Original Article

Phenotypic Clustering of T2DM in Urban India: Insights for Targeting Modifiable Risk Factors

Debalina S[1], Srinivasulu N[2], Ishita B[2], Paramesh S[3], Sumathi ME[2]

[1]Charaka Super Speciality Hospital, Shri Atal Bihari Vajpayee Medical College and Research Institute, Bengaluru, Karnataka, India. [2]Bowring & Lady Curzon Hospital, Shri Atal Bihari Vajpayee Medical College and Research Institute, Bengaluru, Karnataka, India. [3]Bangalore Diabetes Centre and Diagnostic Lab Bengaluru, Karnataka, India

Abstract

Background: Type 2 Diabetes Mellitus (T2DM) is a metabolic disorder characterized by hyperglycemia resulting from insulin resistance and/or impaired insulin secretion. Its rising prevalence particularly in urban Indian populations presents significant healthcare challenges. The aim is to identify phenotypic clusters of T2DM and evaluate the role of modifiable and non-modifiable risk factors in disease progression with a focus on guiding personalized management strategies. **Material and Methods:** This study was conducted at Bangalore Diabetes Centre with 1153 T2DM patients. Using six clinical variables-age, BMI,HbA1c,HOMA-B,HOMA-IR,ALT, phenotypic clustering was performed with k-means clustering (k=4). Cluster stability was validated using the Jaccard bootstrap method. Lifestyle, environmental and genetic factors were collected using validated questionnaires. Biochemical parameters were analyzed and statistical tests, including One-Way ANOVA, Chi-square tests were performed to evaluate associations between clusters and risk factors. **Results:** Four distinct clusters of T2DM were identified: MOD (25%), characterized by high BMI (30.5 kg/m²) and insulin resistance (HOMA-IR: 5); MARD (42%), with older age at diagnosis and better glycemic control and a lower family history prevalence (7.7%); SIRD (16%), marked by significant insulin resistance, hepatic stress (ALT: 72.2 U/L), with a high prevalence of ischemic heart disease. SIDD (17%), the most severe phenotype, with early onset, poor glycemic control HbA1c (12.1%) and severe beta-cell dysfunction. Significant associations were observed between modifiable factors and clusters: higher smoking prevalence in MOD and SIDD ($p < 0.001$), better physical activity in MARD (16.3%, $p = 0.049$), and MOD showing higher unhealthy diet adherence (65.6%) compared to MARD's healthier diet (44.5%, $p = 0.038$). Socioeconomic factors also influenced clustering, with MARD showing a higher percentage of individuals from upper-middle-class incomes ($p = 0.021$). **Conclusion:** This study highlights the importance of targeting modifiable risk factors and leveraging phenotypic clustering to personalise T2DM management. Insights from clustering can guide clinical decision-making by targeting modifiable risk factors. Tailored interventions based on cluster-specific characteristics can improve outcomes, reduce complications and provide a foundation for future diabetes care strategies.

Keywords: Personalized Medicine, Clustering, T2DM subtypes.

Received: 08 August 2025 **Revised: 30 August 2025** **Accepted: 21 September 2025** **Published: 24 December 2025**

Introduction

Type 2 diabetes mellitus (T2DM) is a multi-dimensional metabolic disorder that arises due to a combination of genetic as well as environmental factors. This condition is marked by the presence of hyperglycemia due to impaired insulin secretion and/or insulin resistance. Worldwide, T2DM affects over 537 million adults, with its prevalence rising drastically in recent years.[1] This poses a substantial challenge to healthcare systems, as T2DM is one of the most common non-communicable diseases, associated with complications such as cardiovascular diseases, nephropathy, retinopathy, and neuropathy.[2]

The concept of "risk factors" in diabetes, particularly T2DM, has been extensively explored through epidemiological studies. A notable early effort in this domain is the Framingham Heart Study, which played a pivotal role in popularising the concept of risk factors in cardiovascular diseases, subsequently influencing diabetes research.[3] This study helped establish the framework for identifying and analyzing various lifestyle and environmental factors that contribute to T2DM. Further systematic reviews and meta-analyses of observational studies have expanded on these findings, identifying numerous non-genetic risk factors for T2DM, including obesity, physical inactivity, unhealthy dietary patterns and smoking. The significant role of environmental changes in the T2DM epidemic is highlighted by the rapid increase in its prevalence, despite the stability in genetic factors.[4] Moreover, studies have highlighted variability in the prevalence of risk factors among different ethnic and racial groups, emphasizing the need for tailored preventive strategies.[5]

Address for correspondence: Dr. Debalina Sen, Charaka Super Speciality Hospital, Shri Atal Bihari Vajpayee Medical College and Research Institute, Bengaluru, Karnataka, India
E-mail: debalina.biochem@gmail.com

DOI:
10.21276/amit.2025.v12.i3.266

How to cite this article: Debalina S, Srinivasulu N, Ishita B, Paramesh S, Sumathi ME. Phenotypic Clustering of T2DM in Urban India: Insights for Targeting Modifiable Risk Factors. Acta Med Int. 2025;12(3):1248-1255.

The most significant increases in T2DM prevalence are observed in societies undergoing major dietary changes, reductions in physical activity and increase in overweight and obesity.[6] Diets associated with diabetes are typically energy-dense, high in saturated fatty acids, and low in non-starch polysaccharides (NSP). Overweight and obesity are universally associated with an increased risk of T2DM, especially when excess adiposity is centrally distributed.[7] Conventional body mass index (BMI) categories may not adequately reflect the risk of developing T2DM across different population groups due to ethnic differences in body composition and the distribution of excess adiposity. Thus, all lifestyle-related and environmental factors contributing to weight gain must be considered as risk factors for T2DM.

