.435

Comorbidities: Dg of musculoskeletal disease = Yes	0.357	0.722	0.111	-0.408	0.63
A type of oral antidiabetic drug (metformin) =Yes	2.495	0.015	1.244	0.411	2.077
A type of oral antidiabetic drug (repaglinid or acarbose) = Yes	1.213	0.23	0.557	-0.21	1.325
Hypolipidemic drug statins = Yes	0.654	0.515	0.171	0.266	0.608
Hypolipidemic drugs other than statins = Yes	0.783	0.437	0.311	-0.353	0.975
Antipsychotic drugs = Yes	1.826	0.073	2.413	0.205	4.62

Figure 1: Associations between fasting glucose variance (SD^2) and HbA1c (left) and between postprandial (2 hr after breakfast) glucose variance (SD^2) and HbA1c (right). Q1-Q4 differences (two-sample Kolmogorov-Smirnov test) (p = 0.002, and p = 0.006).

of onset, diagnoses of other chronic diseases (indicating diabetes comorbidities), and information on medications prescription. The laboratory data was used from PC laboratory templates. Laboratory tests were performed as a part of periodic chronic disease surveillance or preventive check-ups.

Information on non-pharmacological regimens and some conditions for which evidence indicated their associations with impaired glucose metabolism, such as sleep and cognitive disorders, was collected by participant self-reports [16,17]. Anthropometric measurements, if not updated in eHRs, were taken from participants during encounters.

Blood glucose values, which are necessary for estimating glucose variability, were obtained from participants by using a glucometer device and in 6-8 routes of random measurements (Appendix Table 1). A value of HbA1c for each participant was taken from eHRs, as measured in the period under study as a part of the periodic surveillance of individuals with DM2 (Appendix Table 1).

Measuring glucose variability

The variance, a statistical measure defined as the square of the standard deviation (SD^2), was used as a measure of inter-day glucose variability, for both fasting and postprandial (2 h after breakfast) glucose variations [10]. This statistical measure demonstrates the extent of variability by taking into account the mean of glucose values.

The box-plot diagrams of HbA1c according to Q1-Q4 of variances of fasting and postprandial glucose values were used to show how HbA1c values depend on de-

