



OPEN ACCESS

EDITED BY

Kulvinder Kochar Kaur,
Kulvinder Kaur Centre For Human
Reproduction, India

REVIEWED BY

Zhen Qi,
Shanghai Jiao Tong University, China
Koushik Guchhait,
Medical College of Wisconsin, United States

*CORRESPONDENCE

Si-Jin Li

✉ lisjnm123@163.com

Zhi-Fang Wu

✉ wuzhifang01@163.com

RECEIVED 18 July 2025

ACCEPTED 25 September 2025

PUBLISHED 13 October 2025

CITATION

Liu B, Hou Y-j, Wu P, Han X, Qi H,
Yang X-Y, Wu Z-F and Li S-J (2025)
Exploring unsupervised learning
techniques for early detection of
myocardial ischemia in type 2 diabetes.
Front. Endocrinol. 16:1668516.
doi: 10.3389/fendo.2025.1668516

COPYRIGHT

© 2025 Liu, Hou, Wu, Han, Qi, Yang, Wu and
Li. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums
is permitted, provided the original author(s)
and the copyright owner(s) are credited and
that the original publication in this journal is
cited, in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Exploring unsupervised learning techniques for early detection of myocardial ischemia in type 2 diabetes

Bing Liu^{1,2}, Yan-jie Hou^{1,2}, Ping Wu^{1,2}, Xiao Han^{1,2}, Hao Qi³,
Xiu-Yun Yang⁴, Zhi-Fang Wu^{1,2*} and Si-Jin Li^{1,2*}

¹Department of Nuclear Medicine, First Hospital of Shanxi Medicinal University, Shanxi Medical University, Taiyuan, Shanxi, China, ²Collaborative Innovation Center for Molecular Imaging of Precision Medicine, Shanxi Medical University, Taiyuan, Shanxi, China, ³Department of Endocrinology, First Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, ⁴Modern Educational Technology Center, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China

Introduction: Myocardial ischemia can result in severe cardiovascular complications. However, the impact of clinical factors on myocardial ischemia in individuals with T2DM remains unclear. We applied a clustering approach to identify the variability in myocardial ischemia evaluated through Single-Photon Emission Computed Tomography.

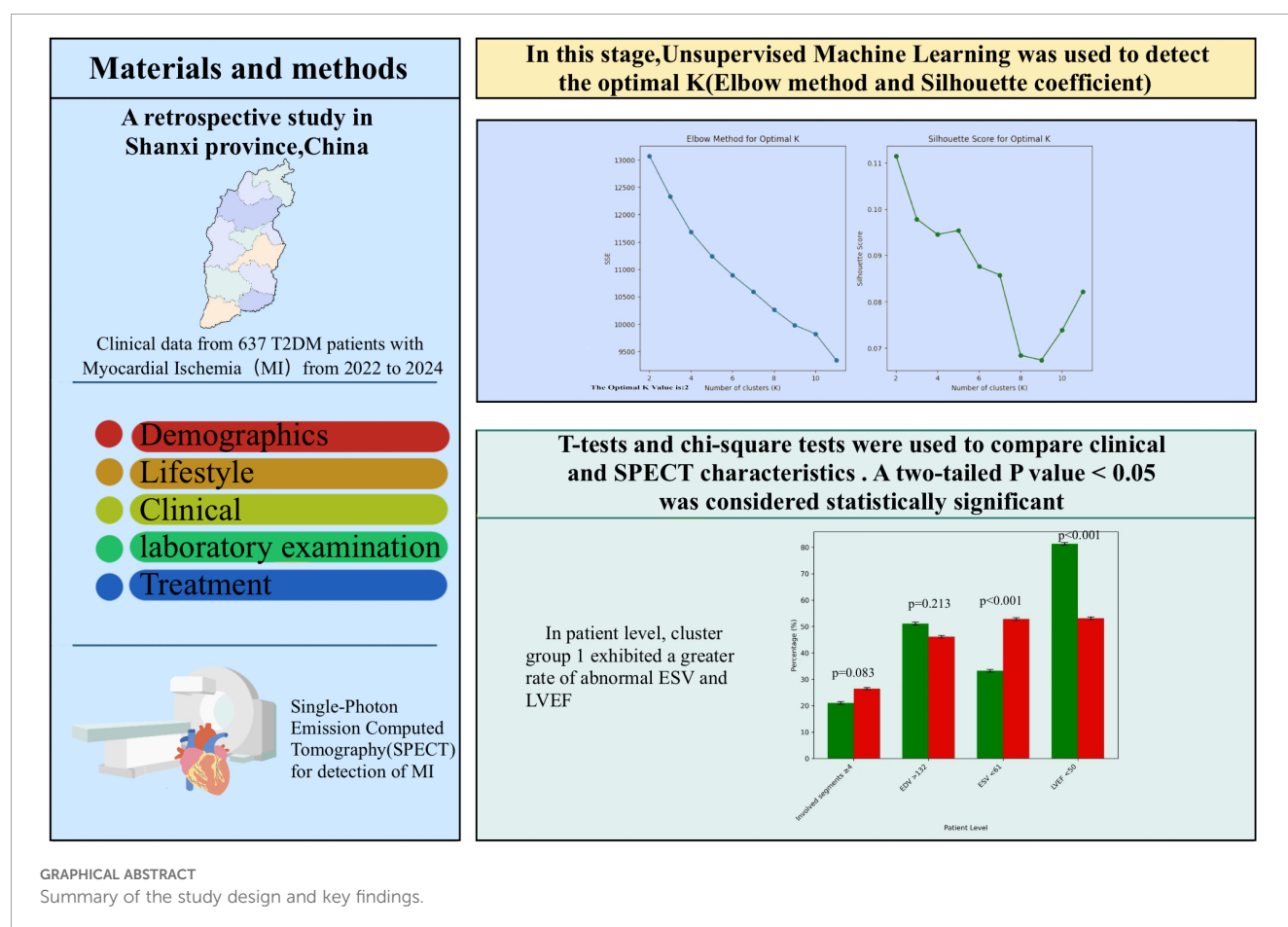
Methods: Retrospective statistics derived from 637 T2DM patients with myocardial ischemia who participated in SPECT imaging at our hospital between January 2022 and September 2024 were gathered. Ischemia areas, cavity size, wall motion, ventricular contraction, cardiac systolic coordination, End-diastolic Volume, End-systolic Volume; Left ventricular injection fraction were assessed and analyzed. Clustering analysis of medical data in unsupervised learning, involving the elbow method and silhouette coefficient (cluster 1: 262; cluster 2: 375).