Recent advancements in diabetic research have demonstrated the value of classifying T2DM into phenotypic clusters based on metabolic characteristics. For example, the study by Ahlqvist et al. (2018) introduced the concept of phenotypic clustering, categorizing T2DM into distinct subgroups and the importance of personalized management approaches. Our study builds on this by focusing on the urban Bangalore population, integrating modifiable risk factors into phenotypic clustering to address the unique challenges in this demographic.[8] The integration of phenotypic clustering with a focus on modifiable risk factors represents a significant step towards optimizing T2DM management. This approach has the potential to improve glycemic control, reduce complications and enhance the quality of life for patients with T2DM.

Modifiable risk factors like obesity, poor dietary habits, physical inactivity, and smoking are strongly linked to poor glycemic control, which in turn increases the risk of complications. Additionally, socio-environmental factors such as education income, and access to healthcare also play a significant role in optimizing diabetic care.[9] Targeting these factors can greatly improve the effectiveness of management strategies for T2DM.[10] Thus, this study aims to integrate phenotypic clustering of T2DM with modifiable risk factors to develop targeted management strategies that can guide personalized treatments and improve patient outcomes.

Materials and Methods

The cross-sectional study was conducted at Bangalore Diabetes Centre, Bangalore, India, with a total sample size of 1,153 patients. Ethical approval was granted by the Bangalore Diabetic Centre's Medisys Clinisearch Ethical Review Board (Ethics No. MCERB/219/2022) on 21st May 2022. Inclusion criteria included patients aged ≥18 years, with Type 2 Diabetes Mellitus for less than three years, who provided consent. Exclusion criteria were pregnant women, individuals with hyperthyroidism, hypothyroidism, autoimmune disorders, cognitive dysfunctions, or any comorbidity that prevented participation.

Demographic and anthropometry measurements were recorded for every patient during their clinic visits. Height was measured by a portable/wall-mounted stadiometer with a movable head piece (resolution 0.1cm). Weight was measured with the help of Omron digital body weight scale HN–283. BMI was calculated by the Quetelet formula i.e. a person's body weight in kilograms divided by the square of their height in meters. Blood pressure is recorded to the nearest 2 mm Hg from the right arm in a sitting position with a mercury sphygmomanometer (Diamond Deluxe BP Apparatus, Pune, India).

Blood samples were collected for the measurement of various biochemical parameters using A15 Biosystem Autoanalyzer (manufactured by Biosystems, Spain). The HOMA-β and HOMA-IR indices were calculated using FBS and fasting insulin levels with the HOMA calculator V2.2.3.[11] HbA1c was measured by high-performance liquid chromatography using D10 Biorad (Bio-Rad, Hercules, California, USA). De novo k-means clustering was conducted using six phenotypic variables—age, BMI, ALT, HbA1c, HOMA-B and HOMA-IR with k=4 to identify clinically meaningful subgroups within the population. These variables were selected based on their established relevance in capturing the heterogeneity and pathophysiology of diabetes. Age and BMI are critical indicators of demographic and obesity-related risks, while ALT reflects liver function and metabolic health. HOMA-IR and HOMA-B assess insulin resistance and beta-cell function, which are central to diabetes progression, and HbA1c provides a reliable measure of long-term glycemic control. The choice of k=4 for clustering was guided by prior studies, such as those by Ahlqvist et al.2018, which demonstrated the clinical utility of identifying distinct subgroups within diabetes populations. The resulting clusters were designated as MOD, MARD, SIRD, and SIDD. Cluster stability was assessed using the Jaccard bootstrap method, involving 2,000 dataset resampling. A Jaccard similarity index >0.75 indicated stable clusters.[12]

A structured questionnaire was prepared to gather data on demographic, lifestyle, and environmental factors. The questionnaire underwent expert validation, which involved a pilot study with 30 participants to ensure clarity, reliability, and construct validity. Modifications were made based on feedback before implementation. Data were collected through face-to-face interviews and physical assessments conducted by trained healthcare professionals. Definitions and categorisation of variables were clearly established. Diet was categorized into two groups: healthy and unhealthy. A healthy diet was defined as regular consumption of fruits, vegetables, and minimally processed foods, while an unhealthy diet included frequent consumption of processed foods and sugary beverages.[13] Physical activity was categorized as moderate or absent according to WHO recommendations, where moderate physical activity referred to engaging in at least 150 minutes of moderate-intensity aerobic activity per week.[14] Participants were asked about the presence of Type 2 Diabetes in immediate family members. A positive family history was recorded for participants reporting at least one affected family member. Smoking status was determined by identifying participants as either smokers or non-smokers based on their current smoking habits. Access to healthcare facilities was assessed by measuring the distance from participants' residences to the nearest healthcare facility, categorized as either within 5km or within 10km. Education level was determined by the highest level of formal education completed, with levels categorized according to ISCED 2011.[15]

Family income, defined as the total monthly household income, was categorized into income levels based on The World Bank's criteria, adjusted for the Indian context. Households earning less than ₹30,883 per month were classified as lower middle-income, while those earning between ₹30,884 and ₹95,678 per month were classified as upper middle-income.[16] Statistical analysis was performed using SPSS version 26 and R version 4.0.3.[17] Descriptive statistics were computed for all variables. One-Way ANOVA, Levene's test, Welch's Robust Test, and Bonferroni correction were used to analyze differences in biochemical parameters across clusters. Chi-square tests were employed for categorical variables, and graphical representations were created using Microsoft Excel.

