

Differential risk of cardiovascular complications in patients with adult type-2 diabetes mellitus in Ghana using clustering analysis: A hospital-based cross-sectional study

Christian Obirikorang^{1,2*}, Evans Asamoah Adu^{1,2}, Anthony Afum-Adjei Awuah^{1,2}, Samuel Nkansah Darko¹, Frank Naku Ghartey³ Samuel Ametepe^{1,4}, Eric NY Nyarko⁵, Enoch Odame Anto^{6,7}

¹Department of Molecular Medicine, School of Medical Sciences, Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana

²Global Health and Infectious Disease, Kumasi Centre for Collaborative Research in Tropical Medicine, Kumasi, Ghana

³Department of Chemical Pathology, School of Medical Sciences, University of Cape Coast, Ghana

⁴Department of Medical Laboratory Sciences, Koforidua Technical University, Koforidua, Ghana

⁵Department of Chemical Pathology, University of Ghana Medical School, University of Ghana

⁶Department of Medical Diagnostics, Faculty of Allied Health Sciences, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

⁷Centre for Precision Health, School of Medical and Health Sciences, Edith Cowan University, Western Australia, Australia

***Corresponding Author**

Email: krisobiri@yahoo.com, (CO)

Authors' Contributions

All authors contributed equally to this research.

Abstract

Aim: To characterize clinically relevant subgroups of patients with T2DM based on adiposity, insulin secretion, and resistance indices. **Methods:** A cross-sectional study was conducted at Eastern Regional Hospital in Ghana from July to October 2021 to investigate long-term patients with T2DM. To select participants, a systematic random sampling method was employed. Demographic data was collected using a structured questionnaire and fasting blood samples were taken to measure glycemic and lipid levels. Blood pressure and adiposity indices were measured during recruitment. The risk of cardiovascular disease (CVD) was defined using Framingham scores and standard low-density lipoprotein thresholds. To analyze the data, k-means clustering algorithms and regression analysis were used. **Results:** The study identified three groups in female patients according to body mass index, relative fat mass, glycated hemoglobin, and triglyceride-glucose index. These groups included the obesity-related phenotype, the severe insulin resistance phenotype, and the normal weight phenotype with improved insulin resistance. Among male patients with T2DM, two groups were identified, including the obesity-related phenotype with severe insulin resistance and the normal weight phenotype with improved insulin sensitivity. The severe insulin resistance phenotype in female patients was associated with an increased risk of high CVD (OR=5.34, 95%CI:2.11-13.55) and MetS (OR=7.07; 95%CI:3.24-15.42). Among male patients, the obesity-related phenotype with severe insulin resistance was associated with an increased intermediate (OR=21.78, 95%CI:4.17-113.78) and a high-risk CVD (OR=6.84, 95%CI:1.45-32.12). **Conclusions:** The findings suggest that there are specific subgroups of patients with T2DM characterized by obesity and uncontrolled insulin resistance leading to poor glycemic control. This underscores the importance of considering differences in adiposity, insulin secretion, and sensitivity indices when making clinical decisions for patients with T2DM.

Keywords: type-2 diabetes, subgroups, cluster analysis, cardiovascular risk.

Introduction

Type-2 diabetes mellitus (T2DM) is a chronic condition characterized by high levels of circulating blood glucose. It is the most common type of diabetes, typically affecting adults with or without obesity-associated complications [1]. While a diagnosis of diabetes can be made based on glycated hemoglobin and/or fasting blood glucose levels [1, 2], this does not provide a definitive classification of the disease. Thus, T2DM diagnosis is made in the absence of type-1 diabetes mellitus, which has unique immune and genetic markers for classification [2]. Unfortunately, the current approach to classifying T2DM falls short in understanding the epidemiological and clinical landscape of the disease. Moreover, multiple risk factors can influence the progression of T2DM, making it difficult to predict prognosis and track disease progression [3].

The prevalence of diabetes has increased worldwide, especially in low and middle-income countries [1]. In the International Diabetes Federation (IDF) report, 1 in 10 adults have diabetes worldwide, with 3 in 4 adults in low- and middle-income countries [4]. In Africa, diabetes is projected to increase by 129% by 2045, with mortality rates increasing by up to 20.8% between 2011 and 2021 [4]. In Ghana, 2.6% of the adult population is estimated to have diabetes, which is projected to increase to 3.3% by 2045 [4]. However, aggregated data from observational studies suggest even higher estimates at 6.5 (4.7–8.3)% [5]. Thus, the impact of diabetes is felt globally, affecting social, economic, and personal livelihoods [6, 7]. According to the Global Health estimate [8], T2DM is a major cause of disability and death.

The above text sheds light on the mounting responsibility of managing diabetes and the shortcomings of the current standard definition. In Ghana, the majority of individuals with T2DM have poor control [9], and hospitalizations are on the rise annually [10, 11]. A large part of this can be attributed to the current clinical definition, which fails to link diagnosis with underlying pathophysiological factors. Consequently, achieving optimal metabolic control and identifying those who require intensified treatment is a challenge [12-15]. Additionally, the current treatment approach for T2DM is dependent on the availability and affordability of laboratory tests and medications at the time of patient presentation [1, 12]. These limitations have a significant impact on the effectiveness of existing treatment guidelines and contribute to increased morbidity and mortality rates associated with T2DM [9, 16].

Recent studies suggest that personalized treatment strategies for T2DM patients could be more effective in preventing cardiovascular complications[17-19]. To achieve this, it is crucial to efficiently analyze patient data to identify metabolic differences [15, 18]. There is a considerable variance in the T2DM population, which is why new classification algorithms based on clinical and

genetic data are being proposed [3, 13, 19-21]. For example, Ahlqvist *et al.*, [13] discovered five groups of diabetic patients with varying risks of complications through unsupervised grouping of patient data, and these findings have been replicated in other studies [22-24]. This demonstrates that unsupervised learning models have the potential to pinpoint clinically relevant subgroups of T2DM patients.

Thus, in this study, we have focused on the use of unsupervised data-driven cluster analysis to identify subgroups of long-term T2DM patients receiving treatment at a tertiary hospital in Ghana. The main objective of our analysis was to provide baseline evidence of the existence of clinically relevant T2DM subgroups, which can influence the scope of future large-scale research. Thus, we based our clustering analysis on indices that reflect inherent pathophysiology in T2DM, particularly, obesity, insulin resistance, and poor glycemic control. The selected indices included body mass index (BMI), relative fat mass (RFM), glycated hemoglobin (HbA1c) and triglyceride-glucose index (TyG).

In this study, we focused on identifying subgroups of long-term T2DM patients who were receiving treatment at a tertiary hospital in Ghana. We used unsupervised data-driven cluster analysis to achieve this objective. Our main aim was to provide baseline evidence of clinically relevant T2DM subgroups that could influence future large-scale research. We based our clustering analysis on indices that reflect inherent pathophysiology in T2DM such as obesity, insulin resistance, and poor glycemic control. The selected indices included Body Mass Index (BMI), Relative Fat Mass (RFM), Glycated Hemoglobin (HbA1c), and Triglyceride-Glucose Index (TyG).

BMI is a well-characterized and most used adiposity index in daily clinical work in patients with T2DM. Despite the obesity paradox in patients with T2DM [25, 26], elevated BMI has widely been replicated in several studies as a risk factor for both the onset and progression of T2DM [27-30]. The RFM represents a newly proposed adiposity index that overcomes the limitations of BMI by accurately defining obesity defined by body fat-defined obesity [31-33]. In several studies, RFM has been shown to be effective for use in the general practice setting to estimate the risk of both the onset and progression of T2DM [34, 35]. The TyG index is a simple method for determining insulin resistance [36]. Some studies reported that the TyG index is superior to the Homeostatic Model Assessment for Insulin Resistance (HOMA-IR) in predicting T2DM [37, 38]. Moreover, TyG has been used for predicting a variety of CVD types [39-44]. HbA1c is an important indicator of long-term glycemic control that is a reliable biomarker in routine practice for the diagnosis and prognosis of diabetes [45]. Therefore, using these indices together in a cluster analysis provides a cumulative cardiometabolic history of the population with T2DM.

Materials and Methods

Study Setting and Design

According to recent statistics, hospitalization trends among patients with diabetes were markedly high in the Eastern compared to the other regions in Ghana [10]. Patients were recruited from the Eastern Regional Hospital, Koforidua (ERHK) in Ghana from July to October 2021 for this study. The diabetes clinic is held twice a week with an average patient attendance of 140 patients per week [46].