Results: The Healthcare information between two groups differed in multiple respects (1) Cluster 1 had the older patient (63.23 ± 12.31), longer average duration of diabetes (10.27 ± 8.77), higher Glycated Hemoglobin (HbA1c) values (7.69 ± 1.76), the higher level of serum creatinine ($115.42 \pm 106.18 \mu\text{mol/L}$); and a higher proportion of patients with insulin treatment (40.5%) (2). Cluster 1 had more males (68.8%), higher proportion of patients with smoking history (44.5%), the higher level of Cholesterol ($3.96 \pm 1.12 \text{mmol/L}$), serum uric acid ($406.78 \pm 135.24 \mu\text{mol/L}$), Low-density lipoprotein cholesterol ($2.08 \pm 0.32 \text{mmol/L}$), and was more prone to statin therapy (6.1%). The SPECT features differed across the various clusters (1): Cluster 1 had higher proportion of Hypokinesis (38.2%), poor ventricular contraction (57.6%), Impaired Cardiac systolic coordination (63.7%), and abnormal LVEF (81.3%) (2). Cluster 2 had a higher proportion of total ischemia (11.5%) and abnormal ESV (52.8%) (3). There was no significant difference in Ischemia areas, Cavity size, Involved segments, and EDV.

Discussion: Although the unsupervised clustering approach revealed differences in various clinical and imaging characteristics, no significant differences were observed in ischemic burden, cavity size, involved segments, or EDV.

KEYWORDS

machine learning, elbow method, silhouette coefficient, myocardial ischemia, diabetes mellitus, single-photon emission computed tomography



1 Introduction

Abbreviations: T2DM, Type 2 diabetes mellitus; SPECT, Single-Photon Emission computed tomography; EDV, End-diastolic Volume; ESV, End-systolic Volume; LVEF, Left Ventricular Ejection Fraction; HbA1c, Glycated Hemoglobin; LDL-C, low-density lipoprotein cholesterol; HDL-C, High-density lipoprotein cholesterol; GLP-1 inhibitor, Glucagon-Like Peptide-1 receptor agonist; DPP-4 inhibitor, Dipeptidyl Peptidase-4 inhibitor; SGLT-2 inhibitor, Sodium-Glucose Co-Transporter 2 inhibitor; DM, diabetes mellitus; LV, left ventricular. EDV, End-diastolic Volume; ESV, End-systolic Volume; LVEF, Left ventricular injection fraction.

Diabetes is a major global health burden, with its increasing prevalence contributing to higher morbidity and mortality rates (1, 2). Data from the International Diabetes Federation (IDF) indicates that the global population of individuals with diabetes has reached 536 million, and it is expected to rise to 783 million by 2045 (3). The most widespread form of diabetes is T2DM. Studies have indicated that T2DM contributes to a higher risk of cardiovascular diseases (4).

Cardiovascular disease (CVD) is the leading cause of death and disability in T2DM (5, 6). Research has demonstrated that compared to individuals without diabetes, those with T2DM have a two to four

times higher risk of cardiovascular disease (CVD) (7, 8). Myocardial ischemia (MI) is a major cardiovascular disease, which is more commonly seen in the diabetic population than in those without diabetes (9, 10). Additional risk factors for cardiovascular death, including hypertension, dyslipidemia, smoking, and visceral obesity are particularly harmful in individuals with diabetes (3, 11). T2DM patients have different clinical features, and the comprehensive effect of these features on myocardial ischemia still needs further study.

Machine learning (ML), a branch of artificial intelligence that enables mining the relationships from complex datasets, has been used to make predictions about future outcomes (12). ML-based techniques have been successfully applied on various types of Coronary artery disease (CAD) datasets (13–19). These algorithms have demonstrated promising performance in the detection and treatment of myocardial ischemia. However, Limited research has been dedicated to exploring the use of non-supervised learning algorithms to differentiate the varied medical features of T2DM and examine the connections across various categories and myocardial ischemia characteristics. Thus, we applied non-supervised learning algorithms to examine The diversity in clinical manifestations of T2DM based on clinical indicators and to better understand the overall impact of medical determinants on myocardial ischemia characteristics detected on SPECT, which could potentially enhance personalized medical intervention.

2 Methods

The Biomedical Research Ethics Committee of our hospital approved this retrospective study, and the requirement for written informed consent was waived.

2.1 Study cohort

Between January 2022 to September 2024, We retrospectively reviewed T2DM patients with myocardial ischemia identified through

SPECT at our hospital. The criteria for exclusion were listed below: Patients previously treated with percutaneous coronary intervention, coronary bypass surgery, and cerebrovascular diseases before SPECT, SPECT image quality was inadequate for ischemia diagnosis; deficient clinical data; Patients with heart disease, respiratory failure, severe liver and kidney diseases, cancer, severe infections, and other illnesses. Finally, 637 patients with T2DM were incorporated into the study.

2.2 Acquisition of resting gated MPI

All patients underwent gated resting SPECT MPI using a single IQ-SPECT dual-probe scanner (Symbia T16, Siemens, Germany). After fasting for at least 4 hours, 99mTc-MIBI (740–925 MBq) was intravenously injected. Fatty meals were consumed 15 to 20 minutes prior to imaging. Electrodes were placed on the chest to capture gated data. Patients remained supine with arms raised, and acquisition began 60 minutes post-injection, lasting 8 minutes. Imaging parameters included a 128×128 matrix, 208° rotation, 34 frames, and 25 seconds per frame. Images were reconstructed using ordered subsets expectation maximization in fully automated mode.

2.3 SPECT analysis

The SPECT and gated SPECT images were anonymized and visually analyzed by two experienced observers. Myocardial perfusion was assessed by consensus using a 17-segment division of the left ventricle (20) and a four-point grading system: 0, normal uptake; 1, equivocal; 2, moderate reduction; and 3, severe reduction. Segments with an uptake score ≥ 2 at stress were classified as having a definite uptake reduction (21–23). Only segments with an uptake score ≥ 2 at stress were considered to have a definite uptake reduction at stress. Segments with an uptake score ≥ 2 at stress were classified as having a definite uptake reduction. Total myocardial counts were determined in each acquisition with a manually adjusted elliptical region of interest (Figure 1).