Results

We have successfully identified four distinct clusters within the Bangalore population with Type 2 Diabetes Mellitus, each exhibiting unique clinical characteristics and metabolic profiles. Cluster 1, termed Mild Obesity-Related Diabetes (MOD), comprises 25% of the population (n=291) and is characterized by a younger mean age at diagnosis (49.8 years), elevated BMI (30.5 kg/m²), and preserved beta-cell function (HOMA-B: 132.6). This cluster also shows high insulin resistance (HOMA-IR: 5), reflecting obesity-driven metabolic dysfunction. The higher prevalence of smokers (10%) and unhealthy dietary habits (65.6%) in this cluster further exacerbates obesity-related risks, positioning it at high risk for long-term complications, particularly cardiovascular diseases. Cluster 2, known as Mild Age-Related Diabetes (MARD), represents the largest group, accounting for 42% of the population (n=479). It predominantly includes older individuals (mean age 57.5 years) with relatively better glycemic control (HbA1c: 7.9%) and the lowest prevalence of family history (7.7%). This cluster exhibits a reduced metabolic burden, as evidenced by lower BMI (26.8 kg/m²) and ALT levels (29.1 U/L). MARD participants benefit from healthier lifestyle factors, such as moderate physical activity (16.3%) and adherence to healthy dietary habits (44.5%), which appear to mitigate the progression of diabetes. Cluster 3, known as Severe Insulin-Resistant Diabetes (SIRD), accounts for 16% of the population (n=181) and is predominantly male (69.1%). Despite a moderate BMI (27.9 kg/m²), this cluster exhibits significant insulin resistance (HOMA-IR: 2.5) and elevated ALT levels (72.2 U/L), indicating hepatic stress, likely linked to non-alcoholic fatty liver disease (NAFLD). Participants in this cluster have a younger mean age at diagnosis (44.9 ± 10.8 years) and face substantial cardiovascular risks, such as elevated systolic blood pressure (128 mmHg) and ischemic heart disease. The lower engagement in moderate physical activity (12.7%) likely contributes to the worsening of insulin resistance. Cluster 4, identified as Severe Insulin-Deficient Diabetes (SIDD), represents 17% of the population (n=202) and is characterized by the youngest mean age at diagnosis (38.8 years) and the highest HbA1c levels (12.1%), indicating severe hyperglycemia and poor long-term glycemic control. This cluster is driven by profound beta-cell dysfunction, as reflected by reduced HOMA-B values, which severely impair glucose uptake. Unhealthy dietary habits (56.9%) and low levels of physical activity (87.1% with no activity) exacerbate the challenges faced by this group, necessitating aggressive glycemic management to mitigate complications [Table 1].

In parallel, the study also assessed non-modifiable risk factors contributing to these clusters. Non-modifiable factors, such as family history of diabetes, educational attainment, income status and access to healthcare were evaluated to understand their underlying role in disease susceptibility [Table 2]. Income status was significantly associated with clusters (p = 0.021), with a higher proportion of individuals from the upper-middle-class income group in the MARD cluster (69.9%), whereas lower-middle-class representation was greater in the MOD and SIDD clusters. Although education level did not show statistical significance (p = 0.065), secondary or higher education was common across all clusters, with the SIRD cluster having the highest proportion (96.1%). Access to healthcare facilities, while not statistically significant (p = 0.099), showed that nearly all participants had healthcare access within 5 km [Table 3 and Figure 1].

Table 1: Baseline characteristics of the novel clusters of Type 2 Diabetes Mellitus

	Cluster 1 MOD (Mean ± S.D.)	Cluster 2 MARD (Mean ± S.D)	Cluster 3 SIRD (Mean ± S.D)	Cluster 4 SIDD (Mean ± S.D)	Significance
n	291	479	181	202	-
Frequency, %	25	42	16	17	-
Men, %	56.0	53.7	69.1	71.3	p < 0.001
Women, %	44.0	46.3	30.9	28.7	p < 0.001
Diabetes Duration, years	2.7	3.0	2.0	2.2	p < 0.001
DBP, mm Hg	79.5 ± 8.3	78.3 ± 6.2	81.4 ± 4.5	82 ± 4.9	p < 0.001
SBP, mm Hg	121 ± 9.9	118 ± 6.2	128 ± 9.6	133 ± 6.9	p < 0.001
Drinking, n (%)	41 (14.4)	73 (18.3)	56 (24.5)	84 (22.8)	p < 0.021
Smoking, n (%)	29 (10.0%)	19 (4.0%)	7 (3.9%)	21 (10.4%)	p < 0.001
Family History of Diabetes, n (%)	38 (13.1%)	37 (7.7%)	16 (8.8%)	22 (10.9%)	p <.0098
Age at diagnosis, years	49.8 ± 11.5	57.5 ± 12.9	44.9 ± 10.8	38.8 ± 7.6	p < 0.001
HbA1c, %	8.4 ± 2.04	7.9 ± 1.7	9.6 ± 2.2	12.1 ± 1.8	p < 0.001
HOMA-B	132.6 ± 60.7	83.3 ± 36.5	64.7 ± 32.9	43.7 ± 22.9	p < 0.001
HOMA-IR	5 ± 2.10	2.1 ± 0.6	2.5 ± 1.0	2.0 ± 0.8	p < 0.001
BMI, kg/m2	30.5 ± 2.8	26.8 ± 3.8	27.9 ± 3.0	25.5 ± 2.2	p < 0.001
ALT, U/L	33.6 ± 19.0	29.1 ± 14.6	72.2 ± 17.9	30.6 ± 10.4	p < 0.001