grees of glucose variability (Figure 1). The Kolmogorov-Smirnov test was used to calculate Q1-Q4 differences.
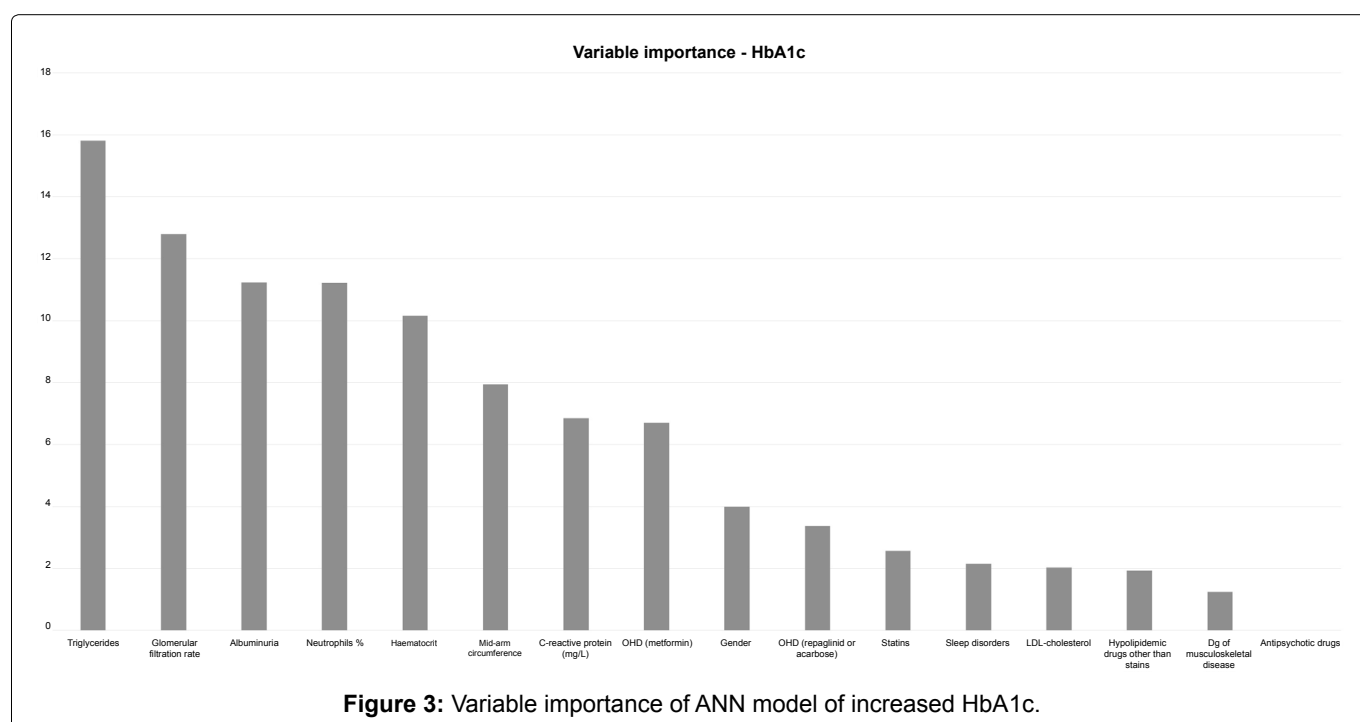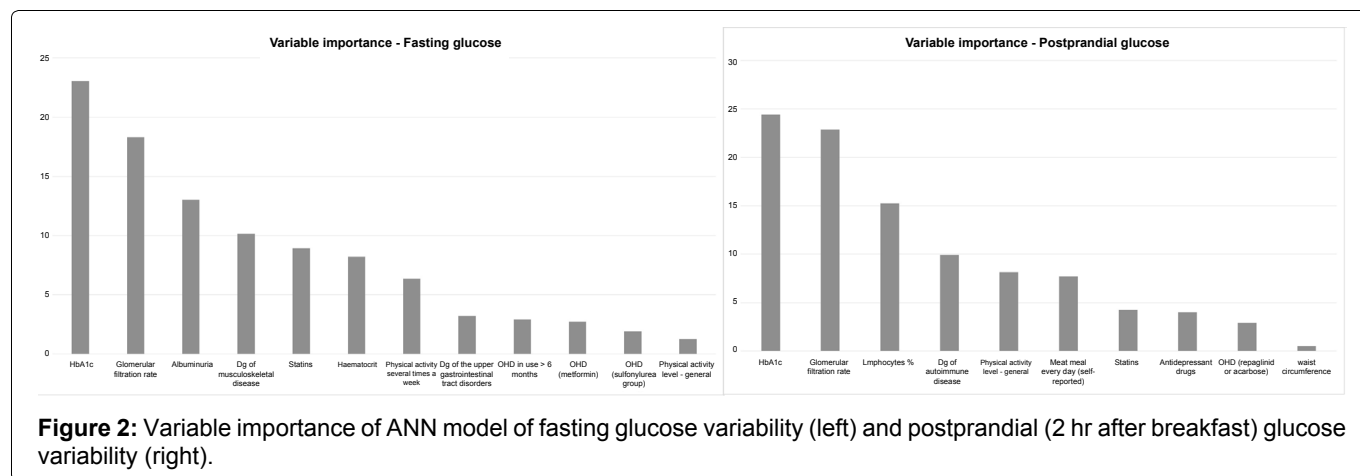
Modelling predictors of glucose variability

Managing older individuals with multiple medical conditions, such as those with DM2, is as dealing with the complex system [18]. In statistical terms, this means non-linear associations, interdependence, collinearity, and networking between the variables. By contrast, classical multiple linear regression (MLR) models are based on the assumptions of independence of the input variables, linearity between dependent and independent variables, normality of the residuals, and the absence of endogenous variables [15].

As there was a need to use more robust regression models, for which there are no theoretical constraints for data modelling, we have performed the artificial neural networks (ANN) algorithm of machine learning (ML) methods [19]. The variances of fasting and postprandial blood glucose measurements were used as the output measures.

The artificial neural networks (ANN) method is known to be robust and to not require strong assumptions on the distributions and interdependence of the predictors. It uses a multi-layered structure to imitate the multiprocessing of biological networks [19]. The ANN model for predicting increased HbA1c (≥ 7.0%) was performed to serve as the benchmark. A specific type of ANN method used in this study was the multilayer perceptron (MLP).

In searching for the most accurate model, different numbers of hidden units, varying from 1 to 20, were used in ANN architectures with one hidden layer or two

Figure 2: Variable importance of ANN model of fasting glucose variability (left) and postprandial (2 hr after breakfast) glucose variability (right).

Figure 3: Variable importance of ANN model of increased HbA1c.

and three hidden layers.

Due to the large number of variables used for participants` description, the variable reduction procedure was conducted using the Spearman`s rank correlation coefficient and the best-first search algorithm [20]. This algorithm selects the subsets of input variables and cross-validates the model by training and testing. Based on this procedure, the number of input variables was significantly reduced: to 12 predictors for fasting glucose variability as an output; to 10 predictors for postprandial glucose variability as an output; and to 16 predictors for HbA1c as an output (Figure 2 and Figure 3).