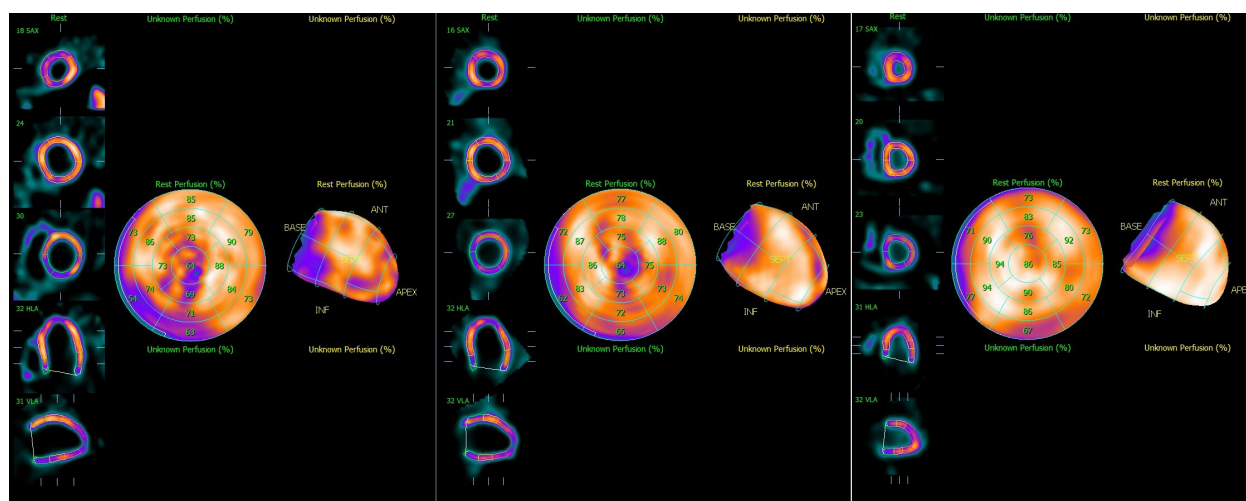


FIGURE 1
Representative Single-Photon Emission computed tomography images of (A) Total ischemia, (B) Part ischemia, and (C) Normal.

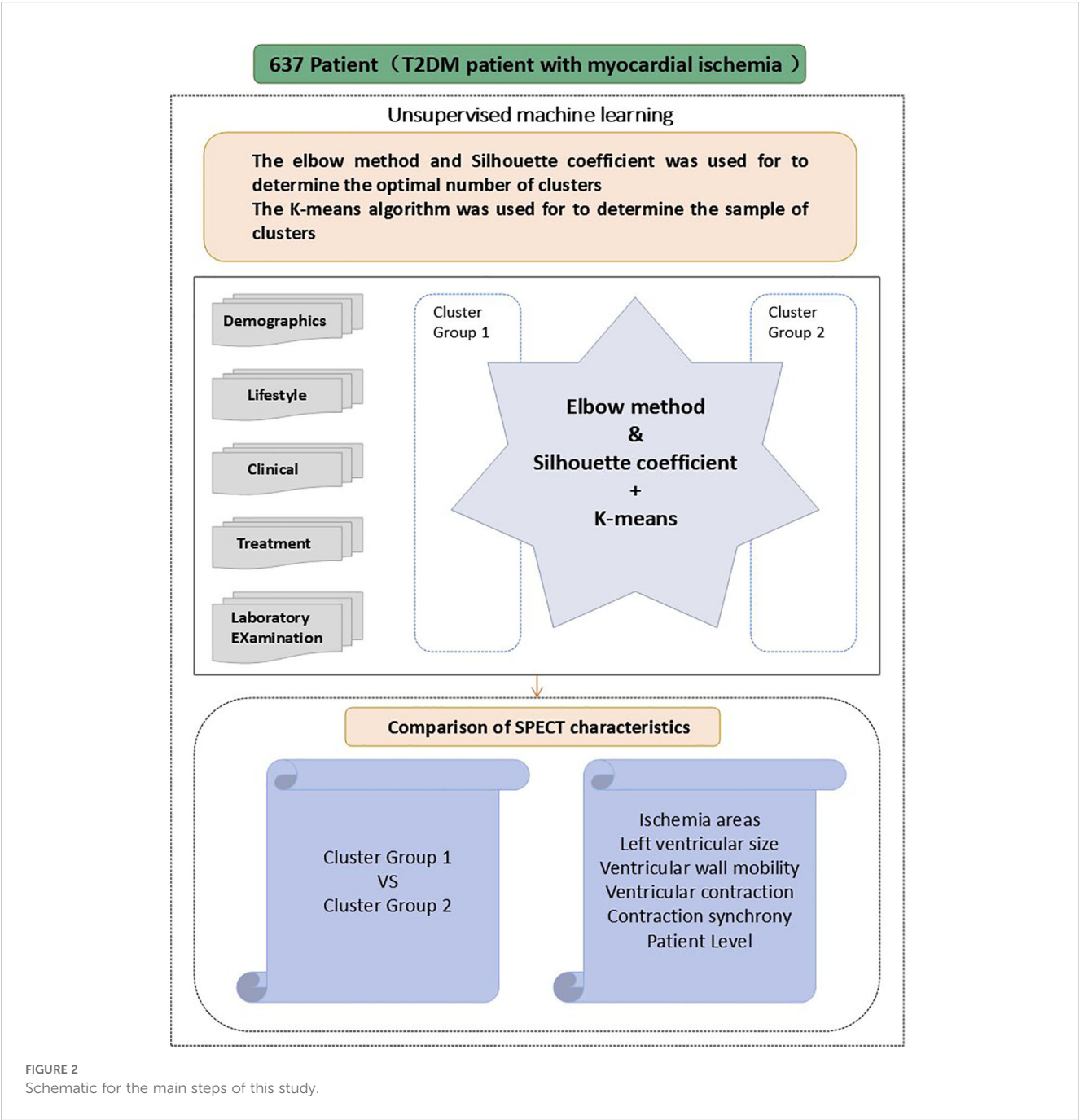
2.4 Unsupervised machine learning

The elbow method combined with the Silhouette coefficient were utilized to cluster the 637 T2DM patients based on their clinical characteristics (Figure 2), to identify the appropriate cluster count. The core principle of this method is to reduce to the smallest possible value the sum of squared errors between the cluster center and the points within each cluster. As K grows, the model's

separation becomes more distinct. However, when K exceeds the best appraisal, The cumulative squared errors no longer shows significant changes. K-means is a clustering technique used for grouping objects that consist of both quantitative and qualitative data (24).

The fundamental procedures of the Silhouette coefficient and the elbow method are described below:

- (1) Select a Range of K values.



Determine a possible range for K, typically starting from 2 (since the Silhouette coefficient is meaningless for K = 1) up to a reasonable maximum value.

(2) Perform the K-means clustering.

For each K value, run the K-means algorithm and record the following results:

SSE (Sum of squared Errors): Used for the Elbow Method.

Silhouette Coefficient: Used to evaluate clustering quality.

(3) Analyze the Elbow Method: Calculate SSE, Plot the Elbow Point, Identify the Elbow point.

(4) Analyze the Silhouette Coefficient: Calculate the Silhouette Coefficient, Plot the Silhouette curve, Select The Optimal K Value.

(5) Compare the Elbow Method and Silhouette Coefficient.

To determine the optimal number of clusters, we evaluated both the elbow method and the silhouette coefficient across a range of K values (K = 2–6). The elbow plot showed that the reduction in the sum of squared errors plateaued after K = 2, while the silhouette coefficient reached its maximum at K = 2. These complementary results indicated that two clusters provided the best balance between separation and stability. Moreover, the two-cluster solution yielded clinically meaningful subgroups with distinct demographic, biochemical, and imaging characteristics, further validating the selection of K = 2.