Sample Design and Population

The study design was a hospital-based cross-sectional study. The sampling frame was patients with established T2DM, on treatment (between 1 year and 30 years) and aged 30 to 70 years. We emphasized our sample size calculation on the proposed existence of T2DM subpopulations with differential cardiometabolic risk profiles. According to previous studies [15, 22, 47] the proportion of the minimum subgroup of the T2DM population ranges between 5% and 21%. Based on these estimates, we assumed a maximum permissible limit of a minimum subgroup prevalence of 21%. Hence, we chose a relative precision of $(21/100) * 21 = 4.4\%$ for our sample size calculation. This means that the sample size represents a minimum prevalence of $18.8\% \pm 2.2\%$ (18.8% to 23.2%). At a 95% confidence interval, and based on the above assumptions, the minimum size of T2DM cases required to detect a proportion of 18.8% to 23.2% of the minimum group was 330 patients. Our calculation was based on the following formulae:

$$n = \frac{z^2 \times p(1-p)}{e^2} \quad [48]. \text{ Where } n = \text{sample size, } z = 1.96, p = 21\%, e = 4.4\%.$$

Inclusion and Exclusion Criteria

Patients were included in the study if they met the following criteria: confirmed clinically diagnosed patients with T2DM (ICD-10 code: E11) who have been treated for at least one year, between 30 and 70 years old, do not have comorbid hypertension or a history of cardiovascular disease, including stroke, myocardial infarction, and coronary artery disease. Patients with T2DM who self-managed or were being treated for microvascular complications were excluded from the study. Additionally, patients whose medical records revealed a probable alteration of the metabolic profile due to conditions including peripheral vascular disease (PVD), inflammatory and autoimmune disorders such as rheumatoid arthritis (RA), and other indications of corticosteroid therapy were excluded from this study.

Sampling and Recruitment

Patient recruitment was integrated into routine clinic visits for patients with diabetes at ERHK. We considered that if an average of 140 patients visits per week [46], then 70 diabetes patients are expected for each clinic visit day. Based on this information, we adopted a systematic random sampling approach for patient recruitment on a first-come-first-chosen basis within the 4-month (16 weeks) time window. After the first patient was selected, an estimated sample interval was calculated based on the expected patients per clinic visit: 70 patients by 2 clinic visits by 16 weeks, divided by our estimated sample size ($n = 330$). That is, every seventh patient was selected after the first patient is recruited on a particular visit day (approximately, 10 patients were sampled per clinic visit). Whenever a patient refused or withdraws consent or did not meet the eligibility criteria, the next available patient was recruited.

Data and Sample Collection

A structured questionnaire was used for data collection. The questionnaire had sections for demographic data for patients (current age, sex, family history of DM, past and current history of smoking and alcohol intake), clinical history (age at the time of diagnosis of diabetes, duration of treatment, symptoms, diagnostic information on complication and medication regime), anthropometric measurements and laboratory tests. It is routine practice for patients to fast overnight before visiting the clinic for health examinations. Thus, laboratory measurements are taken in fasting state.

Anthropometry and Blood pressure measurements

For each recruited participant, height and weight were measured using a stadiometer and an electronic weigh balance while in light cloth (OMRON HEALTHCARE Co., Ltd.), respectively. Waist circumference (WC) was measured at the umbilicus level and the maximum gluteal protrusion using a tape measure. BMI was calculated as body weight/height (kg/m^2). RFM was calculated from WC and height:

$$RFM = (64 - (20 * (\text{height}/\text{waist circumference})) + (12 \times \text{ifelse}(\text{sex} == \text{female}, 1, 0))$$

Blood pressure was checked by trained staff nurses and recovered from the patient's folder. In the laboratory, blood samples (2 millilitres in sodium fluoride tube and 3 millilitres in serum separator tubes) were drawn, processed, and stored (-20°C) for subsequent measurements.

Laboratory Measurements

Laboratory analysis of the samples was performed at the ERHK Clinical Chemistry Laboratory. Sodium fluoride whole blood samples were used to estimate HbA1c with a standard A1c Care System (SD BIOSENSOR, Kyonggi-do, Korea). Fasting plasma glucose (FPG) and lipids were estimated using plasma and serum obtained from sodium fluoride and serum separator tubes, respectively, on the SELECTRAPRO M chemistry analyzer (EliTech Group B.V, The Netherlands). We recorded glucose and lipid concentrations in mmol/L and HbA1c as percentages. The lipid profile included high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), triglycerides (Trig), and total cholesterol (T. Chol). Non-HDL-C was estimated as the difference between T. Chol and HDL-C.

Definition of outcome variables

CVD risk outcomes were defined using the following scores and definitions:

- The updated 2008 Framingham Risk Score for the risk of general Atherosclerotic cardiovascular disease (ASCVD) and the risk of individual CVD events (coronary, cerebrovascular, and peripheral arterial disease, and heart failure) was used to define the 10-year risk of heart diseases [49].
- We adopted the criteria of the American Diabetes Association (ADA) criteria [50] and the 2013 American College of Cardiology (ACC) and the American Heart Association (AHA) guidelines to categorize low-density lipoprotein cholesterol: < 3.4 mmol/L (near optimal); 3.4-4.1 mmol/L (borderline high); ≥ 4.2 mmol / L (high).

Based on the above categories, we defined CVD risk phenotypes on an ordinal scale using **Table 1** below.

Table 1: Adapted definition for cardiovascular risk among patients with T2DM.

LDL-c categories	ASCVD risk scores			Key	
	Low risk	Intermediate risk	High risk		
Optimal				Low	
Borderline high					Intermediate
High					

LDL-c: < 3.4 mmol/L (Optimal); 3.4-4.1 mmol/L (Borderline high); ≥ 4.2-4.9 mmol / L (high).
ASCVD risk scores: <10% (low risk); 10-19.9% (intermediate risk); 20% (high risk).

Definitions of other outcome variables

General obesity was defined based on BMI thresholds for overweight (≥ 25 Kg/m²), and obesity (≥ 30 Kg/m²) according to the World Health Organization's criteria [3]. Central obesity was defined using sex and age-specific cut-off values as described in a previous study [51]:

- Males: 20–39 years ≥ 25 ; 40–59 years ≥ 28 ; 60–79 years > 30 .
- Females: 20–39 years ≥ 39 ; 40–59 years ≥ 40 ; 60–79 years > 42

We created categories of triglyceride and HDL-c levels based on the 2013 American College of Cardiology/American Heart Association guidelines [52]: normal triglyceride (triglyceride < 1.7 mmol/L), high triglyceride (triglyceride 1.7 mmol / L), normal HDL-c (HDL-c 1.0 mmol / L for men and 1.3 mmol / L for women), and low HDL-c (HDL-c < 1.0 mmol/L for men and < 1.3 mmol/L for women). We defined metabolic dyslipidemia as high triglyceride and low HDL according to the definition proposed in a previous study [53].

We adopted the Adult Treatment Panel III (ATP III) of the National Cholesterol Education Program (NCEP) Adult Treatment Panel III (ATP III) [54] for the definition of Metabolic syndrome (MetS). That is, if two of the following criteria are met in addition to the presence of T2DM: waist circumference < 102 cm (men) or < 88 cm (women), blood pressure $< 130/85$ and dyslipidaemia (fasting triglyceride level > 1.7 mmol/L and/or HDL-c less than 1.03 mmol/l (men) or 1.29 mmol/l (women)). Severe insulin resistance was defined as TyG levels ≥ 9.0 for women and ≥ 9.2 for men [55]. Poor glycemic control was defined according to the recommendations of the American Diabetes Association: HbA1c $> 7.0\%$ for both men and women [50].

Cluster Analysis

Variables for the clustering model were selected on the basis that the pathophysiological progression to T2DM includes obesity, metabolic syndrome, and glucose dysregulation, due to systemic insulin resistance. The combined impact of these abnormalities imposes glucolipotoxicity leading to eventual β -cell decline with decreased insulin secretion. At this point, the development of cardiovascular complications is inevitable. Thus, we have used BMI and RFM as proxies for central and general obesity, HbA1c for glucose control and β -cell function, and TyG for systemic insulin resistance. Details of the clustering analysis have been attached as separate HTML and Word documents (T2DM_cluster.html and T2DM_analysis.docx). We performed the cluster analysis for men and women data separately to avoid any stratification due to sex-dependent differences. We tested several methods for an optimal number of clusters and performed internal

validation analysis using the average Silhouette width, Dunn index, and connectivity index. We used K-means analysis according to the Hartigan and Wong algorithm [56] to cluster the data. Clustering analysis was performed with the 'NbClust' package [57] on R software (Version 4.3.3).