The cross-validation procedure is a re-sampling method, used to evaluate the performance of ML models when the amount of data is limited [19]. It provides a model with the opportunity to train on multiple train-test splits, i.e. to predict how well the model will perform on unseen data. The data is separated into two sets, called the training set and the testing set. The predictive model built on the training set is validated on the test set. In our study, we used the 70:30 partition ratio.

The variable importance procedure was conducted on all three ANN models (Figure 2 and Figure 3). In both models of glucose variability, the two top-ranked variables were the same, indicating HbA1c and glomerular filtration rate, a marker of reduced renal function (Figure 2).

The MLR models were performed on the selected variables, to compare to the predictive accuracies of ANN models (Table 1).

For testing differences between the models, the minimum mean squared error (MSE), and a more explanatory symmetric mean average percentage error (SMAPE), were used (Table 2) [21]. The SMAPE is an average of the absolute percentage errors, although these errors are computed using a denominator representing the average of the predicted and observed values. The SMAPE has an upper limit of 200%, offers a well-designed range to judge the level of accuracy, and is less strongly influenced by extreme values. The lowest value shows the best model.

Majnarić et al. Int J Diabetes Clin Res 2020, 7:121

• Page 5 of 12 •

Results

Results of assessing glucose variability

The Q4 of fasting or postprandial glucose variance, compared to Q1, is considered as a higher level of glucose variability.

A considerable part of subjects in Q1 of fasting glucose variance had HbA1c values which are within the range of normal HbA1c values, while a majority of subjects in Q4 had increased HbA1c values (≥ 7%). The Q1-Q4 difference was statistically significant (p = 0.002) (Figure 1, left).

An association between Q1-Q4 of postprandial glucose variance and HbA1c, also showed statistically significant difference (p = 0.006) (Figure 2, right). Compared to Q4 of fasting glucose variance, in Q4 of postprandial glucose variance, more subjects had HbA1c values which are within the range of normal HbA1c values.

Results of modelling predictors of glucose variability

Based on the variable reduction procedure, for fasting glucose variability, the number of variables for modelling by the ANN method, was reduced to twelve. Those were variables indicating:

HbA1c, glomerular filtration rate, albuminuria, dg of musculoskeletal disease, use of statins, haematocrit, physical activity every day, dg of upper gastrointestinal tract disorders, more than 6 months use of OHD, use of metformin, use of sulfonilurea group drugs, and physical activity level in general good.

The number of selected predictors in ANN model of postprandial glucose variability was ten, indicating:

HbA1c, glomerular filtration rate, lymphocyte %, dg of autoimmune disease, activity level in general good, eating meat every day, statins use, use of antidepressant drugs, use of repaglinid or acarbose, and waist circumference.

For ANN model of increased HbA1c, there were sixteen predictors, indicating:

triglycerides, glomerular filtration rate, albuminuria, neutrophil %, haematocrit, mid-arm circumference, C-reactive protein, use of metformin, F gender, use of repaglinid or acarbose, use of statins, sleep disorders, LDL-cholesterol, hypolipidemic drugs other than statins, dg of musculoskeletal disease, and antipsychotic drugs.

Figure 2 and Figure 3 show the importance order of input variables of three ANN models.

It can be seen that the first two-ranked variables are the same in both ANN models of glucose variability, indicating HbA1c and glomerular filtration rate (Figure 2, left).

The third-ranked variable in ANN model of fasting glucose variability indicate albuminuria, while in ANN model of postprandial glucose variability, on the third position is a variable indicating lymphocyte % (Figure 2, right).

In ANN model of increased HbA1c, the top-ranked variables indicate triglycerides, glomerular filtration rate, albuminuria, and neutrophil % (Figure 3).

Results of assessing predictive accuracies of ANN models have shown that the accuracy is better when predicting increased HbA1c than when predicting glucose variability, with MSE = 0.17, and SMAPE = 1.47% (Table 2).

No one of variables in either of the two MLR models of glucose variability showed significance as an independent predictor of glucose variability (Table 1).

The predictive power of MLR model of fasting glucose variability was MSE = 2.47, and SMAPE = 69.50%. For MLR model of postprandial glucose variability, the predictive power was MSE = 6.79, and SMAPE = 9.64% (Table 2).

The MRL model of increased HbA1c showed better predictive accuracy than the previous two, with MSE = 1.66, and SMAPE = 13.44% (Table 2). Some variables, like those indicating neutrophil %, F gender, and therapy with metformin, were significant at the level of 0.05 (Table 1).