By using the unsupervised clustering method, the optimal number of clusters that best explained the overall variance in the data was determined (cluster group 1: n=262, cluster group 2: n=375). The clinical variables and myocardial ischemia characteristics were analyzed and compared across the two cluster subgroups. The analysis was based on Python 3.9.12 with the following libraries: scikit-learn 1.0.2 for implementing K-means clustering and calculating the Silhouette Coefficient, NumPy 1.21.5 for numerical computations, and Matplotlib 3.5.1 for visualizing the Elbow Method and Silhouette Score plots. The Jupyter Notebook environment was used for code development and execution (Figure 3).

2.5 Statistical analysis

Following cluster group identification, clinical information and myocardial ischemia characteristics were compared between the groups. Statistical analyses were conducted with SPSS software (version 25.0; IBM, Armonk, NY, USA). Categorical variables are presented as counts (%), while continuous variables are expressed as mean $\pm$ standard deviation. T-tests and chi-square tests were used to compare clinical and SPECT characteristics between cluster groups. A two-tailed P value < 0.05 was considered statistically significant.

3 Result

3.1 Study population

The study included a total of 637 individuals with T2DM, of whom 68.3% (435/637) were men, with an average age of 61.65 ± 14.05 years old. Unsupervised K-means clustering analysis was performed on 29 clinical parameters of 637 T2DM subjects. The silhouette coefficient for different values of K was calculated, and the K value with the highest silhouette coefficient was selected as the optimal number of clusters. The results revealed the existence of two clinical subtypes of T2DM with myocardial ischemia, which were classified into Cluster 1 (n=262) and Cluster 2 (n=375). Table 1 presents the clinical characteristics of the participants in the two cluster groups.

3.1.1 Cluster group 1

In Cluster 1, the proportions of male patients (67.6%), smoking history (42.4%), and hypertension (34.9%) were lower than those in Cluster 2, and serum uric acid levels were significantly lower. Compared with Cluster 2, Cluster 1 participants had a higher

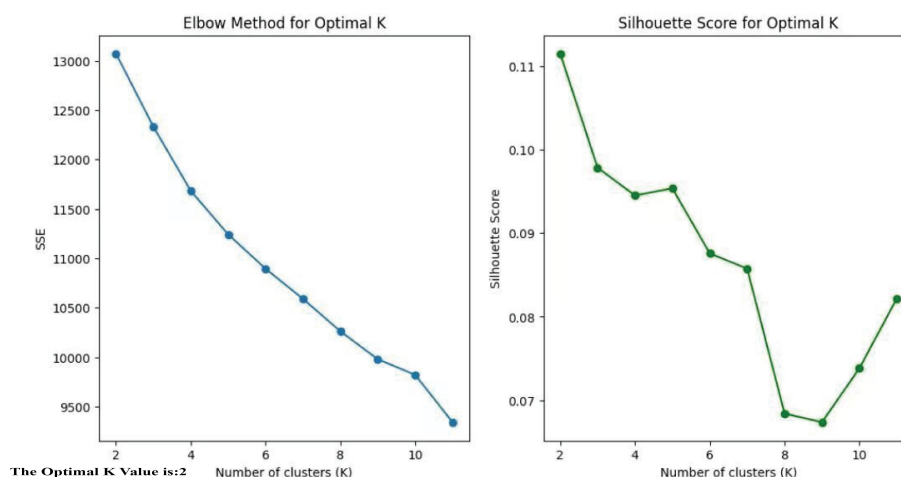


FIGURE 3
The optimal K.

TABLE 1 Baseline characteristics of the study cohort.

Variables	Cluster1 (n=262)	Cluster2 (n=375)	P value
Male(%)	177(67.6%)	258(68.8%)	0.740
Age (years old))	63.23±12.31	60.49±15.04	0.009
BMI (kg/m2)	25.22±4.03	25.46±4.43	0.469
Smoking history (%)	111(42.4%)	167(44.5%)	0.587
Alcohol (%)	74(28.2%)	82(21.9%)	0.065
Hypertension (%)	82(34.9%)	75(35.5%)	0.778
Systolic blood pressure (mmHg)	132.70±22.85	129.84±20.79	0.101
Diastolic blood pressure (mmHg)	77.30±14.71	77.81±15.32	0.666
Pulse (rate)	77.53±14.71	80.29±16.57	0.011
Diabetes duration(year)	10.27±8.77	5.51±7.61	<0.001
HbA1c (%)	7.69±1.76	6.35±1.68	<0.001
Fasting blood glucose(mmol/L)	8.29±7.26	6.44±2.62	<0.001
Cholesterol(mmol/L)	3.90±1.25	3.96±1.12	0.459
Triglyceride(mmol/L)	1.61±1.04	1.57±0.82	0.066
HDL-C(mmol/L)	2.51±1.02	2.52±0.83	0.178
LDL-C(mmol/L)	1.03±0.38	2.08±0.32	0.728
Serum uric acid (μmol/L)	390.56±130.98	406.78±135.24	0.066
Serum Creatinine(μmol/L)	115.42±106.18	106.38±102.70	0.009
Diabetes treatment (%)			
Oral			
Biguanides	64(24.4%)	71(18.9%)	0.095
Non-Sulfonylureas	7(2.7%)	14(3.7%)	0.997
α-Glucosidase inhibitor	61(23.3%)	46(12.3%)	<0.001
Sulfonylureas	21(8.0%)	15(0.4%)	<0.001
GLP-1 receptor agonist	2(0.8%)	26(6.9%)	0.076
DPP-4 inhibitor	11(4.2%)	21(5.6%)	0.713
SGLT-2 inhibitor	35(13.4%)	36(9.6%)	0.082
Stains	1(0.4%)	23(6.1%)	0.013
Without drugs	41(15.6%)	82(21.9%)	<0.001
Insulin	106(40.5%)	58(15.5%)	<0.001

Statins in Cluster 2 were likely prescribed at low-to-moderate intensity, which may partially explain the inconsistency with ischemia outcomes reported in high-dose statin clinical trials.Cluster 1 patients showed higher use of SGLT2 inhibitors and insulin, while Cluster 2 patients were more likely to receive GLP-1RA, DPP-4 inhibitors, and statins. BMI, Body Mass Index; HbA1c, HbA1c, Glycated Hemoglobin; LDL-C, low-density lipoprotein cholesterol; HDL-C, High-density lipoprotein cholesterol; GLP-1,Glucagon-Like Peptide-1receptor agonist; DPP-4 inhibitor, Dipeptidyl Peptidase-4 inhibitor; SGLT-2 inhibitor, Sodium-Glucose Co-Transporter 2 inhibitor;

prevalence of alcohol consumption history (28.2%vs21.9%) (Alcohol consumption history was defined as current drinking at the time of the study, past drinking, or both.)and a longer duration of diabetes (10.27 ± 8.77 years vs5.51 ± 7.61years). Regarding treatment patterns, patients in Cluster 1 were more likely to use metformin, α-glucosidase inhibitors, sulfonylureas, SGLT-2 inhibitors, and insulin, whereas a higher proportion of Cluster 2

patients were untreated or received GLP-1 receptor agonists and statins (Table 1).