Data Management and Statistical Analysis

All data obtained from the patients were managed on Microsoft Spreadsheet: double checked after entry for precision and reliability and locked before usage. The data required for analysis were extracted in a CSV file. Data analysis was performed using the R software (Version 4.3.3) and the Statistical Package for Social Sciences (SPSS version 25). Categorical data were presented as counts and corresponding percentages. The chi-square test was used to compare data differences in categorical variables between the T2DM subgroups. Summary statistics of continuous variables were presented as mean and standard deviations (for normally distributed data) and compared using a t-test and one-way analysis of variance when required. Skewed data was presented as median and interquartile ranges and differences between subgroups computed with Mann-Whitney or Kruskal Wallis test. We used alluvial diagrams to illustrate the flow of cardiometabolic risk factors within each T2DM group. To understand the impact of variables on the existence of clinical subgroups and cardiovascular risk outcomes, we used multiple regression models with a 'logit' link and reported p-values less than 0.05 as statistically significant.

Regulatory and Ethical Approval

The study was carried out according to the ethical principles of the Declaration of Helsinki and the applicable national ethical regulatory requirements. The Human Research, Publication and Ethics Committee (CHRPE) of Kwame Nkrumah University of Science and Technology (KNUST) granted ethical clearance (Ref: CHRPE/AP/217/21) for this study. Additionally, permission was requested and received from the ERHK to use their medical resources for the recruitment of research participants. Under the CHRPE study participant information policies, written information was provided to participants and their written consent was obtained prior to admission to the study. Participation was completely voluntary and all patients with T2DM received equal participation rights.

Results

Demographic, clinical, and laboratory characteristics of female and male T2DM groups

The characteristics of male T2DM clusters are shown in **Table 2**. Two (2) clusters were obtained for male T2DM data. Cluster 1 constituted 62.1% of the male data and cluster 2 constituted 37.9%. Clusters 1 and 2 did not differ in terms of their demographic, medication prescription, and symptom history. Regarding laboratory profile, Cluster 2 had significantly higher levels of triglycerides (p-value = 0.001), total cholesterol (p-value = 0.006) and non-HDL-c (p-value = 0.022) than Cluster 1 (**Table 2**). The waist circumference was elevated in group 2 than in group 1 (p-value <0.001, **Table 2**).

The cluster analysis resulted in three distinct clusters within the female data from T2DM: 36.1% for cluster 1 and 31.9% each for clusters 2 and 3 (**Table 3**). Sociodemographic background did not differ between the three clusters (p-value >0.05). A significantly higher proportion of cluster 2 than cluster 1 had a history of taking Metformin (98.6% vs. 84.2%), while a significantly higher proportion of cluster 1 than cluster 2 (13.2% vs. 2.9%) took Thiazolidinedione (p-value <0.05). Furthermore, a higher proportion of patients in Group 2 and 3 than in Group 1 reported symptoms of easy fatiguability (p-value = 0.03, **Table 3**). In terms of its laboratory profile, cluster 2 had significantly higher fasting blood glucose (p-value <0.001), total cholesterol (p-value < 0.001), non-HDL-c (p-value < 0.001) and diastolic blood pressure (p-value = 0.004) than clusters 1 and 3. All three groups were distinct in terms of their mean triglyceride (p-value < 0.001) and waist circumference (p-value < 0.001). Clusters 2 and 3 had significantly elevated mean LDL-c compared to cluster 1 (p-value <0.001).

Table 2 Demographic, clinical and laboratory characteristics of male subgroups of T2DM

Variable	Male (n = 103)		p-value
	Cluster 1 N = 64(62.1%)	Cluster 2 N= 39(37.9%)	
Education			0.092
Non-formal	5 (7.8)	0	
Basic	19 (29.7)	7 (17.9)	
Secondary	24 (37.5)	16 (41.0)	
Tertiary	16 (25.0)	16 (41.0)	
Family history of DM			
No	29 (45.3)	14 (35.9)	
Yes	35 (54.7)	25 (64.1)	0.347
Currently smoking			
No	59 (92.2)	38 (97.4)	
Yes	5 (7.8)	1 (2.6)	0.404
Alcohol intake			
No	49 (76.6)	32 (84.2)	
Yes	15 (23.4)	6 (15.8)	0.451
Medication history			
Metformin	53 (88.3)	32 (86.5)	0.788
Insulin	28 (46.7)	14 (37.8)	0.394
Thiazolidinediones	1 (1.7)	2 (5.4)	0.556
Sulphonyl urea	50 (83.3)	34 (91.9)	0.229
DPP-4	2 (3.3)	3 (8.1)	0.366
Statins	16 (25.0)	16 (41.0)	0.088
Symptom history			
Easy fatiguability	8 (12.7)	2 (5.1)	0.310
Breathlessness	1 (1.6)	0	n/a
Chest pain	8 (12.5)	3 (7.7)	0.527
Palpitations	11 (17.2)	2 (5.1)	0.124
Metabolic profiles[^]			
FBS (mmol/L)	8.40±2.96	8.96±2.45	0.322
Trig (mmol/L)	0.88±0.34	1.26±0.62	0.001
HDL-c (mmol/L)	1.37±0.48	1.39±0.37	0.861
T. Chol (mmol/L)	4.59±0.99	5.11±1.16	0.016
LDL-c (mmol/L)	3.04±0.86	3.45±1.16	0.063
Non-HDL-c (mmol/L)	3.22±0.94	3.73±1.27	0.022
Anthropometric indices[^]			
WC (cm)	83.72±6.40	99.29±6.17	<0.001
Systolic BP (mmHg)	119.58±13.32	124.77±9.63	0.024
Diastolic BP (mmHg)	71.47±8.68	73.44±7.37	0.241
Age (years)	51.77±9.20	51.15±8.39	0.736
Age at T2DM onset (years)	44.03±9.26	42.95±6.51	0.589
Duration of T2DM (years) [#]	6.0 (3.0-13.0)	7.0 (3.0-13.0)	0.586

Values are presented as frequencies and compared using chi-square test, unless otherwise specified. Variable with [^] symbol are presented as mean and standard deviation and compared using student t-test. Variables with # symbol are presented as median (interquartile ranges) and compared using Mann-Whitney test. DM- diabetes mellitus; BP- blood pressure; DPP-4- inhibitors of dipeptidyl peptidase 4; WC- waist circumference.

Table 3 Demographic, clinical, and laboratory characteristics of the female subgroups of T2DM.

Variable	Male (n = 238)			P-value
	Cluster 1 N = 86(36.1%)	Cluster 2 N= 76(31.9%)	Cluster 3 N = 76(31.9%)	
Education				0.693
Non-formal	7 (8.2)	7 (9.2)	6 (7.9)	
Basic	33 (38.8)	24 (31.6)	24 (31.6)	
Secondary	42 (49.4)	37 (48.7)	40 (52.6)	
Tertiary	3 (3.5)	8 (10.5)	6 (7.9)	
Family history of DM				0.162
No	39 (45.3)	25 (32.9)	25 (32.9)	
Yes	47 (54.7)	51 (67.1)	51 (67.1)	
Currently smoking				n/a
No	86 (100.0)	76 (100.0)	76 (100.0)	
Yes	0	0		
Alcohol intake				0.822
No	83 (96.5)	72 (94.7)	72 (94.7)	
Yes	3 (3.5)	4 (5.3)	4 (5.3)	
Medication				
Metformin	64 (84.2) ^a	68 (98.6) ^b	65 (92.9) ^{a,b}	0.007
Insulin	32 (42.1)	20 (29.0)	20 (28.6)	0.141
Thiazolidinediones	10 (13.2) ^a	2 (2.9) ^b	6 (8.6) ^{a,b}	0.083
Sulphonyl urea	63 (82.9)	58 (84.1)	55 (78.6)	0.674
DPP-4	0	2 (2.9)	0	n/a
Statins	37 (43.0)	35 (46.1)	35 (46.1)	0.903
Symptom history				
Easy fatiguability	11 (12.8) ^a	19 (25.0) ^b	22 (28.9) ^b	0.03
Breathlessness	5 (6.0)	6 (8.0)	4 (5.3)	0.783
Chest pain	9 (10.8)	5 (6.6)	4 (5.4)	0.400
Palpitations	10 (11.6)	4 (5.3)	10 (13.2)	0.227
Metabolic profiles [^]				
FBS (mmol/L)	7.83±3.02 ^a	10.35±3.16 ^b	8.05±2.77 ^a	<0.001
Trig (mmol/L)	0.81±0.30 ^a	1.49±0.49 ^b	1.05±0.32 ^c	<0.001
HDL-c (mmol/L)	1.47±0.35	1.43±0.48	1.33±0.50	0.126
T. Chol (mmol/L)	4.69±1.11 ^a	5.74±1.42 ^b	4.72±1.10 ^a	<0.001
LDL-c (mmol/L)	2.92±0.95 ^a	3.69±1.30 ^b	3.05±1.00 ^b	<0.001
Non-HDL-c (mmol/L)	3.22±1.05 ^a	4.31±1.50 ^b	3.40±1.07 ^a	<0.001
Anthropometric indices[^]				
WC (cm)	86.59±6.30 ^a	96.12±7.50 ^b	104.39±9.14 ^c	<0.001
Systolic BP (mmHg)	119.44±11.12	119.39±10.20	115.71±11.34	0.052
Diastolic BP (mmHg)	71.66±6.44 ^a	75.08±7.74 ^b	71.99±6.74 ^a	0.004
Age (years)	50.85±8.73	50.28±7.42	50.12±7.78	0.829
Age at T2DM onset (years)	41.81±8.56	42.41±6.82	43.03±7.56	0.612
Duration of T2DM (years) [#]	7.0 (5.0-12.0)	6.0 (4.0-12.0)	6.0 (3.0-10.5)	0.067