Discussion

This preliminary research showed that it is possible, in individuals with DM2 and receiving OHDs, by using routinely collected data from PC eHRs, to identify predictors of a higher risk for long-term negative outcomes. A ML-based approach (ANN models) was used to model data describing many clinical characteristics of participants. The classical MLR method was used to check the predictive accuracy of variables selected in ANN models.

Both type of methods were more successful in predicting increased HbA1c than in predicting glucose variability. The reason may be the fact that the outcome measure, used to indicate glucose variability, lacks standardisation regarding cut-off values that would be complementary to the threshold of increased HbA1c. The only statistical measure for which it was done, is a coefficient of variation (CV%), with the cut-off of 36% [12]. To create the adequate model, a large sample would be necessary.

Although both type of models (MLR and ANN) were based on the same set of selected variables, ANN models, especially that indicating increased HbA1c, showed as to be preferable over MLR models (Table 2). In addition, the variable importance procedure, performed on ANN models, has enabled better insights into clinical contexts (Figure 2 and Figure 3). That ML methods are suitable for issues for which there is a need to cope with

Majnarić et al. Int J Diabetes Clin Res 2020, 7:121

• Page 6 of 12 •

Table 2: Results of testing differences in prediction accuracies between the created ANN, and MLR models.

ANN model	ANN architecture	MSE on the test sample	SMAPE on the test sample (%)
Fasting glucose variability	2 hidden layers	5.23	23.69
Postprandial glucose variability	1 hidden layer	30.58	33.24
Increased HbA1c	1 hidden layer	0.17	**1.47**

ANN: Artificial neural networks; HbA1c: Glycosylated hemoglobin A1c; MSE: Mean squared error; SMAPE: Symmetric mean average percentage error.

MLR model	Model's size	MSE on the test sample	SMAPE on the test sample (%)
Fasting glucose variability	number of explanatory variables = 14**	2.47	69.50
Postprandial glucose variability	number of explanatory variables = 10	6.79	99.64
Increased HbA1c	number of explanatory variables = 16	1.66	**13.44**

MLR: Multiple linear regression; **The variables "Dg of the upper gastrointestinal tract disorders" and "Antidiabetic drugs in use > 6 months" was split into the two components.

the biological complexity, we know from our previous research [22,23].

The study results indicate that two type of measures of glycaemic control, HbA1c, and variances of fasting and postprandial glucose measurements, are partially complementary to each other. One implication is that both measures should be used for risk prediction, The recent longitudinal study demonstrated that individuals with higher visit-to-visit variability of glycaemia (based on measuring CV) have an increased risk of developing micro- and macrovascular complications and mortality [24]. Another implication of the complementarity between the outcome measures is that HbA1c-based model and glucose variability-based models share common sets of variables.

In this terms, results of QQ analysis showed that the principle, the higher variances of fasting and postprandial glucose, the higher HbA1c, is only partly true, as there are participants with a higher glucose variability (Q4) who do not have increased HbA1c, and *vice versa*, a considerable part of those with increased HbA1c are at a lower level of glucose variability (Q1) (Figure 1). In addition, from models indicating increased HbA1c, it can be seen that HbA1c might be influenced by other factors than the level of glucose exposure (Table 1 and Figure 3). And oppositely, glucose variability measures are likely to be influenced by a wide range of factors, each having a small contribution but neither reaching a power as an independent predictor, thus demonstrating a behaviour as a complex system (Table 1 and Figure 2). Among these factors, HbA1c plays an important role (ANN models of glucose variability) (Figure 2) but still not as a dominating, independent predictor (MLR models of glucose variability) (Table 1).

Evidence suggests chronic renal impairment as a co-morbid condition which attenuates HbA1c as a measure of glycaemic control by weakening its association with mean glucose values [25]. This is also associated with higher dispersion of mean glucose value patterns and thus with increased glucose variability [26].

Chronic renal impairment in older adults is a consequence of comorbidities of diabetes and related metabolic and cardiovascular (CV) disorders, which, in turn, has an effect on changes in body shape and whole-body homeostasis, by mechanisms such as chronic inflammation, protein catabolism and malnutrition [27]. The hallmark of diabetes related renal impairment is albuminuria, which is also known as an independent CV risk factor [28]. By mechanisms such as chronic inflammation and disturbed rheology, chronic renal impairment may futher worsen insulin resistance and glucose related metabolism [29].