3.1.2 Cluster group 2

Cluster group 2 patients had a reduced period of diabetes(5.51 ± 7.61years) and had the lower Rate of patients involving alcohol consumption (21.9%).A larger fraction of patients in cluster group 2

had received diabetes treatment, like non-Sulfonylureas, Glucagon-Like Peptide-1(GLP-1) receptor agonist, Dipeptidyl Peptidase-4 (DPP-4) inhibitor, Statins. Moreover, without drug treatment, Cluster 1 had 141 patients (15.6%), while Cluster 2 had 82 patients (21.9%), showing a significant difference ($p < 0.001$).

3.2 Connection between cluster identity and myocardial ischemia

The characteristics of myocardial ischemia are compared across the two clustering groups in [Table 2](#) and [Figure 4](#). In terms of Ischemia areas, cluster group 1 had the higher proportion of patients with incomplete ischemia (69.8%), Cluster group 2 had higher proportion of complete ischemia (11.5%) ([Figure 4A](#)). Wall motion in cluster group 1 tend to had higher proportion of Hypokinesis and Akinesis, while normal wall motion in cluster

group 2 seems more common ([Figure 4C](#)). For Ventricular contraction, cluster group 1 had worse contraction ([Figure 4D](#)). In terms of Cardiac systolic coordination, group 1 had a worse proportion ([Figure 4E](#)). From the patient perspective, cluster group 1 exhibited a greater rate of abnormal ESV and LVEF ([Figure 4F](#)). No significant variations were noted in the involved segments, EDV, ischemia areas, and cavity size ([Figure 4B](#)) between the two cluster groups (all P values >0.05).

4 Discussion

In this research, an unsupervised machine learning method was applied to classify T2DM patients into subgroups with different clinical profiles. Machine learning methods provide innovative approaches to integrate and analyze diverse omics data, facilitating disease-Specific biomarker discovery. These biomarkers provide the

TABLE 2 Comparison of myocardial ischemia characteristics of two cluster groups in T2DM.

	Cluster(n=262)	Cluster2(n=375)	P value
Ischemia areas			0.065
All	26(9.9%)	43(11.5%)	
Part	183(69.8%)	229(61.1%)	
Normal	53(20.2%)	103(27.5%)	
Cavity size			0.979
Enlarge	115(43.9%)	165(44%)	
Normal	147(56.1%)	210(56%)	
Wall Motion			0.005
Normal	65(24.8%)	137(36.5%)	
Hypokinesis	100(38.2%)	112(29.9%)	
Akinesis	97(37.0%)	126(33.6%)	
Ventricular contraction			0.010
Reduced	151(57.6%)	177(47.2%)	
Normal	111(42.4%)	198(52.8%)	
Cardiac systolic coordination			0.004
Impaired	167(63.7%)	196(52.3%)	
Normal	95(36.3%)	179(47.7%)	
Patient Level			
Involved segments ≥ 4	55(21.0%)	99(26.4%)	0.083
Involved segments <4	207(79.0%)	276(73.6%)	
EDV >132	134(51.1%)	173(46.1%)	0.213
EDV <132	128(48.9%)	202(53.9%)	
ESV >61	175(66.8%)	177(47.2%)	<0.001
ESV <61	87(33.2%)	198(52.8%)	
LVEF <50	213(81.3%)	199(53.1%)	<0.001
LVEF >50	49(18.7%)	176(46.9%)	