Values are presented as frequencies and compared using chi-square test, unless otherwise specified. Variable with [^] symbol are presented as mean and standard deviation and compared using student t-test. Variables with # symbol are presented as median (interquartile ranges) and compared using Mann-Whitney test. DM- diabetes mellitus; BP- blood pressure; DPP-4- inhibitors of dipeptidyl peptidase 4; WC- waist circumference. Superscript alphabets (a, b, c) represent differences between the three cluster according to multiple testing with Bonferroni's correction. Clusters with the same superscript alphabet do not differ significantly.

Features of the female and male T2DM clusters.

As shown in **Fig 1a**, the three groups were distinct in terms of their median levels of BMI, RFM, HbA1c and TyG. Clusters 2 and 3 had characteristically distinct elevated median lipid accumulation product index compared to Cluster 1 (p-value <0.01). Cluster 2 had a significantly elevated median coronary and atherogenic risk score than Cluster 1 but was comparable to Cluster 3.

As shown in **Fig 1b**, Cluster 2 from male T2DM data had distinctively high BMI and RFM compared to Cluster 1 (p-value <0.001). Again, the lipid accumulation product index and TyG for cluster 2 were significantly elevated compared to cluster 1 (p-value <0.001). However, both clusters 1 and 2 were comparable in terms of their HbA1c and ASCVD risk scores (p-value >0.05,).

Figure 1. Characteristics of the (a) female and (b) male T2DM groups. *The Mann-Whitney test was used to compute probability values. Asymptomatic probability values (2-sided tests) are indicated with the * symbol. *** P-value <0.001; ** p-value <0.01, *p-value <0.05, ns p-value >0.05. HbA1c = glycated hemoglobin, TyG = triglyceride-glucose index; BMI = body mass index; RFM = relative fat mass index; LAP= lipid accumulation product index; ascvd-10y-frs = risk of ASCVD 2008 (model with lipid labs); ascvd-10y-frs-simple = risk of ASCVD 2008 (model with BMI).*

Cardiometabolic profile of female patients with T2DM

The cardiometabolic profile of the population with T2DM is shown in **Table 4**. Among female patients with T2DM, Cluster 3 was distinctively characterized by an excess prevalence of general (80.3%) and central (100.0%) obesity than Clusters 1 and 2. Cluster 2 had a distinct characteristic of a high prevalence of poor glycemic control (96.1%), severe insulin resistance (78.9%), and hypertriglyceridemia (39.5%) compared to clusters 1 and 3. The prevalence of MetS syndrome was higher in Clusters 2 (60.5%) and 3 (61.8%) compared to Cluster 1 (24.4%). Furthermore, higher proportions of Cluster 2 than of Clusters 3 and 1 had high phenotype of cardiovascular disease (p-value =0.001). The prevalence of metabolic dyslipidemia was 10.5% in Cluster 2 but was absent in Clusters 1 and 3 (**Table 4**).

Fig 2 provides a visual representation of the different cardiometabolic phenotypes present in patients with T2DM. For example, in cluster 1 of female patients with T2DM, we can observe two phenotypes: normal weight with normal cardiometabolic profile and overweight with poor glycemic control, but a normal cardiometabolic profile. In Cluster 2 of female patients with T2DM, two phenotypes are obvious: obesity with a poor cardiometabolic profile and a high risk of CVD incidence and overweight with a poor cardiometabolic profile but low risk of CVD incidence. In

Cluster 3 of women with T2DM, we can observe obese individuals with controlled insulin resistance with or without poor glycaemic status but low CVD risk.

Figure 2. Spatial view of the cardiometabolic and cardiovascular risk profile of the female T2DM clusters. Gen obesity: 1 = normal weight, 2 = overweight, 3 = obese; insulin resistance: 1 = optimal, 2 = severe; glycaemic control: 1 = normal, 2 = poor; blood pressure: 1 = normal, 2 = high; MetS: 1 = absent, 2 = present; CVD risk: 1 = low, 2 = intermediate, 3 = high.

Table 4. Cardiometabolic risk profile of male and female patients with T2DM

Variable	Male (n = 103)			Female (n = 235)			
	Cluster 1 N = 64	Cluster 2 N = 39	p-value	Cluster 1 N = 86	Cluster 2 N = 76	Cluster 3 N = 76	p-value
Cardiometabolic risk							
General obesity	0.0	28.2%	<0.001	2.3% ^a	25.0% ^b	80.3% ^c	<0.001
Central obesity	7.8%	79.5%	<0.001	26.7% ^a	81.6% ^b	100.0% ^c	<0.001
High triglycerides	4.7%	17.9%	0.027	1.2% ^a	39.5% ^b	5.3% ^a	<0.001
Low HDL-c	64.1%	38.5%	0.011	31.4%	36.8%	44.7%	0.214
Metabolic dyslipidemia	0.0	0.0		0.0	10.5%	0.0	<0.001
High blood pressure	29.7%	30.8%	0.908	19.8%	21.1%	21.1%	0.973
Poor glycaemic control	54.7%	51.3%	0.737	48.8% ^a	96.1% ^b	31.6% ^a	<0.001
Severe insulin resistance	7.8%	28.2%	0.006	10.7% ^a	78.9% ^b	23.7% ^a	<0.001
MetS	14.1%	10.3%	0.573	24.4%	60.5%	61.8%	<0.001
ASCVD-10-yr risk							
≤10% increased risk	73.4%	71.8%	0.860	100.0	100.0	100.0	n/a
10%-20% increased risk	18.8%	28.2%	0.269	0.0	0.0	0.0	
>20% increased risk	7.8%	0.0	0.075	0.0	0.0	0.0	
ASCVD-10-yr risk[§]							0.521
≤10% increased risk	31.1%	20.5%	0.243	59.3%	61.8%	57.3%	
10%-20% increased risk	32.8%	43.6%	0.273	34.9%	35.5%	41.3%	
>20% increased risk	35.9%	35.9%	1.00	5.8%	2.6%	1.3%	
CVD risk category							
Low	59.4%	28.2%	0.002	80.2% ^a	59.2% ^b	77.3% ^a	
Intermediate	20.3%	46.2%	0.006	8.1%	1.3%	4.0%	
High	20.3%	25.6%	0.533	11.6% ^a	39.5% ^b	18.7% ^a	<0.001

MetS metabolic syndrome; ASCVD risk = Framingham 2008 10-year ASCVD risk (model with lipid labs); ASCVD-10-year risk[§] = Framingham 2008 10-year ASCVD risk (model with BMI). Superscript alphabets (a, b, c) represent differences between the three cluster according to multiple testing with Bonferroni's correction. Clusters with the same superscript alphabet do not differ significantly.

Cardiometabolic profile of male patients with T2DM

Among male patients with T2DM (Table 4), cluster 2 was distinctively characterized by a higher prevalence of hypertriglyceridemia, general obesity, central obesity, and severe insulin resistance (17.9%, 28.2%, 79.5% and 28.2%, respectively) than cluster 1 (4.7%, 0%, 7.8% and 7.8%, respectively). Cluster 1 was characterised by a higher prevalence of hypoalphalipoproteinemia (64.1% vs. 38.5%) than cluster 2. Metabolic dyslipidemia was absent in both groups, while equivalent proportions of groups 1 and 2 had high blood pressure, poor glycemic control, metabolic syndrome, and high risk of cardiovascular disease. Furthermore, higher proportions of Cluster 2 than of Cluster 1 had an intermediate risk of cardiovascular disease (p-value = 0.006).