All these pieces of evidence, indicating a pathophysiology scenario which can be described as **chronic renal impairment - chronic inflammation - disturbed blood rheology**, are likely to be suggested by the top ranked variables in ANN models. These variables include: triglycerides, glomerular filtration rate, albuminuria, neutrophil %, haematocrit, mid-arm circumference, and C-reactive protein (in ANN-HbA1c model) (Figure 3); glomerular filtration rate and albuminuria (in ANN-fasting glucose variability model) (Figure 2); and glomerular filtration rate and lymphocyte % (in ANN-postprandial glucose variability model) (Figure 2). The neutrophil-to-lymphocyte ratio has recently been recognised as a marker of chronic renal impairment, and as a good measure of glycaemic control, being comparable to HbA1c [4,30]. The importance of this derived parameter is likely to be highlighted by these results, and especially with the result of MLR-HbA1c model, which indicates the variable "neutrophil %" as an independent predictor (Table 1).

The small predictive power of variables selected in the created models is in contrast to the models` satisfactory performances. This is especially case in ANN-HbA1c model, which is characterised with very good

predictive performances (small errors). It is therefore reliable to use sets of complementary variables to describe particular contexts, than to realise the meaning of the single variables. Considered this way, several sets of variables other than those indicating chronic renal impairment, although with a minor predictive power, can be also recognised in the ANN models. These sets of variables can be labelled as: **drugs**, **comorbidities**, and **lifestyles**.

The variables indicating **drugs** include old-fashioned OHDs, such as metformin, repaglinid/acarbose and sulfonilurea-group OHDs, as well as statins, hypolipidemic drugs (all three ANN models) (Figure 2 and Figure 3). The choice of pharmacological treatment may be of the special importance for its effect on fasting glucose variability, and specifically in individuals with a long-term diabetes duration (the variable "OHDs in use of more than 6 months") (ANN model of fasting glucose variability) (Figure 2). Although there are no clear recommendations for pharmacological treatment of older individuals with DM2 and chronic renal impairment, an awareness is rising on the exceptional complexity of these individuals, and the need for their individualised treatment and careful monitoring [31]. A special attention is also needed to those who in addition use antidepressants and antipsychotic drugs, due to the side effects of these drugs on glucose related metabolism (Figure 2 and Figure 3) [32].

The **comorbidities** suggested by ANN models are musculoskeletal diseases, including also rheumatoid arthritis, a commonly found inflammatory joint disease in older population (the variable „autoimmune disease", in ANN model of postprandial glucose variability), together with the upper gastro-intestinal tract disorders and sleep disorders (Figure 2 and Figure 3). Musculoskeletal diseases are less frequent than CV and metabolic disorders but still make a prominent part of the comorbidity patterns of patients with DM2 [33]. The association with frailty and sleep disorders, due to chronic pain, are factors which may relate musculoskeletal diseases to higher risk of DM2 complications [34].

Of **lifestyle behaviours**, our results suggest the importance of physical activities for glucose variability, in particular for fasting glucose variability, as more sensitive to hypoglycaemic events (the variables "physical activities - general" and "physical activities - every day") (ANN models of fasting and postprandial glucose variability) (Figure 2) [6].

Conclusions

Models of glucose variability can be created using data from PC eHRs, to help PC providers recognising clinical profiles of individuals with DM2, and on treatment with OHDs, who might be at higher risk for negative outcomes. As model outcome measures, sporadically recorded visit-to-visit fasting and postprandial (2h after breakfast) glucose measurements, widely avail-

able in PC, can be used. All predictive variables showed small contribution to the overall model`s predictive power, reflecting the complexity of individuals in the sample. For models, to become useful in a practice, it is necessary that outcome measures are dichotomised and standardised according to the thresholds of HbA1c or some standardised measures of glucose variability, such as the coefficient of variation.

Author Disclosure Statement

No competing financial interests exist.

Funding Source

References

1. Zimmet PZ (2017) Diabetes and its drivers: The largest epidemic in human history? Clin Diabetes Endocrinol 3: 1.

2. Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, et al. (2020) 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD: The Task Force for diabetes, pre-diabetes, and cariovascular diseases of the European Society of Cardiology (ESC) and the European Association for the Study of Diabetes (EASD). Eur Heart J 41: 255-323.

3. Rehman K, Akash MSH (2016) Mechanisms of inflammatory responses and development of insulin resistance: How are they interlinked? J Biomed Sci 23: 87.

4. Hussain M, Babar MZM, Akhtar L, Hussain MS (2017) Neutrophil lymphocyte ratio (NLR): A well assessment tool of glycemic control in type-2 diabetic patients. Pak J Med Sci 33: 1366-1370.

5. Tsujimoto T, Sugiyama T, Shapiro MF, Noda M, Kajio H (2017) Risk of cardiovascular events in patients with diabetes mellitus on β-blockers. Hypertesnion 70: 103-110.