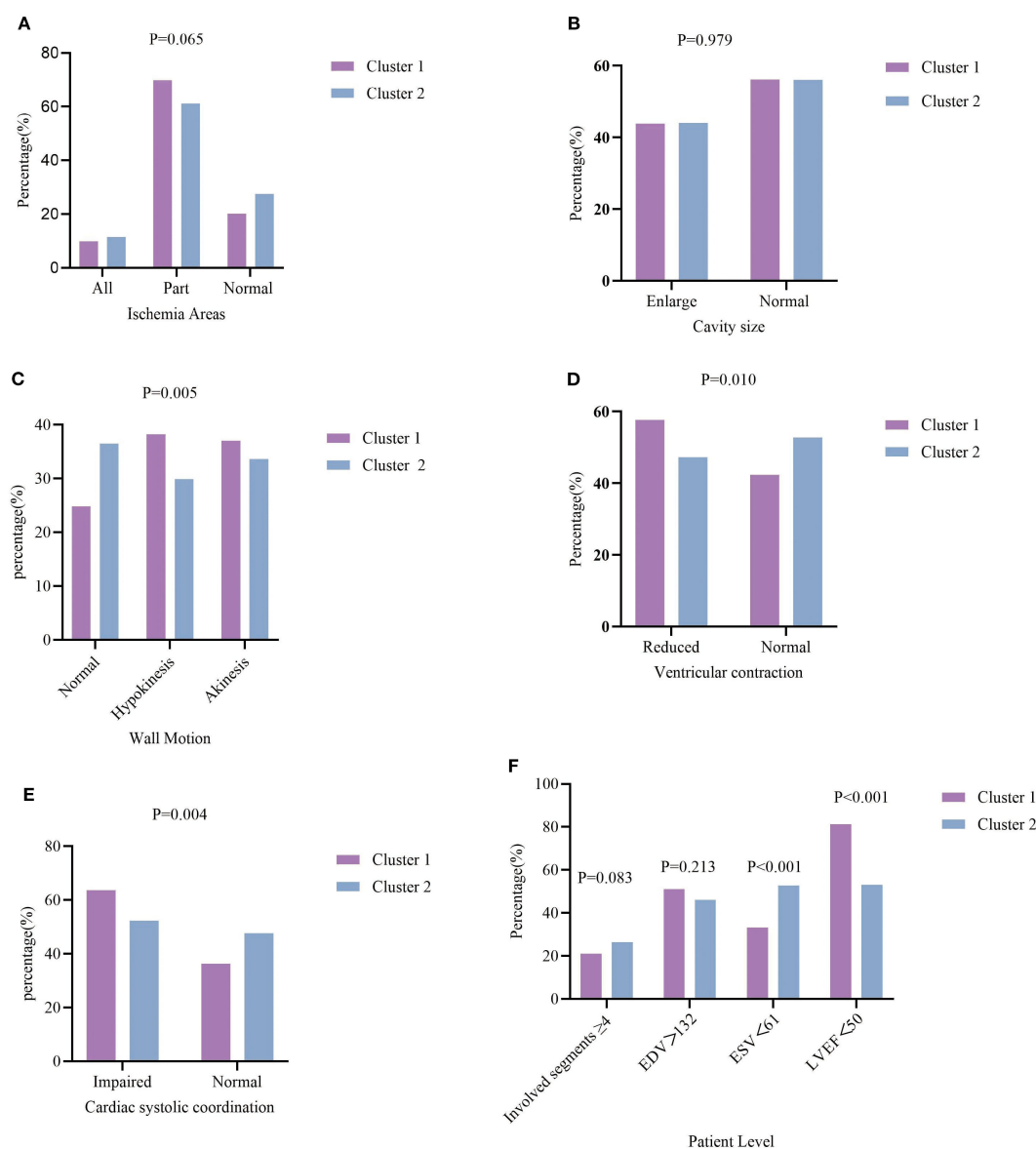


FIGURE 4 Characteristics of myocardial ischemia among the two clustering groups. (A) percentage of T2DM patients with different extent of ischemia; (B) percentage of Cavity size; (C) percentage of wall Motion; (D) percentage of ventricular contraction; (E) percentage of cardiac systolic coordination; (F) percentage of T2DM patients with involved segments ≥ 4 , EDV, ESV, LVEF.

opportunity to enhance prognostic assessment accuracy, stratified healthcare, and the delivery of precision medicine (25, 26). This study showed that unsupervised learning techniques can be applied to analyze integrated healthcare data and enable the possibility of identify distinct T2DM patient categories with varying ischemia areas and degrees of co myocardial ischemia.

4.1 Unsupervised machine learning for processing clinical data

It is widely recognized that diabetic patients are at greater risk for more aggressive vascular disease, including diffuse coronary

atherosclerosis, and exhibit a significantly increased occurrence of heart failure, myocardial infarction (MI), and cardiovascular mortality (27, 28). Diabetic patients have multitude of characteristic features, the interaction of these contributor on myocardial ischemia are attracting more focus. Preceding studies primarily aimed at the limited factors for myocardial ischemia (29–31). The target of unsupervised machine learning is to uncover clusters of patients with analogous combinations of features, free from biases introduced by clinical experts or information on future outcomes. As clinical data continue to grow rapidly, clustering methods may become increasingly valuable for analyzing the varied and multifaceted data available processing the diverse and heterogeneous data found in digital health data.

4.2 Correlation between myocardial ischemia areas and clusters

The clustering approach demonstrated the ability to not only differentiate T2DM patients with diverse clinical profiles but also to indirectly identify distinct subgroups exhibiting various types of myocardial ischemia. The findings revealed that cluster 2 had higher proportion of complete ischemia. This can be attributed to the evidence that cluster group 2 consisted of a larger number of males, exhibited more unhealthy habits such as nicotine use and alcohol consumption, and had elevated levels of LDL-C. Present smoking has been classified as a threat for myocardial ischemia (32). Mild alcohol consumption is known to be cardio protective compared with either heavy drinking or complete abstinence (33, 34). Treatment to reduce LDL cholesterol HDL-C levels is beneficial to improve ischemia (35). This result draws attention to the crucial role of effective management for T2DM patients exhibiting these risk factors for myocardial ischemia.

In line with these risk profiles, patients in Cluster 2 were more frequently prescribed GLP-1 receptor agonists and statins. GLP-1 receptor agonists have been reported to improve coronary microvascular function, while statins effectively reduce LDL-C concentrations. Nevertheless, in this cohort statin therapy was primarily administered at low-to-moderate intensities, which may have attenuated their cardioprotective effect. This limitation could partly explain the persistently higher prevalence of complete ischemia in Cluster 2, despite the seemingly more optimized pharmacological regimen.

4.3 Heart motion function in clusters

Group1 had a higher proportion of hypokinesia and akinesia wall motion, and ventricular contraction and cardiac systolic coordination also worse in group 1. This might be attributed to the fact that patients in Cluster 1 were older (63.23 ± 12.31 vs 60.49 ± 15.04 years), had a longer duration of diabetes (10.27 ± 8.77 vs 5.51 ± 7.61 years), and exhibited poorer glycemic control, as reflected by higher HbA1c levels (7.69 ± 1.76 vs 6.35 ± 1.68). A previous study showed that wall motion is an independent predictor of ischemic heart (36, 37). In both cross-sectional and short-term longitudinal studies involving older adults, the status of glycemic dysregulation add to the risk of wall motion (38, 39). Additionally, inadequate control of diabetes mellitus (DM) was associated with subclinical left ventricular (LV) dysfunction (40). Traditionally, aging is regarded as a risk factor for myocardial ischemia. We infer from our data that longstanding glycemic abnormality produces a compounded harmful influence on LV wall motion.

Although SGLT2 inhibitors were more commonly prescribed in Cluster 1 as a cardioprotective strategy, the prevalence of reduced exercise tolerance and exercise incapacity remained high. This apparent inconsistency highlights that, despite the demonstrated efficacy of SGLT2 inhibitors in randomized controlled trials, real-world effectiveness may be attenuated due to suboptimal adherence, heterogeneity in therapeutic responsiveness, or patient-specific

factors. These findings underscore the need for further evaluation of treatment strategies in routine clinical practice to mitigate LV dysfunction in high-risk T2DM populations.

4.4 Association of extent of ESV and LVEF with ischemia

Although some differences were observed in ventricular contraction, cardiac systolic coordination, ESV and LVEF, there was no statistically significant difference in the ischemic areas between the two cluster groups. A previous study demonstrated that T2DM increase the risk of death among patients with ischemic heart disease. Another studies showed that risk factors for myocardial ischemia, such as hypertension, hyperlipidemia, diabetes, smoking, obesity, age, gender, family history, can exacerbate the severity of myocardial ischemia through different mechanisms (41–43). These risk factors collectively contribute to the increased severity of myocardial ischemia. Therefore, early intervention and comprehensive management targeting these factors are crucial.

Beyond systolic function, diastolic impairment also represents a critical concern in T2DM. The disease promotes myocardial fibrosis and increases ventricular stiffness, thereby contributing to diastolic dysfunction—a hallmark mechanism of heart failure with preserved ejection fraction (HFpEF). In line with this, Cluster 1 patients demonstrated more pronounced abnormalities in left ventricular end-diastolic volume (EDV) and E/E' ratio, indicating a heightened susceptibility to diastolic dysfunction. Importantly, these patients also exhibited higher mean systolic blood pressure, longer disease duration, and poorer glycemic control, whereas Cluster 2 patients presented with slightly lower diastolic pressure. Consequently, a wider pulse pressure was evident in Cluster 1. Given that both hypertension and diabetes are key risk factors for HFpEF, these pathophysiological distinctions provide additional mechanistic support for the clustering results.