As shown in Fig 3, Cluster 1 of male patients with T2DM, three phenotypes were observed: individuals with normal cardiometabolic and low CVD risk profile; those with poor glycemic control but normal cardiometabolic profile, those with poor glycemic control with intermediate-high risk of CVD incidence. Within cluster 2 of male patients with T2DM diabetes, three phenotypes were observed: patients with overweight/obesity, severe insulin resistance, poor glycemic controls, and high risk of CVD incidence; overweight/obesity patients with poor glycemic control and intermediate risk of CVD; and overweight/obesity patients with normal cardiometabolic profile and low CVD risk profile (**Fig 3**).

Figure 3. Spatial view of the cardiometabolic and cardiovascular risk profile of the T2DM male clusters. *Gen obesity: 1 = normal weight, 2 = overweight, 3 = obese; insulin resistance: 1 = optimal, 2 = severe; glycemic control: 1 = normal, 2 = poor; blood pressure: 1 = normal, 2 = high; MetS: 1 = absent, 2 = present; CVD risk: 1= low, 2 = intermediate, 3 = high, 4 = very high.*

Table 5 shows the association between cardiometabolic and cardiovascular risk outcomes and T2DM clusters. Among female patients with T2DM, Cluster 2 was associated with increased high risk of CVD (OR = 5.34, 95% CI: 2.11-13.55) and MetS (OR =7.07; 95%CI: 3.24-15.42). Furthermore, cluster 3 was significantly associated with an increased risk of developing a metabolic syndrome.

Among male patients with T2DM, Cluster 2 was also associated with increased intermediate (OR = 21.78, 95% CI 4.17-113.78) and high risk of CVD (OR = 6.84, 95% CI 1.455-32.116), but not MetS and high blood pressure.

Table 5 Regression analysis showing the association between CVD risk outcomes, T2DM clusters, and other covariates.

Risk Outcomes	Variables	P-value	Exp(B)	Lower Bound	Upper Bound
Female T2DM patients					
Intermediate CVD risk [^]	Age	<0.001	1.28	1.12	1.46
	Cluster 2	0.402	0.39	0.04	3.59
	Cluster 3	0.818	0.83	0.18	3.95
High CVD risk [^]	Age	<0.001	1.12	1.06	1.18
	Cluster 2	<0.001	5.34	2.11	13.55
	Cluster 3	0.204	1.88	0.71	5.00
MetS [#]	Age	<0.001	1.08	1.038	1.124
	Cluster 2	<0.001	7.066	3.237	15.424
	Cluster 3	<0.001	5.944	2.755	12.823
High blood pressure [#]	Age	0.001	1.079	1.03	1.13
Male T2DM Patients					
Intermediate CVD risk [^]	Age	<0.001	1.339	1.185	1.514
	Cluster 2	<0.001	21.775	4.167	113.783
High CVD risk [^]	Age	<0.001	1.229	1.104	1.369
	Cluster 2	0.015	6.836	1.455	32.116
MetS [#]	Insulin therapy	0.045	12.325	1.061	143.155
High blood pressure [#]	Alcohol intake	0.034	4.226	1.115	16.02

All regression models included prescription of medications, patient age, duration of disease, smoking, and alcohol intake status, family history of diabetes and T2DM groups. The symbols represent multinomial regression analysis and # represent logistic regression analysis. In all model T2DM clusters were treated as an entry term, whereas all other variables were conditioned to enter the model using the stepwise forward conditional method.

Discussion

In this study, we have characterised subgroups of T2DM marked by obesity, uncontrolled insulin resistance, and poor glycemic control, similar to what has been previously described [20, 58-60]. Therefore, we identified three distinct phenotypes among female patients with T2DM: obesity-

related phenotype with intermediate CVD risk, severe insulin resistance phenotype with high CVD risk, and 'normal weight-improved insulin resistance' phenotype with low CVD risk. Within the female clusters, there was marked heterogeneity in terms of glycemic control, blood pressure, and metabolic syndrome status. Like the female T2DM patients, two phenotypes were observed in the male population: obesity-related phenotype with severe insulin resistance; and normal weight phenotype with improved insulin sensitivity. Like the female clusters, heterogeneity was observed in the male clusters in terms of glycaemic control, MetS and CVD risk.

Therefore, the T2DM subgroups observed in our study overlap most of the clinically relevant subgroups identified among newly diagnosed T2DM patients [3, 13, 19-21]. In a recent publication including 541 Ghanaians with adult-onset diabetes mellitus reported subgroups including obesity-related (73%), insulin-resistant (5%) and insulin-deficient (7%), age-related (10%) and autoimmune-related (5%) were reported [47]. Similar phenotypes including the severe insulin resistance, severe insulin-deficient and obesity-related phenotypes have been replicated using T2DM clinical data obtained from North America, Canada, and Europe [22-24].

The emphasis of many of the recently proposed clustering algorithms for T2DM data assumes that individual patients present a combination of defects in different metabolic pathways, placing them on a quantitative spectrum of metabolic disturbances. Using data from the Outcome Reduction with the Initial Glargine Intervention (ORIGIN) trial [58], Individual Support & Resources for Diabetes (INSPIRED) study [61] and Swedish All New Diabetics in Scania cohort [13], five replicable clusters have been defined for newly diagnosed patients with T2DM: severe autoimmune diabetes, severe insulin-deficient diabetes, severe insulin resistant diabetes, mild obesity-related diabetes and mild age-related diabetes. These classifications bring out somewhat measurable differences in clinical features between patients and have promising benefits for the practice of precision medicine and clinical research. Using these features in a trial, Dennis *et al.*, [59] reported that precision medicine in T2DM will have a broader clinical utility if it is based on an approach of using specific phenotypic measures to predict specific outcomes. Similarly Bancks *et al.*, [62] employed these subgrouping on a secondary analysis of clinical data and presented the benefits subgroup to select intensive lifestyle interventions for patients. From the pool of literature [60] it is uncertain, which variables and algorithms are appropriate to generate clinically significant T2DM subtypes. It is also uncertain whether the subgroupings identified for newly diagnosed patients will be applicable in patients with longer duration of diabetes. Furthermore, to what extent these groups can be applicable to predict the prognosis in patients with a longer duration of

diabetes, and the chances that patients can move from one group to another in the progression of T2DM are unknown.

One thing is certain however, that β -Cell function decline, loss of glycemic control, and weight gain are inherent progressive abnormalities in T2DM, even with treatment [63-66]. Our findings reflect these observations by presenting that the individual with T2DM who has the obesity phenotype and uncontrolled insulin resistance phenotype, has a substantial risk of MetS and develops CVD events. In the report from the ORIGIN trial [58] patients with severe insulin-resistant diabetes subtype presented with a higher incidence of chronic kidney disease and macroalbuminuria, but also benefited more from receiving glargine. Similarly, Xiong *et al.*, [67] reported that patient with a severe insulin-resistant diabetes subtype have increased risk of diabetic retinopathy, diabetic peripheral neuropathy, and hypertension. These findings, together with our present report, indicate that worse outcomes in T2DM may be associated with progressive β -Cell failure not meeting the increasing demand of insulin resistance. It also emphasizes the logic to consider insulin secretion and sensitivity indices in the routine clinical classification and treatment decisions of patients with T2DM.

One of the strengths of the study is that the variables used for clustering were chosen based on their clinical utility, potential pathophysiological proximity, and previous use. The primary limitation is the sample size, which reduces adequate power to back the comparisons between the subgroups. Another limitation is the varied treatment regimens received by the study enrolled patients which may have a significant effect on the measured parameters. However, effect of such confounders was minimized employing adjusted models. Although findings from this study are based on observation, it begs the discourse to be validated with more robust designs.

Conclusion

The study suggests that there is a wide range of patients with T2DM in Ghana, and clinicians must consider individualized care and treatment options based on the patient's specific characteristics. It further highlights the importance of considering insulin secretion and sensitivity indices in clinical decisions regarding patients with T2DM which may be accounting for unexplained high levels of uncontrolled diabetes and related hospitalization rates. Overall, the study underscores the need for continued research and the development of effective treatment strategies for patients with T2DM in the sub-Saharan population.

Acknowledgments

We would like to thank the staff at the Eastern Regional Hospital, Koforidua, diabetes clinic for their support.