6. Brownlee M, Hirsch IB (2006) Glycemic variability: A hemoglobin A1c-independent risk factor for diabetic complications. JAMA 295: 1707-1708.

7. Qu Y, Jacober SJ, Zhang Q, Wolka LL, DeVries JH (2012) Rate of hypoglycemia in insulin-treated patients with type 2 diabetes can be predicted from glycemic variability data. Diabetes Technol Ther 14: 1008-1012.

8. Beck RW, Connor CG, Mullen DM, Wesley DM, Bergenstal RM (2017) The fallacy of average: How using HbA1c alone to assess glycemic control can be misleading. Diabetes Care 40: 994-999.

9. Fradkin JE, Hanlon MC, Rodgers GP (2016) NIH precision medicine initiative: Implications for diabetes research. Diabetes Care 39: 1080-1084.

10. Kovatchev BP (2017) Metrics of glycemic control - from HbA1c to continuous glucose monitoring. Nat Rev Endocrinol 13: 425-436.

11. Olafsdottir AF, Polonsky W, Bolinder J, Hirsch IB, Dahlqvist

S, et al. (2018) A randomized clinical trial of the effect of continuous glucose monitoring on nocturnal hypoglycemia, daytime hypoglycemia, glycemic variability, and hypoglycemia confidence in persons with type 1 diabetes treated with multiple daily insulin injections (GOLD-3). Diabetes Technol Ther 20: 274-284.

12. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, et al. (2019) Clinical targets for continuous glucose monitoring data interpretation: Recommendations from the international consensus on time in range. Diabetes Care 42: 1593-1603.

13. Taylor PJ, Thompson CH, Brinkworth GD (2018) Effectiveness and acceptability of continuous glucose monitoring for type 2 diabetes management: A narrative review. J Diabetes Investig 9: 713-725.

14. Carlson AL, MD, Mullen DM, Bergenstal RM (2017) Clinical use of continuous glucose monitoring in adults with type 2 diabetes. Diabetes TechnolTher 19: S4-S11.

15. Johnson JL (2011) Probability and statistics for computer science. Whiley, New York.

16. Hall MH, Okun ML, Sowers MF, Matthews KA, Kravitz HM, et al. (2012) Sleep is associated with the metabolic syndrome in a multi-ethnic cohort of midlife women: The SWAN sleep study. Sleep 35: 783-790.

17. Panza F, Frisardi V, Capurso C, Imbimbo BP, Vendemiale G, et al. (2010) Metabolic syndrome and cognitive impairment: Current epidemiology and possible underlying mechanisms. J Alzheimers Dis 21: 691-724.

18. Sturmberg JP, Bennett JM, Martin CM, Picard M (2017) 'Multimorbidity' as the manifestation of network disturbances. J Eval Clin Prac 23: 199-208.

19. Russel S, Norvig P (2003) Artificial intelligence: A modern approach. Prentice Hall, New York.

20. Tofallis C (2015) A better measure of relative prediction accuracy for model selection and model estimation. J Oper Res Soc 66: 1352-1362.

21. Makridakis S (1993) Accuracy measures: Theoretical and practical concerns. Int J of Forecast 9: 527-529.

22. Majnarić Lj, Zekic-Susac M (2012) Elucidating clinical context of lymphopenia by nonlinear modelling. Expert Syst Appl 39: 10889-10897.

23. Sabanovic S, Majnaric Trtica Lj, Babic F, Vadovský M, Paralič J, et al. (2018) Metabolic syndrome in hypertensive women in the age of menopause: A case study on data from general practice electronic health records. BMC Med Inform Decis Mak 18: 24.

24. Slieker RC, van der Heijden AAWH, Nijpels G, et al. (2019) Visit-to-visit variability of glycemia and vascular complications: The hoorn diabetes care system cohort. Cardiovasc Diabetol 18: 170.

25. Bloomgarden Z, Handelsman Y (2018) How does CKD affect HbA1c? J Diabetes 10: 270.

26. Jung M, Warren B, Grams M, Kwong YD, Shafi T, et al. (2018) Performance of non-traditional hyperglycemia biomarkers by chronic kidney disease status in older adults with diabetes: Results from the atherosclerosis risk in communities study. J Diabetes 10: 276-285.

27. Walker SR., Wagner M, Tangri N (2014) Chronic kidney disease, frailty and successful aging: A review. J Ren Nutr 24: 364-370.

28. Gerstein HC, Mann JF, Yi Q, Zinman B, Dinneen SF, et al. (2001) Albuminuria and risk of cardiovascular events, death, and heart failure in diabetic and nondiabetic individuals. JAMA 286: 421-426.