4.5 Clinical implications of unsupervised clustering

The present study demonstrates the feasibility of applying unsupervised clustering to classify T2DM patients with myocardial ischemia and underscores its potential clinical relevance. By identifying subgroups with distinct ischemic and metabolic features, clustering provides an evidence-based approach for personalized decision-making, enabling physicians to tailor therapies to specific risk profiles. Its integration of SPECT-derived functional parameters further highlights potential for early ischemia detection, thereby improving diagnostic accuracy and guiding timely interventions in high-risk populations. Beyond diagnosis, clustering may assist in predicting disease progression and stratifying cardiovascular risk, offering a valuable tool for long-term management and proactive prevention. Overall, these findings suggest that unsupervised learning could complement conventional

risk assessment and foster a more precise and individualized model of care for patients with T2DM.

5 Limitations

This study has several limitations. First, given that it is a monocentric study, Systematic bias in selection is unavoidable, and future multi-center studies are required to confirm these findings. Second, since it was a retrospective analysis, ongoing data were not available. The gradual onset and progression of myocardial ischemia in T2DM patients requires in-depth study. However, SPECT, a noninvasive examination, is one of the most common imaging modalities used for assessment of ischemia (44).

6 Conclusions

This study demonstrates that unsupervised clustering methods can effectively analyze heterogeneous clinical and imaging data from T2DM patients with myocardial ischemia. By revealing distinct patient subgroups, this approach provides a novel pathway for stratifying individuals with varying ischemic burdens. These findings highlight the methodological value of unsupervised learning for integrating complex healthcare data and suggest promising directions for future research aimed at refining risk stratification and improving outcomes in T2DM populations.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: Medical data, involving ethics and privacy. Requests to access these datasets should be directed to BL, liubing@sxmu.edu.cn.

Ethics statement

The studies involving humans were approved by Ethics Committee of Scientific Research at the First Hospital of Shanxi Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable

images or data included in this article because The written informed consent was waived due to the retrospective nature of this study.

Author contributions

BL: Writing – review & editing, Writing – original draft. Y-JH: Investigation, Writing – review & editing. PW: Writing – review & editing, Investigation. XH: Writing – review & editing. HQ: Writing – original draft. X-YY: Writing – review & editing, Software. Z-FW: Writing – review & editing. S-JL: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was supported by the National Natural Science Foundation (U22A6008) and Shanxi Province Higher Education "Billion Project" Science and Technology Guidance Project.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. . *Lancet*. (2020) 396:1204–22. doi: 10.1016/S0140-6736(20)32226-1
2. Islam SM, Purnat TD, Phuong NT, Mwingira U, Schacht K, Fröschl G. Non-communicable diseases (NCDs) in developing countries: a symposium report. *Global Health*. (2014) 10:81. doi: 10.1186/s12992-014-0081-9