References

1. WHO. Diabetes. World Health Organization; 2022 16 September 2022.
2. Diagnosis and classification of diabetes mellitus. *Diabetes care*. 2010;33 Suppl 1(Suppl 1):S62-9.
3. Lugner M, Gudbjörnsdóttir S, Sattar N, Svensson A-M, Miftaraj M, Eeg-Olofsson K, et al. Comparison between data-driven clusters and models based on clinical features to predict outcomes in type 2 diabetes: nationwide observational study. 2021;64(9):1973-81.
4. Diabetes around the world in 2021 [Internet]. International Diabetes Federation. 2021. Available from: <https://diabetesatlas.org/>.
5. Asamoah-Boaheng M, Sarfo-Kantanka O, Tuffour AB, Eghan B, Mbanaya JCJH. Prevalence and risk factors for diabetes mellitus among adults in Ghana: a systematic review and meta-analysis. 2019;11(2):83-92.
6. Jalilian H, Heydari S, Imani A, Salimi M, Mir N, Najafipour F. Economic burden of type 2 diabetes in Iran: A cost-of-illness study. 2023;6(2):e1120.
7. Soares Andrade CA, Shahin B, Dede O, Akpeji AO, Ajene C-L, Albano Israel FE, et al. The burden of type 2 diabetes mellitus in states of the European Union and United Kingdom at the national and subnational levels: A systematic review. 2023;24(9):e13593.
8. Global Health Estimates: Life expectancy and leading causes of death and disability [Internet]. World Health Organization. 2019 [cited 11 January 2023]. Available from: <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates>.
9. Mobula LM, Sarfo FS, Carson KA, Burnham G, Arthur L, Ansong D, et al. Predictors of glycemic control in type-2 diabetes mellitus: evidence from a multicenter study in Ghana. 2018;1:1-8.
10. Narh CT, Der JB, Ofosu A, Blettner M, Wollschlaeger D. Trends in hospitalization of patients with diabetes mellitus in Ghana from 2012 to 2017 with predictions to 2032. *International Health*. 2021;14(6):588-96.
11. Sarfo-Kantanka O, Sarfo FS, Oparebea Ansah E, Eghan B, Ayisi-Boateng NK, Acheamfour-Akowuah EJPO. Secular trends in admissions and mortality rates from diabetes mellitus in the central belt of Ghana: a 31-year review. 2016;11(11):e0165905.
12. MoH-GNDP. Standard Treatment Guidelines. Seventh Edition (7th) ed. Republic of Ghana: Yamens Press Limited; 2017. p. 1-708.
13. Ahlqvist E, Storm P, Käräjämäki 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. 2018;6(5):361-9.
14. Danquah I, Bedu-Addo G, Terpe KJ, Micah F, Amoako YA, Awuku YA, et al. Diabetes mellitus type 2 in urban Ghana: characteristics and associated factors. *BMC public health*. 2012;12:210.

15. Ahima RS. Editorial: Rethinking the definition of diabetes for precision medicine. *Molecular endocrinology* (Baltimore, Md). 2015;29(3):335-7.
16. Djonor SK, Ako-Nnubeng IT, Owusu EA, Akuffo KO, Nortey P, Agyei-Manu E, et al. Determinants of blood glucose control among people with Type 2 diabetes in a regional hospital in Ghana. 2021;16(12):e0261455.
17. Brownlee MJ. The pathobiology of diabetic complications: a unifying mechanism. 2005;54(6):1615-25.
18. Gloyn AL, Drucker DJ. Precision medicine in the management of type 2 diabetes. 2018;6(11):891-900.
19. Prasad RB, Groop L. Precision medicine in type 2 diabetes. 2019;285(1):40-8.
20. Xing L, Peng F, Liang Q, Dai X, Ren J, Wu H, et al. Clinical characteristics and risk of diabetic complications in data-driven clusters among type 2 diabetes. 2021;12.
21. Del Prato S. Heterogeneity of diabetes: heralding the era of precision medicine. 2019;7(9):659-61.
22. Safai N, Ali A, Rossing P, Ridderstråle M. Stratification of type 2 diabetes based on routine clinical markers. 2018;141:275-83.
23. Zaharia OP, Strassburger K, Strom A, Bönhof GJ, Karusheva Y, Antoniou S, et al. Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: a 5-year follow-up study. 2019;7(9):684-94.
24. Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT. Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data. 2019;7(6):442-51.
25. Tobias DK, Pan A, Jackson CL, O'Reilly EJ, Ding EL, Willett WC, et al. Body-mass index and mortality among adults with incident type 2 diabetes. *The New England journal of medicine*. 2014;370(3):233-44.
26. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Li W, Johnson J, et al. Body mass index and the risk of all-cause mortality among patients with type 2 diabetes mellitus. *Circulation*. 2014;130(24):2143-51.
27. Karin A, Jon E, Martin A, Lena B, Martin L, Naveed S, et al. Body mass index in adolescence, risk of type 2 diabetes and associated complications: A nationwide cohort study of men. *eClinicalMedicine*. 2022;46.
28. Teufel F, Seiglie JA, Geldsetzer P, Theilmann M, Marcus ME, Ebert C, et al. Body-mass index and diabetes risk in 57 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data in 685 616 adults. *Lancet* (London, England). 2021;398(10296):238-48.
29. Edqvist J, Rawshani A, Adiels M, Björck L, Lind M, Svensson AM, et al. Contrasting Associations of Body Mass Index and Hemoglobin A1c on the Excess Risk of Acute Myocardial Infarction and Heart Failure in Type 2 Diabetes Mellitus. *Journal of the American Heart Association*. 2019;8(24):e013871.
30. Jayedi A, Soltani S, Motlagh SZ, Emadi A, Shahinfar H, Moosavi H, et al. Anthropometric and adiposity indicators and risk of type 2 diabetes: systematic review and dose-response meta-analysis of cohort studies. *BMJ* (Clinical research ed). 2022;376:e067516.

31. Woolcott OO, Bergman RN. Relative fat mass (RFM) as a new estimator of whole-body fat percentage — A cross-sectional study in American adult individuals. *Scientific reports*. 2018;8(1):10980.
32. Woolcott OO, Seuring T. Temporal trends in obesity defined by the relative fat mass (RFM) index among adults in the United States from 1999 to 2020: a population-based study. *2023;13(8):e071295*.
33. Guzmán-León AE, Velarde AG, Vidal-Salas M, Urquijo-Ruiz LG, Caraveo-Gutiérrez LA, Valencia MEJPO. External validation of the relative fat mass (RFM) index in adults from north-west Mexico using different reference methods. *2019;14(12):e0226767*.
34. Jambart I, Sleilaty G, Jambart S, Medlej R, Halaby GJJoE, Disorders. Relative Fat Mass, a new tool for measuring obesity, could predict the complications of type 2 diabetes. *2021;5(3):2640-1045*.
35. Suthahar N, Wang K, Zwartkruis VW, Bakker SJL, Inzucchi SE, Meems LMG, et al. Associations of relative fat mass, a new index of adiposity, with type-2 diabetes in the general population. *European journal of internal medicine*. 2023;109:73-8.
36. Tahapary DL, Pratisthita LB, Fitri NA, Marcella C, Wafa S, Kurniawan F, et al. Challenges in the diagnosis of insulin resistance: Focusing on the role of HOMA-IR and Tryglyceride/glucose index. *Diabetes & metabolic syndrome*. 2022;16(8):102581.
37. Park HM, Lee HS, Lee YJ, Lee JH. The triglyceride-glucose index is a more powerful surrogate marker for predicting the prevalence and incidence of type 2 diabetes mellitus than the homeostatic model assessment of insulin resistance. *Diabetes research and clinical practice*. 2021;180:109042.
38. Lee SH, Kwon HS, Park YM, Ha HS, Jeong SH, Yang HK, et al. Predicting the development of diabetes using the product of triglycerides and glucose: the Chungju Metabolic Disease Cohort (CMC) study. *PloS one*. 2014;9(2):e90430.
39. Alizargar J, Bai CH, Hsieh NC, Wu SV. Use of the triglyceride-glucose index (TyG) in cardiovascular disease patients. *Cardiovascular diabetology*. 2020;19(1):8.
40. Li H, Zuo Y, Qian F, Chen S, Tian X, Wang P, et al. Triglyceride-glucose index variability and incident cardiovascular disease: a prospective cohort study. *Cardiovascular diabetology*. 2022;21(1):105.
41. Tao LC, Xu JN, Wang TT, Hua F, Li JJ. Triglyceride-glucose index as a marker in cardiovascular diseases: landscape and limitations. *Cardiovascular diabetology*. 2022;21(1):68.
42. Wang X, Xu W, Song Q, Zhao Z, Meng X, Xia C, et al. Association between the triglyceride-glucose index and severity of coronary artery disease. *Cardiovascular diabetology*. 2022;21(1):168.
43. Xu L, Wu M, Chen S, Yang Y, Wang Y, Wu S, et al. Triglyceride-glucose index associates with incident heart failure: A cohort study. *Diabetes & metabolism*. 2022;48(6):101365.
44. Anto EO, Frimpong J, Boadu WIO, Korsah EE, Tamakloe VCKT, Ansah E, et al. Cardiometabolic syndrome among general adult population in Ghana: The role of lipid accumulation product, waist circumference-triglyceride index, and triglyceride-glucose index as surrogate indicators. *2023;6(7):e1419*.
45. Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of HbA1c Test in Diagnosis and Prognosis of Diabetic Patients. *Biomarker insights*. 2016;11:95-104.