Data are presented as mean, or n (%). Chi-square test was used for between group comparisons of categorical variables. BMI, body mass index; HbA1c, Glycated hemoglobin; HOMA-B, homeostasis model assessment of beta-cell function; HOMA-IR, homeostasis model assessment of insulin resistance; ALT, Alanine Aminotransferase. The baseline characteristics of four distinct T2DM clusters: MOD (Mild Obesity-related Diabetes), MARD (Mild Age-Related Diabetes), SIRD (Severe Insulin-Resistant Diabetes), and SIDD (Severe Insulin-Deficient Diabetes). All parameters are statistically significant ($p < 0.05$). Variables in bold are used for clustering.

Table 2: Distribution of risk factors of T2DM across the clusters

Risk Factor		MOD	MARD	SIRD	SIDD
Current smoker, n (%)	Count (n)	29	19	7	21
	% within cluster	10.0%	4.0%	3.9%	10.4%
Non smoker, n (%)	Count (n)	262	460	174	181
	% within cluster	90.0%	96.0%	96.1%	89.6%
Moderate physical activity, n (%)	Count (n)	39	78	23	26
	% within cluster	13.4%	16.3%	12.7%	12.9%
No physical activity, n (%)	Count (n)	252	401	158	176
	% within cluster	86.6%	83.7%	87.3%	87.1%
Family history with type 2 diabetes , n (%)	Count (n)	38	37	16	22
	% within cluster	13.1%	7.7%	8.8%	10.9%
No family history of diabetes, n (%)	Count (n)	253	442	165	180
	% within cluster	86.9%	92.3%	91.2%	89.1%
Healthy diet , n (%)	Count (n)	100	213	70	87
	% within cluster	34.4%	44.5%	38.7%	43.1%
Unhealthy diet , n (%)	Count (n)	191	266	111	115
	% within cluster	65.6%	55.5%	61.3%	56.9%
Upper middle class , n (%)	Count (n)	189	335	108	121
	% within cluster	64.9%	69.9%	59.7%	59.9%
Lower middle class , n (%)	Count (n)	102	144	73	81
	% within cluster	35.1%	30.1%	40.3%	40.1%
Secondary School & above , n (%)	Count (n)	276	436	174	188
	% within cluster	94.8%	91.0%	96.1%	93.1%
Primary School , n (%)	Count (n)	15	43	7	14
	% within cluster	5.2%	9.0%	3.9%	6.9%
Healthcare Access within 5 km , n (%)	Count (n)	288	474	181	202
	% within cluster	99%	99%	100%	100%
Healthcare Access within 10 km , n (%)	Count (n)	10	5	0	0
	% within cluster	1%	1%	0%	0%

Distribution of lifestyle, socioeconomic, and healthcare-related risk factors among the four diabetes clusters: Mild Obesity-related Diabetes (MOD), Mild Age-Related Diabetes (MARD), Severe Insulin-Resistant Diabetes (SIRD), and Severe Insulin-Deficient Diabetes (SIDD). Values are presented as counts with percentages within each cluster.

Table 3: Statistical analysis of risk factors

Current Smokers	**Value**	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square	17.671a	3	<.001
Likelihood Ratio	17.517	3	<.001
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 11.93.			
Moderate Physical Activity	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	2.418a	3	.0490
Likelihood Ratio	2.402	3	.0493
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 26.06.			
Family History	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	6.290a	3	.0098
Likelihood Ratio	6.156	3	.0104
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 17.74.			
Diet	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	8.429a	3	.038
Likelihood Ratio	8.503	3	.037
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 73.78.			
Income Status	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	9.695a	3	.021
Likelihood Ratio	9.687	3	.021
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 62.79.			
Education	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	7.231a	3	.065
Likelihood Ratio	7.506	3	.057
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 12.40.			
Access to Health Care	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	0.005a	3	.099

Likelihood Ratio	0.012	3	.098
a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 99.75			

Statistical analysis conducted using the Pearson Chi-Square test to assess the association between various Lifestyle factors and health outcomes. Key factors analyzed include smoking, physical activity, family history, diet, income status, education, and access to healthcare. The degrees of freedom and asymptotic significance levels (p-values) for each factor were examined to identify statistically significant relationships.