3. Magliano DJ, Boyko EJ. committee IDFDates: IDF diabetes atlas. In: *Idf diabetes atlas*, vol. 2021. international Diabetes Federation© International Diabetes Federation, Brussels (2021).
4. Wong ND, Sattar N. Cardiovascular risk in diabetes mellitus: epidemiology, assessment and prevention. *Nat Rev Cardiol.* (2023) 20:685–95. doi: 10.1038/s41569-023-00877-z
5. Einarson TR, Acs A, Ludwig C, Panton UH. Prevalence of cardiovascular disease in type 2 diabetes: a systematic literature review of scientific evidence from across the world in 2007–2017. *Cardiovasc Diabetol.* (2018) 17:83. doi: 10.1186/s12933-018-0728-6
6. Tsao CW, Aday AW, Almarazooq ZI, Anderson CAM, Arora P, Avery CL, et al. Buxton AE et al: Heart Disease and Stroke Statistics–2023 Update: A Report From the American Heart Association. *Circulation.* (2023) 147:e93–e621. doi: 10.1161/CIR.0000000000001123
7. Sarwar N, Gao P, Seshasai SR, Gobin R, Kaptoge S, Di Angelantonio E, et al. Stampfer M et al: Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet.* (2010) 375:2215–22. doi: 10.1016/S0140-6736(10)60484-9
8. Gregg EW, Sattar N, Ali MK. The changing face of diabetes complications. *Lancet Diabetes Endocrinol.* (2016) 4:537–47. doi: 10.1016/S2213-8587(16)30010-9
9. Janand-Delenne B, Savin B, Habib G, Bory M, Vague P, Lassmann-Vague V. Silent myocardial ischemia in patients with diabetes: who to screen. *Diabetes Care.* (1999) 22:1396–400. doi: 10.2337/diacare.22.9.1396
10. Wackers FJ, Young LH, Inzucchi SE, Chyun DA, Davey JA, Barrett EJ, et al. Filipchuk N et al: Detection of silent myocardial ischemia in asymptomatic diabetic subjects: the DIAD study. *Diabetes Care.* (2004) 27:1954–61. doi: 10.2337/diacare.27.8.1954
11. Rosenberg DE, Jabbour SA, Goldstein BJ. Insulin resistance, diabetes and cardiovascular risk: approaches to treatment. *Diabetes Obes Metab.* (2005) 7:642–53. doi: 10.1111/j.1463-1326.2004.00446.x
12. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* (2019) 20:e262–73. doi: 10.1016/S1470-2045(19)30149-4
13. Alizadehsani R, Habibi J, Hosseini MJ, Mashayekhi H, Boghrati R, Ghandeharioun A, et al. A data mining approach for diagnosis of coronary artery disease. *Comput Methods Programs BioMed.* (2013) 111:52–61. doi: 10.1016/j.cmpb.2013.03.004
14. Nahar J, Imam T, Tickle KS, Chen Y-PP. Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. *Expert Syst Appl.* (2013) 40:96–104. doi: 10.1016/j.eswa.2012.07.032
15. Nahar J, Imam T, Tickle KS, Chen Y-PP. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Syst Appl.* (2013) 40:1086–93. doi: 10.1016/j.eswa.2012.08.028
16. Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, Yarifard AA. Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. *Comput Methods Programs BioMed.* (2017) 141:19–26. doi: 10.1016/j.cmpb.2017.01.004
17. Tsiouras MG, Exarchos TP, Fotiadis DI, Kotsia AP, Vakalis KV, Naka KK, et al. Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. *IEEE Trans Inf Technol BioMed.* (2008) 12:447–58. doi: 10.1109/TITB.2007.907985
18. Anooj PK. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *J King Saud Univ - Comput Inf Sci.* (2012) 24:27–40. doi: 10.1016/j.jksuci.2011.09.002
19. Pal D, Mandana KM, Pal S, Sarkar D, Chakraborty C. Fuzzy expert system approach for coronary artery disease screening using clinical parameters. *Knowledge-Based Syst.* (2012) 36:162–74. doi: 10.1016/j.knsys.2012.06.013
20. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement healthcare professionals Cardiac Imaging Committee Council Clin Cardiol Am Heart Assoc Circ. (2002) 105:539–42. doi: 10.1081/JCMR-120003946
21. Marie PY, Mercennier C, Danchin N, Djabballah K, Grentzinger A, Zannad F, et al. Virion JM et al: Residual exercise SPECT ischemia on treatment is a main determinant of outcome in patients with coronary artery disease treated medically at long-term with beta-blockers. *J Nucl Cardiol.* (2003) 10:361–8. doi: 10.1016/S1071-3581(03)00429-X
22. Marie PY, Danchin N, Branly F, Angioi M, Grentzinger A, Virion JM, et al. Juillière Y et al: Effects of medical therapy on outcome assessment using exercise thallium-201 single photon emission computed tomography imaging: evidence of a protective effect of beta-blocking antianginal medications. *J Am Coll Cardiol.* (1999) 34:113–21. doi: 10.1016/S0735-1097(99)00156-4
23. Hesse B, Tägil K, Cuocolo A, Anagnostopoulos C, Bardiés M, Bax J, et al. Dondi M et al: EANM/ESC procedural guidelines for myocardial perfusion imaging in nuclear cardiology. *Eur J Nucl Med Mol Imaging.* (2005) 32:855–97. doi: 10.1007/s00259-005-1779-y
24. Huang Z. Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Min Knowledge Discov.* (1998) 2:283–304. doi: 10.1023/A:1009769707641
25. Lancaster SM, Sanghi A, Wu S, Snyder MP. A customizable analysis flow in integrative multi-omics. *Biomolecules.* (2020) 10(12):1606. doi: 10.3390/biom10121606
26. Williams MC, Bednarski BP, Pieszko K, Miller RJH, Kwiecinski J, Shanbhag A, et al. Unsupervised learning to characterize patients with known coronary artery disease undergoing myocardial perfusion imaging. *Eur J Nucl Med Mol Imaging.* (2023) 50:2656–68. doi: 10.1007/s00259-023-06218-z
27. Giri S, Shaw LJ, Murthy DR, Travin MI, Miller DD, Hachamovitch R, et al. Impact of diabetes on the risk stratification using stress single-photon emission computed tomography myocardial perfusion imaging in patients with symptoms suggestive of coronary artery disease. *Circulation.* (2002) 105:32–40. doi: 10.1161/hc5001.100528
28. Hammoud T, Tanguay JF, Bourassa MG. Management of coronary artery disease: therapeutic options in patients with diabetes. *J Am Coll Cardiol.* (2000) 36:355–65. doi: 10.1016/S0735-1097(00)00732-4
29. Yang ZJ, Liu J, Ge JP, Chen L, Zhao ZG, Yang WY. Prevalence of cardiovascular disease risk factor in the Chinese population: the 2007–2008 China National Diabetes and Metabolic Disorders Study. *Eur Heart J.* (2012) 33:213–20. doi: 10.1093/eurheartj/ehz205
30. Rassi CH, Churchill TW, Tavares CA, Fahel MG, Rassi FP, Uchida AH, et al. Use of imaging and clinical data to screen for cardiovascular disease in asymptomatic diabetics. *Cardiovasc Diabetol.* (2016) 15:28. doi: 10.1186/s12933-016-0334-4
31. Schinkel AF, Poldermans D, Rizzello V, van Domburg RT, Valkema R, Elhendy A, et al. Impact of diabetes mellitus on prediction of clinical outcome after coronary revascularization by 18F-FDG SPECT in patients with ischemic left ventricular dysfunction. *J Nucl Med.* (2006) 47(1):68–73.
32. Law MR, Morris JK, Wald NJ. Environmental tobacco smoke exposure and ischaemic heart disease: an evaluation of the evidence. *Bmj.* (1997) 315:973–80. doi: 10.1136/bmj.315.7114.973
33. Elmadhun NY, Sabe AA, Lassaletta AD, Sellke FW. Ethanol promotes new vessel growth in remote nonischemic myocardium. *J Surg Res.* (2015) 193:536–42. doi: 10.1016/j.jss.2014.05.048
34. Lassaletta AD, Elmadhun NY, Liu Y, Feng J, Burgess TA, Karlson NW, et al. Ethanol promotes arteriogenesis and restores perfusion to chronically ischemic myocardium. *Circulation.* (2013) 128:S136–143. doi: 10.1161/CIRCULATIONAHA.112.000207
35. Mayor S. Moderate LDL cholesterol reduction with statins shows similar benefit to aggressive lipid lowering in ischaemic heart disease. *BMJ.* (2016) 353:i3437. doi: 10.1136/bmj.i3437
36. Levy D, Garrison RJ, Savage DD, Kannel WB, Castelli WP. Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. *N Engl J Med.* (1990) 322:1561–6. doi: 10.1056/NEJM199005313222203
37. Koren MJ, Devereux RB, Casale PN, Savage DD, Laragh JH. Relation of left ventricular mass and geometry to morbidity and mortality in uncomplicated essential hypertension. *Ann Intern Med.* (1991) 114:345–52. doi: 10.7326/0003-4819-114-5-345
38. Turkbey EB, McClelland RL, Kronmal RA, Burke GL, Bild DE, Tracy RP, et al. The impact of obesity on the left ventricle: the Multi-Ethnic Study of Atherosclerosis (MESA). *JACC Cardiovasc Imaging.* (2010) 3:266–74. doi: 10.1016/j.jcmg.2009.10.012
39. Heckbert SR, Post W, Pearson GD, Arnett DK, Gomes AS, Jerosch-Herold M, et al. Traditional cardiovascular risk factors in relation to left ventricular mass, volume, and systolic function by cardiac magnetic resonance imaging: the Multiethnic Study of Atherosclerosis. *J Am Coll Cardiol.* (2006) 48:2285–92. doi: 10.1016/j.jacc.2006.03.072
40. Gerstein HC, Miller ME, Ismail-Beigi F, Largay J, McDonald C, Lochnan HA, et al. Effects of intensive glycaemic control on ischaemic heart disease: analysis of data from the randomised, controlled ACCORD trial. *Lancet.* (2014) 384:1936–41. doi: 10.1016/S0140-6736(14)60611-5
41. Volpe M, Gallo G. Hypertension, coronary artery disease and myocardial ischemic syndromes. *Vascul Pharmacol.* (2023) 153:107230. doi: 10.1016/j.vph.2023.107230
42. Brown DF, Kinch SH, Doyle JT. Serum triglycerides in health and in ischemic heart disease. *N Engl J Med.* (1965) 273:947–52. doi: 10.1056/NEJM196510282731802
43. Magnussen C, Ojeda FM, Leong DP, Alegre-Diaz J, Amouyel P, Aviles-Santa L, et al. Bobak M et al: Global Effect of Modifiable Risk Factors on Cardiovascular Disease and Mortality. *N Engl J Med.* (2023) 389:1273–85. doi: 10.1056/NEJMoa2206916
44. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J.* (2020) 41:407–77. doi: 10.1093/eurheartj/ehz425