46. Amon SK, Aikins MKSJDM. Economic burden of type 2 diabetes mellitus complications among patients in the eastern region of Ghana: A descriptive cross-sectional cost-of-illness study. 2017;7(5):367-76.
47. Danquah I, Mank I, Hampe CS, Meeks KAC, Agyemang C, Owusu-Dabo E, et al. Subgroups of adult-onset diabetes: a data-driven cluster analysis in a Ghanaian population. *Scientific reports*. 2023;13(1):10756.
48. Charan J, Biswas T. How to calculate sample size for different study designs in medical research? *Indian journal of psychological medicine*. 2013;35(2):121-6.
49. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General Cardiovascular Risk Profile for Use in Primary Care. 2008;117(6):743-53.
50. Standards of Medical Care in Diabetes-2016 Abridged for Primary Care Providers. *Clinical diabetes* : a publication of the American Diabetes Association. 2016;34(1):3-21.
51. Jelavic MM, Babic Z, Pintaric H. Interrelation between the relative fat mass index and other obesity indices in predicting clinical severity and prognosis of acute myocardial infarction. *Postepy w kardiologii interwencyjnej = Advances in interventional cardiology*. 2020;16(2):198-201.
52. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Journal of the American College of Cardiology*. 2014;63(25 Pt B):2889-934.
53. Kaze AD, Santhanam P, Musani SK, Ahima R, Echouffo-Tcheugui JB. Metabolic Dyslipidemia and Cardiovascular Outcomes in Type 2 Diabetes Mellitus: Findings From the Look AHEAD Study. 2021;10(7):e016947.
54. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation*. 2005;112(17):2735-52.
55. Unger G, Benozzi SF, Perruzza F, Pennacchiotti GL. Triglycerides and glucose index: A useful indicator of insulin resistance. *Endocrinología y Nutrición (English Edition)*. 2014;61(10):533-40.
56. Ja H. A k-means clustering algorithm. *The Journal of the Royal Statistical Society, Series C*. 1979;28:100-8.
57. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an R package for determining the relevant number of clusters in a data set. *Journal of statistical software*. 2014;61:1-36.
58. Pigeyre M, Hess S, Gomez MF, Asplund O, Groop L, Paré G, et al. Validation of the classification for type 2 diabetes into five subgroups: a report from the ORIGIN trial. *Diabetologia*. 2022;65(1):206-15.
59. Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT. Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data. *The Lancet Diabetes & Endocrinology*. 2019;7(6):442-51.
60. Sarría-Santamera A, Orazumbekova B, Maulenkul T, Gaipov A, Atageldiyeva K. The Identification of Diabetes Mellitus Subtypes Applying Cluster Analysis Techniques: A Systematic Review. 2020;17(24):9523.

61. Anjana RM, Baskar V, Nair ATN, Jebarani S, Siddiqui MK, Pradeepa R, et al. Novel subgroups of type 2 diabetes and their association with microvascular outcomes in an Asian Indian population: a data-driven cluster analysis: the INSPIRED study. 2020;8(1):e001506.
62. Bancks MP, Chen H, Balasubramanyam A, Bertoni AG, Espeland MA, Kahn SE, et al. Type 2 Diabetes Subgroups, Risk for Complications, and Differential Effects Due to an Intensive Lifestyle Intervention. *Diabetes Care*. 2021;44(5):1203-10.
63. Kahn SE, Haffner SM, Heise MA, Herman WH, Holman RR, Jones NP, et al. Glycemic Durability of Rosiglitazone, Metformin, or Glyburide Monotherapy. 2006;355(23):2427-43.
64. Matthews DR, Cull CA, Stratton IM, Holman RR, Turner RC. UKPDS 26: sulphonylurea failure in non-insulin-dependent diabetic patients over six years. 1998;15(4):297-303.
65. Cnop M, Vidal J, Hull RL, Utzschneider KM, Carr DB, Schraw T, et al. Progressive Loss of β -Cell Function Leads to Worsening Glucose Tolerance in First-Degree Relatives of Subjects With Type 2 Diabetes. *Diabetes Care*. 2007;30(3):677-82.
66. Fonseca VA. Defining and characterizing the progression of type 2 diabetes. *Diabetes Care*. 2009;32 Suppl 2(Suppl 2):S151-6.
67. Xiong X-f, Yang Y, Wei L, Xiao Y, Li L, Sun L. Identification of two novel subgroups in patients with diabetes mellitus and their association with clinical outcomes: A two-step cluster analysis. 2021;12(8):1346-58.

Data availability

The data supporting the conclusions of this study can be found within manuscript and the supplementary materials. The raw data is available upon request from the corresponding author without restrictions.

Conflicts of interest

Submitting author are responsible for co-authors declaring their interests.

Funding Statement

The research did not receive any funding.

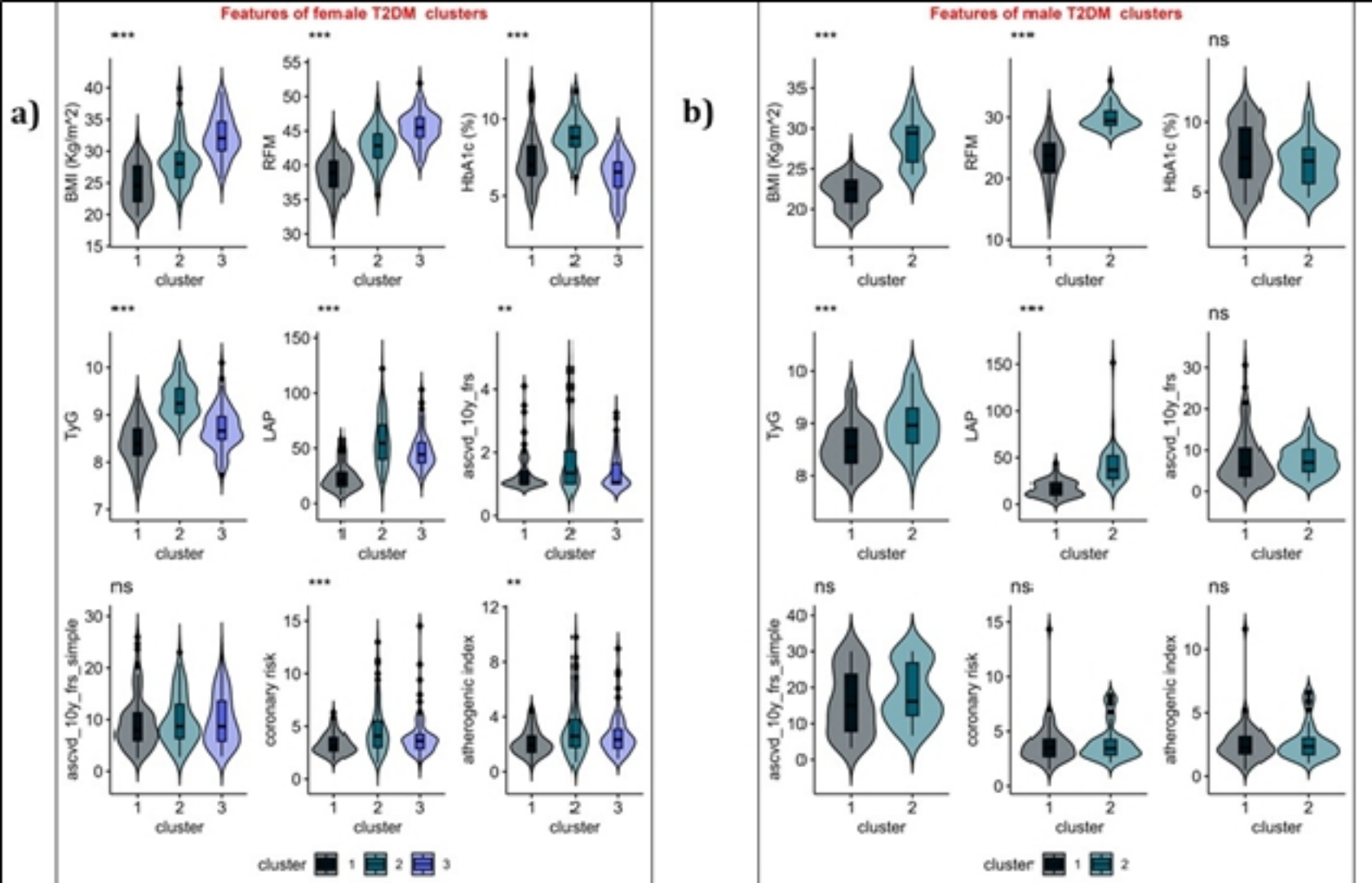
Supplementary data

S1 Fig. Boxplots showing the distribution of the data used for the clustering analysis. Data for female T2DM (left); data for male T2DM (right).