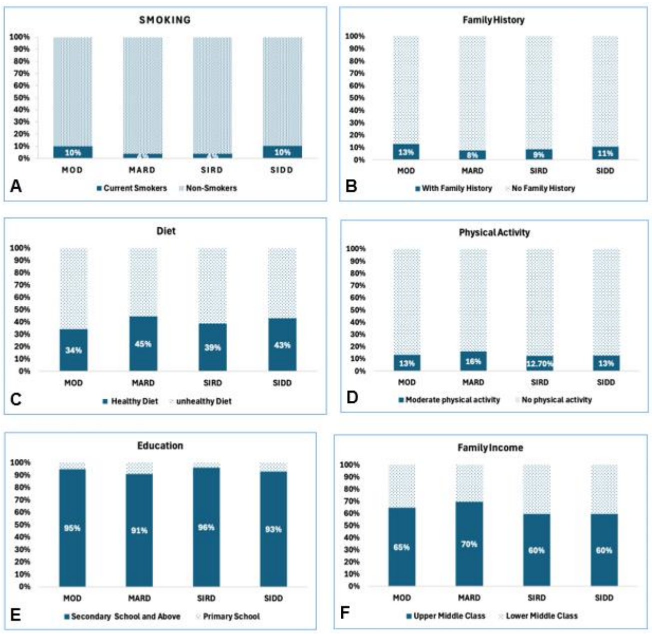

Figure 1: Graphical representation of modifiable and non-modifiable risk factors among the clusters. (A) highlights smoking behavior, with MOD and SIDD showing the highest smoking rates. (B) shows family history of diabetes, with SIDD having the highest percentage. (C) examines dietary habits, with MARD having the highest proportion of healthy diet followers. (D) illustrates physical activity, where MARD leads in moderate physical activity. (E) presents education levels, showing high rates of secondary education across clusters, and (F) compares family income, with MARD having the largest share of upper middle-class individuals

Discussion

Type 2 Diabetes Mellitus (T2DM) is a complex metabolic disorder characterized by chronic hyperglycemia resulting from insulin resistance and/or inadequate insulin secretion. This condition presents a growing global healthcare challenge due to its associated complications, including cardiovascular diseases, nephropathy, and neuropathy. In India, the burden of T2DM is particularly alarming, driven by both genetic predisposition and lifestyle factors.[18] The rising trend of sugar consumption coupled with decreased physical activity holds a greater metabolic significance contributing to high tendency of Indians to develop insulin resistance, abdominal adiposity, ectopic fat deposition, hyperglycemia and atherosclerosis and cardiovascular diseases.[19]

Environmental and lifestyle changes, coupled with the aging of populations, have been generally believed to account for the rapid global increase in the prevalence and incidence of type 2 diabetes in recent decades.[20] A high-calorie diet and sedentary lifestyle have been widely recognized as the primary causes of T2DM.[21] These factors have also been implicated in the global obesity epidemic, which is closely linked to the rising rates of T2DM.[22] Upon closer examination, it has been observed that a high body mass index contributes less to the risk of T2DM compared to increased visceral fat and/or ectopic fat, particularly in the liver.[23] This explains why some obese individuals who do not have metabolic problems typically exhibit low levels of visceral or liver fat.[24] Conversely, individuals in Asia who develop T2DM despite being only overweight or within a normal weight range often show signs of visceral obesity, ectopic fat deposition, and reduced muscle mass, which results in a normal or near-normal BMI.[25]

To address this growing problem, we conducted clustering using k means clustering method (N=1153), six clinical variables - age, BMI, HbA1c, HOMA-B, HOMA-IR, ALT were used and four clusters were identified within the Bangalore population (MOD, MARD, SIRD, SIDD). The identified clusters further provide insights into the interplay between biological, lifestyle, and socioeconomic factors in shaping diabetes phenotypes, emphasizing the heterogeneity of T2DM and the critical need for personalized disease management. These clusters provide valuable insights into the interplay between biological, lifestyle and socioeconomic factors in shaping diabetes phenotypes, highlighting the heterogeneity of T2DM. By examining modifiable risk factors and their association with these clusters, our study highlights the need for personalized disease management strategies to improve outcomes for individuals with T2DM.

In our study, we found that the MARD cluster had the highest percentage (44.5%) of individuals following a healthy diet, whereas the MOD cluster had the greatest percentage (65.6%) of individuals with an unhealthy diet. Diet plays an important role in determining the standard of nutrition received by the body. A high-calorie diet rich in refined carbohydrates and unhealthy fats increases the risk of Type 2 diabetes, while a more balanced diet helps to control blood sugar levels. Foods rich in fibre, proteins and healthy fats improve insulin sensitivity, while consuming food rich in sugar and processed foods leads to glucose spikes. Our study found a significant association between dietary habits and T2DM prevalence, which is supported by Menon et al 2021,[26] who emphasize the impact of urban dietary transitions on diabetes prevalence. It was also demonstrated that Breakfast skipping is associated with a significantly increased risk of T2D. This suggests that regular breakfast consumption is potentially important for the prevention of T2DM, for a good nutritional status along with quality and quantity of food it is also important to maintain regular meal times for better management of diabetes.[27]

Regular physical activity enhances insulin sensitivity by increasing glucose uptake in skeletal muscles. In contrast, inactivity leads to decreased muscle glucose utilization, promoting hyperglycaemia. Sedentary lifestyle also contributes to weight gain, fat accumulation, and systemic inflammation, exacerbating insulin resistance. Moreover, physical inactivity impairs lipid metabolism and mitochondrial function, further disrupting glucose homeostasis.[28] Voluntary weight loss has been shown to improve insulin sensitivity, and several trials have demonstrated a reduced risk of progression from impaired glucose tolerance to T2DM.[29] It was observed in our study that the MARD cluster had the highest engagement in moderate physical activity (16.3%), while SIDD clusters exhibited the lowest participation (12.9%). Obesity and physical inactivity are well-established contributors to the prevalence of Type 2 Diabetes Mellitus (T2DM). Studies have shown that T2DM patients in Bangalore exhibit a high rate of obesity, often accompanied by a lack of physical activity.[30] A study conducted in U.S.A suggested that doing vigorous exercise—working out at 80-90% of maximum heart rate based on age—for at least 20 minutes, five times a week, can greatly improve how the body responds to insulin.[31] It was found that replacing 30 minutes of sitting time with moderate to vigorous physical activity resulted in a 15% improvement in insulin sensitivity.[32] Studies have shown that increased physical activity improves insulin sensitivity and lowers the risk of Type 2 Diabetes.[33]