29. Spoto B, Pisano A, Zoccali C (2016) Insulin resistance in chronic kidney disease: A systematic review. Am J Physiol Renal Physiol 311: F1087-F1108.

30. Tonyali S, Ceylan C, Yahsi S, Sebnem Karakanb M, et al. (2018) Does neutrophil to lymphocyte ratio demonstrate deterioration in renal function? Ren Fail 40: 209-212.

31. Clemens KK, O'Regan N, Rhee JJ (2019) Diabetes management in older adults with chronic kidney disease. Curr Diab Rep 19: 11.

32. Zhang Y, Liu Y, Su Y, You Y, Ma Y, et al. (2017) The metabolic side effects of 12 antipsychotic drugs used for the treatment of schizophrenia on glucose: A network meta-analysis. BMC Psychiatry 17: 373.

33. Lang VB, Markovic BB (2016) Prevalence of comorbidity in primary care patients with type 2 diabetes and its association with elevated HbA1c: A cross-sectional study in Croatia. Scand J Prim Health Care 34: 66-72.

34. Otones Reyes P, Garcia Perea E, Pedroz Marcos A (2019) Chronic pain and frailty in community-dwelling older adults: A systematic review. Pain Manag Nurs 20: 309-315.

Appendix Table 1: Input parameters description.

	Variable name*	Variable type	Descriptive statistics
Part I. Parameters which describe patients characteristics			
	Age (years)	Continuous	Min.: 40.00, Max.: 86.00, Mean: 62.69
	Diabetes duration (years)	Continuous	Min.: 1.0, Max.: 57.0, Mean : 8.3
	Age diabetes diagnosed (years)	Continuous	Min.: 35.0, Max.: 74.0, Mean : 54.47
	Body mass index (kg/cm²)	Continuous	Min.: 23.0, Max.: 42.82, Mean : 31.27
	Waist circumference (cm)	Continuous	Min.: 80.0, Max.: 149.0, Mean : 103.89
	Mid-arm circumference (cm)	Continuous	Min.: 27.0, Max.: 78.0, Mean : 40.19
	Glomerular filtration rate (mL/min/1.73 m²)	Continuous	Min.: 36.0, Max.: 106.0, Mean : 77.25
	Hemoglobin (g/L)	Continuous	Min.: 97.0, Max.: 164.0, Mean : 136.00
	C-reactive protein (mg/L)	Continuous	Min.: 0.30, Max.: 9.8, Mean : 2.951
	Lymphocytes %	Continuous	Min.: 10.40, Max.: 43.4, Mean : 30.08
	Monocytes %	Continuous	Min.: 1.2, Max.: 12.4, Mean : 7.69
	Neutrophils %	Continuous	Min.: 25.0, Max.: 82.7, Mean : 55.04
	Hematocrit (%)	Continuous	Min.: 0.19, Max.: 0.5, Mean : 0.3883
	Total cholesterol (mmol/L)	Continuous	Min.: 2.5, Max.: 9.800, Mean : 5.136
	LDL-cholesterol (mmol/L)	Continuous	Min.: 1.1, Max.: 5.0, Mean : 2.797
	HDL-cholesterol (mmol/L)	Continuous	Min.: 0.7, Max.: 2.70, Mean : 1.315
	Triglycerides (mmol/L)	Continuous	Min.: 0.6, Max.: 17.70, Mean : 2.654
	HbA1c (%)	Continuous	Min.: 5.4, Max.: 12.20, Mean : 7.363
	Gender (M,F)	Categorical, M; F	1=52 (47.27%), 2=58 (52.73%)
	Walking properties, mobility (self-reported and assessed by a physician)	-Completely independently -Decreased -Significantly decreased	-94 (85.45%) -12 (10.91%) -4 (3.64%)
	Cognitive dysfunction (self-reported)	No Yes	-100 (90.91%) -10 (9.09%)
	Sleep disorders (self-reported)	No Yes	-80 (72.73%) -30 (27.27%)
	Albuminuria (mg/24h)	No Yes	-97 (88.18%) -13 (11.82%)
	Dg of hypertension or using antihypertensive drugs	-No -≤ 2 years (y) duration -3-10 y -11-20 y ->20 y	-13 (11.82%) -8 (7.27%) -49 (44.55%) -34 (30.91%) -6 (5.45%)
	Dg of Clinically manifest cardiac insufficiency or Infarctus myocardii	No Yes	-93 (85.45%) -17 (15.45%)
	Dg of Cerebrovascular incident or transitory ischaemic attack	No Yes	-106 (96.36%) -4 (3.64%)
	Dg of Peripheral ischemic vascular disease	No Yes	-97 (88.18%) -13 (11.82%)
	Dg of retinopathy	No Yes	-96 (87.27%) -14 (12.73%)
	Dg of polyneuropathy	No Yes	-92 (83.64%) -18 (16.36%)