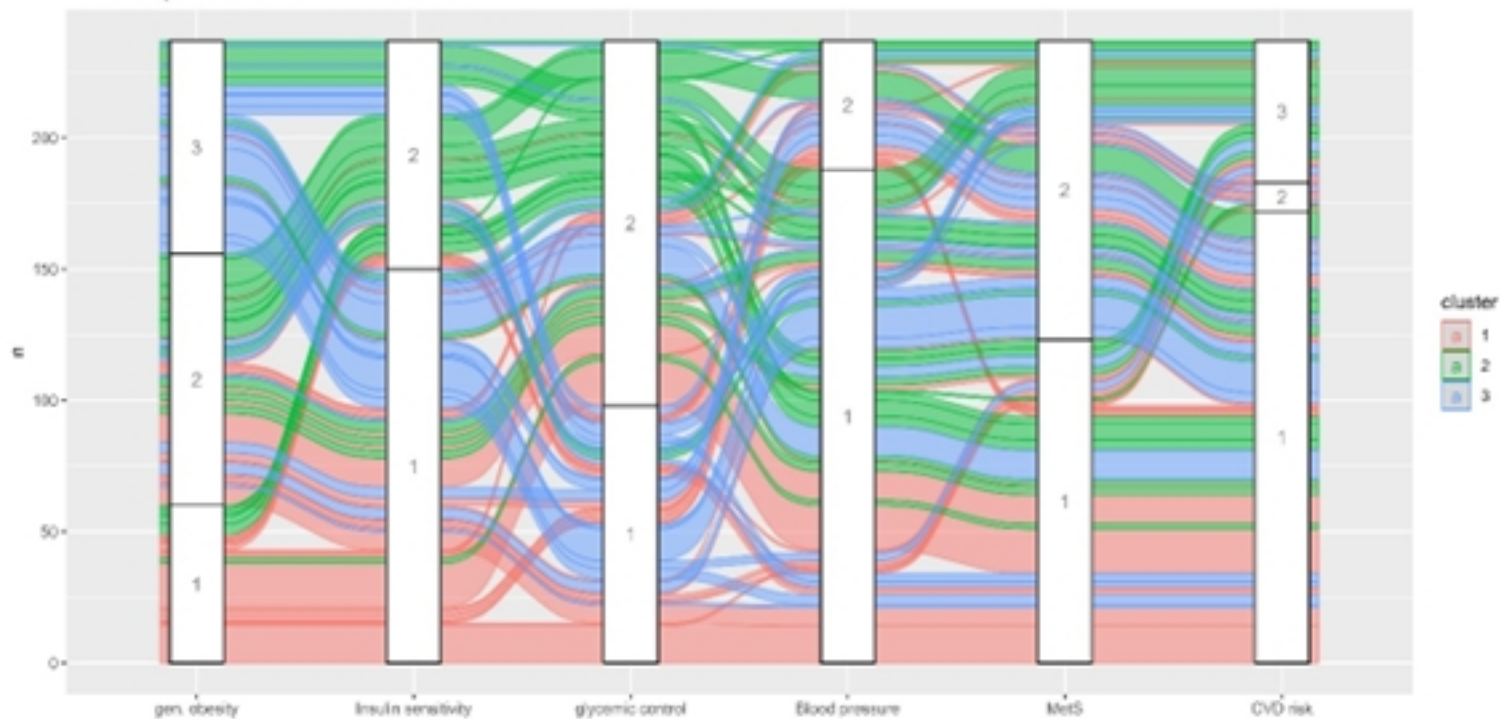
S2 Fig: Bar graph showing the optimal number of clusters that exist in male and female data according to the majority rule.

S3 Fig. A 2D PCA plot showing how the optimal clusters within the male and female T2DM dataset.

S1 Table: Internal validation estimates of optimal clusters.

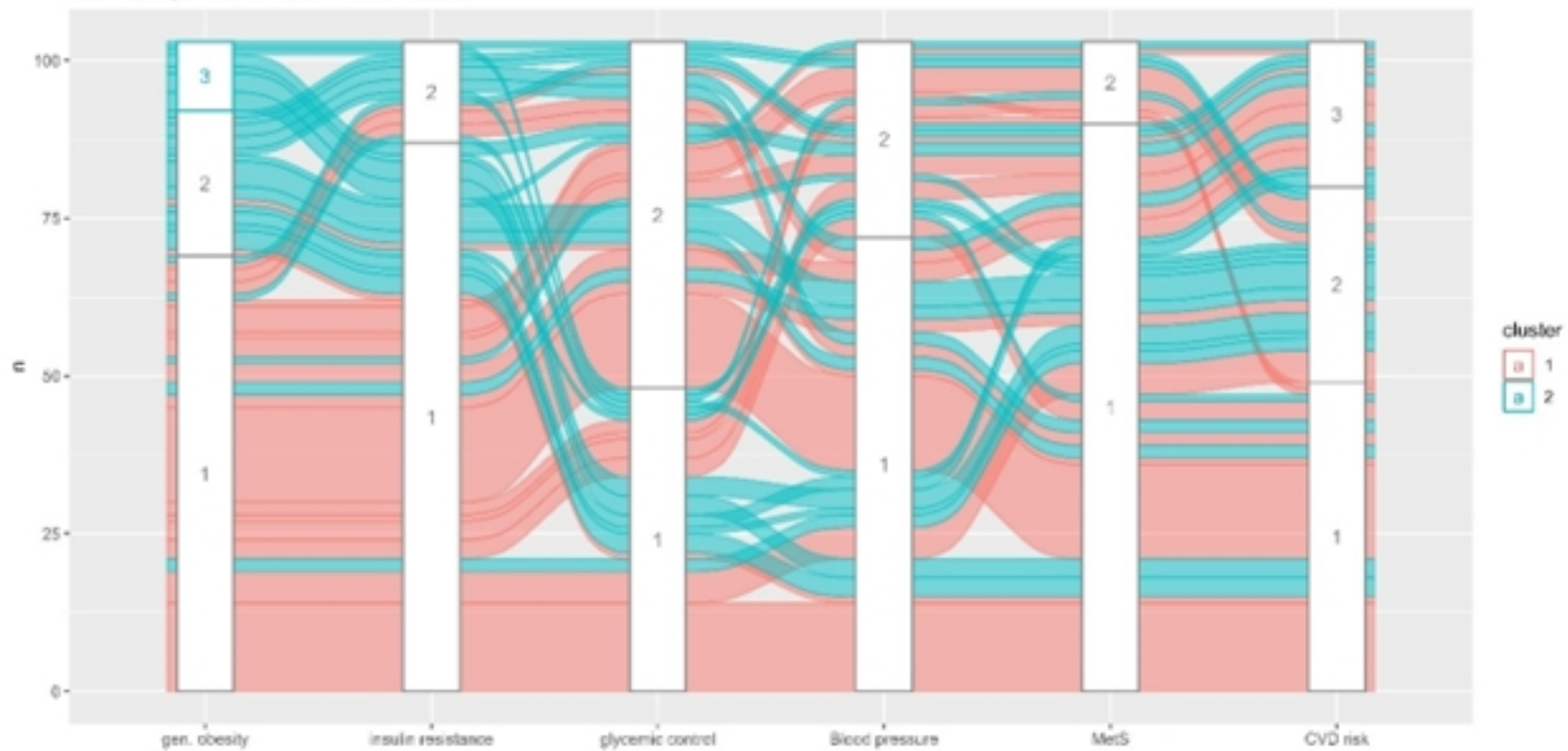


CVD risk profile of female T2DM cluster



Figure

CVD risk profile of male T2DM cluster



Figure