Alcohol consumption plays a significant role in the incidence and progression of Type 2 Diabetes Mellitus (T2DM), as evidenced by the varying drinking patterns across the identified clusters. In the Bangalore population, the prevalence of drinking was highest in the SIRD (24.5%) and SIDD (22.8%) clusters, followed by MARD (18.3%) and MOD (14.4%). Chronic alcohol consumption impairs insulin signaling by inducing hepatic steatosis, increasing systemic inflammation, and promoting oxidative stress, all of which contribute to insulin resistance and hyperglycemia. Prolonged alcohol use can impair pancreatic beta-cell function, reducing insulin secretion and exacerbating metabolic dysfunction.[34] Interestingly, moderate alcohol intake has been suggested to enhance insulin sensitivity, but its effects are highly dependent on dose and individual factors.[35]

Smoking is a significant and modifiable risk factor for T2DM known to increase diabetes risk by promoting insulin resistance and inflammation, which exacerbate hyperglycemia and metabolic dysfunction.[36] Our analysis revealed strong statistical correlations between smoking and the identified T2DM clusters. The prevalence of smoking was highest in the MOD (10%) and SIDD (10.4%) clusters, while the MARD (4%) and SIRD (3.9%) clusters exhibited comparatively lower rates. Smoking contributes to T2DM progression by impairing insulin sensitivity and promoting inflammatory pathways. Furthermore, research indicates that both active and passive smoking significantly elevate the risk of developing T2DM.[37]

Family history has been identified as a critical non-modifiable risk factor for Type 2 Diabetes Mellitus (T2DM), with the MOD cluster in this study exhibiting the highest prevalence (13.1%). This observation aligns with the ICMR-INDIAB study, which emphasizes the significant role of genetic predisposition in the development of T2DM. The risk of developing T2DM increases substantially if a first-degree relative (parent, sibling, or child) has the disease. Research indicates that individuals with a family history of diabetes are at a significantly higher risk of developing the condition, primarily due to a combination of genetic predisposition and shared environmental factors.[38]

Socioeconomic status, measured by family income, was also found to influence T2DM risk. The MARD cluster had the highest percentage (69.9%) of individuals in the upper-middle class, suggesting a higher income level within this group. Lower socioeconomic status had been associated with higher diabetes prevalence due to limited access to healthcare and healthy lifestyle options.[39] The ICMR-INDIAB study offers valuable context, demonstrating that urbanization and higher socioeconomic status are linked to a greater prevalence of T2DM.

The study's findings highlight the complex interplay between diet, physical activity, smoking, family history, and socioeconomic status in the risk of T2DM. It is important to recognize that environmental and lifestyle factors generally have a limited direct impact on β-cell function. However, exceptions do exist, such as elevated levels of nutrients or their metabolites in the blood, which can induce metabolic stress and indirectly influence β-cells by affecting other organs and systems. These include the immune system, blood vessels, adipose tissue, liver, muscle, brain, intestines, and gut microbiota.[40]

This study's strength lies in its comprehensive approach, integrating clinical, lifestyle, and socioeconomic factors to identify distinct T2DM phenotypic clusters within an urban Indian population. The large sample size (n=1153), use of robust statistical methods like k-means clustering, and the inclusion of both modifiable (diet, physical activity, smoking) and non-modifiable (family history, socioeconomic status) risk factors enhance its reliability and relevance for personalized disease management. However, the cross-sectional design limits the ability to establish causal relationships or track transitions between clusters over time. Additionally, reliance on self-reported lifestyle data introduces potential recall bias. The exclusion of genetic, epigenetic and psychosocial factors presents an important limitation.

Future studies should focus on longitudinal analyses to track how individuals move between clusters over time and their associated outcomes. Incorporating genetic and emerging biomarkers like gut microbiota could deepen the understanding of T2DM heterogeneity and its underlying mechanisms. Expanding research to rural and diverse populations would provide insights into regional and cultural differences in diabetes risk.

Conclusion

In conclusion, this study emphasizes the importance of targeting modifiable risk factors, such as obesity, poor dietary habits and physical inactivity, alongside leveraging phenotypic clustering to personalise Type 2 Diabetes Mellitus (T2DM) management. By identifying four distinct T2DM clusters within an urban Indian population, this research highlights the heterogeneity of the disease and the critical need for personalized care strategies. Addressing both lifestyle and socioeconomic factors, while considering genetic predispositions, offers a holistic approach to mitigating the burden of T2DM. These insights provide a foundation for future research and interventions aimed at mitigating the growing impact of diabetes globally.