	Frequent (≥ 3 times a year) urinary tract infection	No	-84 (76.36%)
		Yes	-26 (23.64%)
	Hypertension (arterial blood pressure) is being well regulated?	No	-43 (39.09%)
		Yes	-67 (60.91%)
	Comorbidities: Dg of musculosceletal disease	No	-77 (70.00%)
		Yes	-33 (30.00%)
	Dg of autoimmune disease	No	-103 (93.64%)
		Yes	-7 (6.36%)
	Comorbidities: significant hepatic disease (Dg assigned)	No	-109 (99.09%)
		Yes	-1 (0.91%)
	Comorbidities: Dg of the upper gastrointestinal tract disorders	-No	-67 (60.91%)
		-Dyspepsio	-25 (22.73%)
		-Gastritis (confirmed by gastroscopy)	-18 (16.36%)
	Dg of Chr pulmonary disease (COPB)	No	-99 (90.00%)
		Yes	-11 (10.00%)
Part II. Parameters which describe the regimen treatment (medications, non-pharmacological treatment)			
	A type of oral antidiabetic drugs (metformin)	No	-14 (12.73%)
		Yes	-96 (87.27%)
	A type of oral antidiabetic drugs (gliptins)	No	-79 (71.82%)
		Yes	-31 (28.18%)
	A type of oral antidiabetic drugs (sulfonlurea group)	No	-68 (61.82%)
		Yes	-42 (38.18%)
	A type of oral antidiabetic drugs (repaglinid or acarbose)	No	-97 (88.18%)
		Yes	-13 (11.82%)
	Antidiabetic drug (s) in use < 6 months	No	-98 (89.09%)
		Yes 1	-9 (8.18%)
		Yes 2 or more	-3 (2.73%)
	Antidiabetic drug(s) in use > 6 months	No	-12 (10.91%)
		Yes	-50 (45.45%)
		Yes 2 or more	-48 (43.64%)
	A dose level of an antidiabetic drug	Optimal	-66 (60.00%)
		Maximal	-44 (40.00%)
	Hypolipidemic drug statins	No	-55 (50.00%)
		Yes	-55 (30.00%)
	Hypolipidemic drugs other than statins	No	-96 (87.27%)
		Fenofibrat	-14 (12.73%)
	Beta-blockers	No	-74 (67.27%)
		Yes	-36 (32.73%)
	Diuretics	No	-74 (67.27%)
		Yes	-36 (32.73%)
	Calcium channels blockers	No	-77 (70.00%)
		Yes	-33 (30.00%)
	Ace-inhibitors or Angiotensin receptor blockers	No	-29 (26.36%)
		Yes	-81 (73.64%)
	Drugs with anti-inflammatory properties	No	-68 (61.82%)
		Yes	-42 (38.18%)

	Anticoagulant drugs	No	-100 (90.91%)
		Yes	-10 (9.09%)
	Anxiolytic drugs	No	-67 (60.91%)
		Yes	-43 (39.09%)
	Antidepressant drugs	No	-96 (87.27%)
		Yes	-14 (12.73%)
	Antipsychotic drugs	No	-108 (98.18%)
		Yes	-2 (1.82%)
	Diat recommendations follow-up (self-reported)	-Insufficient	-56 (50.91%)
		-Fairly well, sufficient	-54 (49.09%)
	Meat meal every day (self-reported)	No	-65 (59.09%)
		Yes	-46 (40.91%)
	Fresh fruits and vegetables every day (self-reported)	No	-29 (26.36%)
		Yes	-81 (73.64%)
	Physical activity every days (self-reported)	No	-68 (61.82%)
		Yes	-42 (38.18%)
	Physical activity several times a week (self-reported)	No	-73 (66.36%)
		Yes	-37 (33.64%)
	Physical activity level - general (self-reported)	Weak	-78 (70.91%)
		Fairy well, good	-32 (29.09%)
	Output variables		
	Fasting blood glucose-variance	Continuous	Min.: 0.0200, Max.: 12.2000, Mean: 1.6540
	Postprandial blood glucose-variance	Continuous	Min.: 0.0000, Max.: 20.7100, Mean: 2.2692
	HbA1c	Continuous	Min.: 5.4, Max.: 12.20, Mean : 7.363

*Dg- diagnosis, Chr- chronic, HbA1c - glycosylated hemoglobin A1c.