Acknowledgements: We thank the Bangalore Diabetes Centre and Diagnostic Lab Bengaluru, Karnataka, India and Shri Atal Bihari Vajpayee Medical College and Research Institute, Bengaluru, Karnataka for supporting the study.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. “IDF Diabetes Atlas.” Accessed: Dec. 03, 2024. [Online]. Available: https://diabetesatlas.org/.
2. D. Prabhakaran et al., “The changing patterns of cardiovascular diseases and their risk factors in the states of India: the Global Burden of Disease Study 1990–2016,” The Lancet Global Health, vol. 6, no. 12, pp. e1339–e1351, Dec. 2018, doi: 10.1016/s2214-109x(18)30407-8.
3. W. B. Kannel, “Diabetes and cardiovascular disease. The Framingham study,” JAMA: The Journal of the American Medical Association, vol. 241, no. 19, pp. 2035–2038, May 1979, doi: 10.1001/jama.241.19.2035.
4. V. Bellou, L. Belbasis, I. Tzoulaki, and E. Evangelou, “Risk factors for type 2 diabetes mellitus: An exposure- wide umbrella review of meta-analyses,” PLOS ONE, vol. 13, no. 3, p. e0194127, Mar. 2018, doi: 10.1371/ journal.pone.0194127.
5. “2. Classification and Diagnosis of Diabetes:Standards of Medical Care in Diabetes—2022,” Diabetes Care, vol. 45, no. Supplement_1, pp. S17–S38, Dec. 2021, doi: 10.2337/dc22-s002.
6. L. J. Aronne and K. R. Segal, “Adiposity and Fat Distribution Outcome Measures: Assessment and Clinical Implications,” Obesity Research, vol. 10, no. S11, Nov. 2002, doi: 10.1038/oby.2002.184.
7. D. Parker, S. Weiss, R. Troisi, P. Cassano, P. Vokonas, and L. Landsberg, “Relationship of dietary saturated fatty acids and body habitus to serum insulin concentrations: the Normative Aging Study,” The American Journal of Clinical Nutrition, vol. 58, no. 2, pp. 129–136, Aug. 1993, doi: 10.1093/ajcn/58.2.129.
8. E. Ahlqvist et al., “Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables,” The Lancet Diabetes & Endocrinology, vol. 6, no. 5, pp. 361–369, May 2018, doi: 10.1016/s2213-8587(18)30051-2.
9. R. J. Walker, B. L. Smalls, J. A. Campbell, J. L. Strom Williams, and L. E. Egede, “Impact of social determinants of health on outcomes for type 2 diabetes: a systematic review,” Endocrine, vol. 47, no. 1, pp. 29–48, Feb. 2014, doi: 10.1007/s12020-014-0195-0.
10. “5. Facilitating Behavior Change and Well-being to Improve Health Outcomes:Standards of Medical Care in Diabetes—2021,” Diabetes Care, vol. 44, no. Supplement_1, pp. S53–S72, Dec. 2020, doi: 10.2337/dc21-s005.
11. J. C. Levy, D. R. Matthews, and M. P. Hermans, “Correct Homeostasis Model Assessment (HOMA) Evaluation Uses the Computer Program,” Diabetes Care, vol. 21, no. 12, pp. 2191–2192, Dec. 1998, doi: 10.2337/ diacare.21.12.2191.
12. C. Hennig, “Dissolution point and isolation robustness: Robustness criteria for general cluster analysis methods,” Journal of Multivariate Analysis, vol. 99, no. 6, pp. 1154–1176, Jul. 2008, doi: 10.1016/j.jmva.2007.07.002.
13. C. A. Monteiro, G. Cannon, J.-C. Moubarac, R. B. Levy, M. L. C. Louzada, and P. C. Jaime, “The UN Decade of Nutrition, the NOVA food classification and the trouble with ultra-processing,” Public Health Nutrition, vol. 21, no. 1, pp. 5–17, Mar. 2017, doi: 10.1017/s1368980017000234.
14. H. P. van der Ploeg and F. C. Bull, “Invest in physical activity to protect and promote health: the 2020 WHO guidelines on physical activity and sedentary behaviour,” International Journal of Behavioral Nutrition and Physical Activity, vol. 17, no. 1, Nov. 2020, doi: 10.1186/s12966-020-01051-1.
15. International Standard Classification of Education (ISCED) 2011. UNESCO Institute for Statistics, 2012. Accessed: Dec. 04, 2024. [Online]. Available: https://doi.org/10.15220/978-92-9189-123-8-en
16. N. Fantom and U. Serajuddin, The World Bank’s Classification of Countries by Income. World Bank, Washington, DC, 2016. Accessed: Dec. 04, 2024. [Online]. Available: https://doi.org/10.1596/1813-9450-7528
17. “Posit,” Posit. [Online]. Available: http://www.posit.co/
18. E. L. O’Keefe, J. J. DiNicolantonio, H. Patil, J. H. Helzberg, and C. J. Lavie, “Lifestyle Choices Fuel Epidemics of Diabetes and Cardiovascular Disease Among Asian Indians,” Progress in Cardiovascular Diseases, vol. 58, no. 5, pp. 505–513, Mar. 2016, doi: 10.1016/j.pcad.2015.08.010.
19. S. Gulati and A. Misra, “Sugar Intake, Obesity, and Diabetes in India,” Nutrients, vol. 6, no. 12, pp. 5955– 5974, Dec. 2014, doi: 10.3390/nu6125955.
20. H. Kolb and S. Martin, “Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes,” BMC Medicine, vol. 15, no. 1, Jul. 2017, doi: 10.1186/s12916-017-0901-x.
21. S. Chatterjee, K. Khunti, and M. J. Davies, “Type 2 diabetes,” The Lancet, vol. 389, no. 10085, pp. 2239– 2251, Jun. 2017, doi: 10.1016/s0140-6736(17)30058-2.
22. G. H. Rao, “Global Epidemic of Obesity and Diabetes: World Diabetes Day-2018,” Diabetes & Obesity International Journal, vol. 3, no. 4, pp. 1–3, 2018, doi: 10.23880/doij-16000189.
23. E. Kotsiliti, “NAFLD prevalence in older patients with T2DM,” Nature Reviews Gastroenterology & Hepatology, vol. 20, no. 2, pp. 65–65, Dec. 2022, doi: 10.1038/s41575-022-00733-4.
24. M. Würfel et al., “Adipokines as Clinically Relevant Therapeutic Targets in Obesity,” Biomedicines, vol. 11, no. 5, p. 1427, May 2023, doi: 10.3390/biomedicines11051427.
25. R. Unnikrishnan, R. M. Anjana, and V. Mohan, “Diabetes mellitus and its complications in India,” Nature Reviews Endocrinology, vol. 12, no. 6, pp. 357–370, Apr. 2016, doi: 10.1038/nrendo.2016.53.
26. Menon, “Study on the prevalence of type 2 diabetes mellitus among women population of Bangalore, India,” Journal of International Women’s Studies, vol. 22, no. 6, pp. 43–54, 2021.
27. H. Bi, Y. Gan, C. Yang, Y. Chen, X. Tong, and Z. Lu, “Breakfast skipping and the risk of type 2 diabetes: a meta-analysis of observational studies,” Public Health Nutrition, vol. 18, no. 16, pp. 3013–3019, Feb. 2015, doi: 10.1017/s1368980015000257.
28. L. Qin, M. J. Knol, E. Corpeleijn, and R. P. Stolk, “Does physical activity modify the risk of obesity for type 2 diabetes: a review of epidemiological data,” European Journal of Epidemiology, vol. 25, no. 1, pp. 5–12, Oct. 2009, doi: 10.1007/s10654-009-9395-y.
29. J. Lindström et al., “Prevention of Diabetes Mellitus in Subjects with Impaired Glucose Tolerance in the Finnish Diabetes Prevention Study,” Journal of the American Society of Nephrology, vol. 14, no. suppl_2, pp. S108– S113, Jul. 2003, doi: 10.1097/01.asn.0000070157.96264.13.
30. J. Aravinda, “Risk factors in patients with type 2 diabetes in Bengaluru: A retrospective study,” World Journal of Diabetes, vol.

10, no. 4, pp. 241–248, Apr. 2019, doi: 10.4239/wjd.v10.i4.241.

31. E. J. Mayer-Davis, “Intensity and Amount of Physical Activity in Relation to Insulin Sensitivity,” JAMA, vol. 279, no. 9, p. 669, Mar. 1998, doi: 10.1001/jama.279.9.669.
32. T. Yates et al., “Objectively measured sedentary time and associations with insulin sensitivity: Importance of reallocating sedentary time to physical activity,” Preventive Medicine, vol. 76, pp. 79–83, Jul. 2015, doi: 10.1016/ j.ypmed.2015.04.005.
33. C. Y. Jeon, R. P. Lokken, F. B. Hu, and R. M. van Dam, “Physical Activity of Moderate Intensity and Risk of Type 2 Diabetes,” Diabetes Care, vol. 30, no. 3, pp. 744–752, Mar. 2007, doi: 10.2337/dc06-1842.
34. A. Ghorpade et al., “Diabetes in rural Pondicherry, India: a population-based studyof the incidence and risk factors,” WHO South-East Asia Journal of Public Health, vol. 2, no. 3, p. 149, 2013, doi: 10.4103/2224-3151.206761.
35. I. C. Schrieks, A. L. J. Heil, H. F. J. Hendriks, K. J. Mukamal, and J. W. J. Beulens, “The Effect of Alcohol Consumption on Insulin Sensitivity and Glycemic Status: A Systematic Review and Meta-analysis of Intervention Studies,” Diabetes Care, vol. 38, no. 4, pp. 723–732, Mar. 2015, doi: 10.2337/dc14-1556.
36. D. J. Debnath, J. Ray, S. M. Jah, and Y. Marimuthu, “Smoking and the Risk of Type 2 Diabetes: A Cross- sectional Analytical Study,” Indian Journal of Community Medicine, vol. 49, no. 4, pp. 588–592, Jul. 2024, doi: 10.4103/ijcm.ijcm_1009_22.
37. A. Pan, Y. Wang, M. Talaei, F. B. Hu, and T. Wu, “Relation of active, passive, and quitting smoking with incident type 2 diabetes: a systematic review and meta-analysis,” The Lancet Diabetes & Endocrinology, vol. 3, no. 12, pp. 958–967, Dec. 2015, doi: 10.1016/s2213-8587(15)00316-2.
38. R. M. Anjana et al., “Prevalence of diabetes and prediabetes (impaired fasting glucose and/or impaired glucose tolerance) in urban and rural India: Phase I results of the Indian Council of Medical Research–INdia DIABetes (ICMR– INDIAB) study,” Diabetologia, vol. 54, no. 12, pp. 3022–3027, Sep. 2011, doi: 10.1007/s00125-011-2291-5.
39. E. Agardh, P. Allebeck, J. Hallqvist, T. Moradi, and A. Sidorchuk, “Type 2 diabetes incidence and socio- economic position: a systematic review and meta-analysis,” International Journal of Epidemiology, vol. 40, no. 3, pp. 804–818, Feb. 2011, doi: 10.1093/ije/dyr029.
40. G. E. Rizzo et al., “Diabetes Mellitus and Microbiota: Knowledge and Perspectives,” in Healthy Ageing and Longevity, Cham: Springer International Publishing, 2023, pp. 131–151. Accessed: Dec. 05, 2024. [Online]. Available: https://doi.org/10.1007/978-3-031-14023-5_7.