Echocardiographic phenotypes of Chinese patients with type 2 diabetes may indicate early diabetic myocardial disease

Zheng Wang1,2, Chen Chen Wang1,2, Zuo Ling Xie1,2, Xi Huang1, Hai Yan Shang Guan2,3, Wen Wen Zhu2 and Shao Hua Wang1*

1 Department of Endocrinology, The Affiliated Zhong Da Hospital of Southeast University, Nanjing, China; 2 School of Medicine, Southeast University, Nanjing, China; and 3 Nanjing Central Hospital, Nanjing, China

Abstract

Aim Type 2 diabetes may impair cardiac structure and function at very early stage, other factors, for example, obesity and hypertension, can induce aforementioned abnormalities individually. This study aimed to explore precise prevention and treatment of diabetic cardiomyopathy (DCM) by using cluster analysis of echocardiographic variables.

Methods and results A total of 66,536 inpatients with diabetes from 2013 to 2018 were investigated, and 7,112 patients were available for analysis after nadir. The cluster analysis was performed on echocardiographic variables to assess the clinical profiles and risk factors of clusters. Two clusters were identified. Cluster 1 with 3,576 patients (50.3%, including 62.5% female) had hypertension in 62.4%, while the lower rate of obesity (13.7%). Ultrasound findings showed that 79.9% of them had left ventricular diastolic dysfunction (LVDD), the most characteristic change in the early stages of DCM. Systolic blood pressure (SBP), uric acid and antithrombin III were independent risk factors for LVDD ($P < 0.0001$); 64.0% of the 3,536 patients in the second group were male, with a high prevalence of obesity (30.1%) and a higher prevalence of hypertension (79.5%), In particular, decreased systolic function and a high rate of LV hypertrophy (46.8%) represented the progressive phase of DCM ($P < 0.0001$). SBP, diastolic blood pressure, BMI and creatinine were independent correlates of LV mass index ($P < 0.05$).

Conclusion The cluster analysis of echocardiographic variables may improve the identification of groups of patients with similar risks and different disease courses and will facilitate the achievement of targeted early prevention and treatment of DCM.

Keywords T2DM; Diabetic cardiomyopathy; Risk factors; Cluster analysis

Received: 10 February 2022; Revised: 19 May 2022; Accepted: 21 June 2022

*Correspondence to: Shao Hua Wang, Department of Endocrinology, The Affiliated Zhong Da Hospital of Southeast University, Nanjing 210009, China.
Email: gyjwsh@126.com

Introduction

In China, the prevalence of type 2 diabetes mellitus (T2DM) is increasing yearly, reaching 11.2% of the global diabetic population,1 whereas the prevalence values of hypertension (HTN) and obesity increase yearly.2,3 HTN and obesity are common complications of T2DM. Furthermore, the abnormalities in cardiac geometry and function seen in patients with early diabetic cardiomyopathy (DCM) have many similarities to those seen in patients with HTN and obesity.2,4–6 In addition, age, a factor in T2DM, influences the early clinical features and functional changes of DCM.7,8 These causative factors can lead to changes in the ultrasound structure of the patient’s heart. The specific contribution of these causative factors and their synergistic contribution to cardiac dysfunction in T2DM are unclear. Classic statistical analyses are built on the a priori hypothesis, and cluster analysis may improve cardiac phenotyping and provide new insights into heterogeneous patient groups, such as those with T2DM, by providing an innovative exploratory analysis.9,10 Therefore, we hypothesized that cluster analysis may allow the identification of patient groups (clusters) with T2DM and common cardiac phenotypes on the basis of ultrasound data and different clinical features, providing new insights and exploring...
strategies for the management of diabetic cardiac changes and cardiomyopathy.\textsuperscript{10,11}

Methods

Study population

The study subjects were patients hospitalized in ZhongDa Hospital affiliated to Southeast University from July 2013 to the end of 2018. To avoid selection bias, we included all patients consecutively and enrolled 66,536 patients with diabetes based on primary discharge diagnosis. After excluding 5,411 cases of type 1 diabetes, gestational diabetes, and other specific types of diabetes, there were 61,125 cases of T2DM. For patients hospitalized multiple times during this period, only the first hospitalization data were used \((n = 31,112)\). After excluding 23,299 patients due to significant data gap \((n = 419)\) and cardiac ultrasound data gap \((n = 22,880)\), the remaining 78,13 patients were included. After excluding obvious rheumatic heart disease, congenital heart disease, valvular disease, coronary atherosclerotic heart disease, pericardial disease, and hyperthyroidism, 7,112 patients \((3,601 \text{ men and } 3,511 \text{ women})\) were enrolled (Supporting Information, Figure S1). The procedure of this study was approved by the Research Ethics Committee of ZhongDa Hospital affiliated to Southeast University (Approved No. of ethic committee: 2020DSYLL028-P01).

Clinical and biological data

Relevant physical examination, laboratory biochemical examination and echocardiography were included on the same day. Age, sex, diabetes course, smoking history, alcohol drinking history, height, weight, and drug use were collected from clinical data. The patient is given a complete physical examination, height, weight, and drug use were collected from clinical data. The patient is given a complete physical examination, blood pressure (systolic and diastolic) is measured, and a history of diabetes is collected. The diagnostic criteria for HTN are based on the recommendations of the American College of Cardiology (ACC)/American Heart Association (AHA) and the European Society of Cardiology (ESC)/European Society of HTN (ESH) blood pressure guidelines,\textsuperscript{12,13} and blood pressure \(>140/90\) mmHg measured more than two times in resting state after admission, based on medical history. The diagnostic criteria of obesity are based on the recommendations of the World Health Organization (WHO) and the Working Group on Obesity in China.\textsuperscript{14,15} Obesity is defined as body mass index (BMI) \(>28.0 \text{ kg/m}^2\), while overweight is defined as body mass index (BMI) \(>24.0 \text{ kg/m}^2\).

All patients had venous blood (approximately 2–4 mL of venous blood and 1–2 mL of serum extracted) collected and laboratory parameters measured on the second day of admission. Fasting blood glucose was detected by hexokinase method. HbA1c was determined by cation exchange high performance liquid chromatography. Total cholesterol (T-CHOL) was measured by cholesterol esterase/peroxidase enzyme method, and triglyceride by lipase glycerol kinase method. High density lipoprotein cholesterol (HDL-C) and low density lipoprotein cholesterol (LDL-C) were determined by direct antibody separation and elimination methods, respectively. The concentrations of apolipoprotein-A1, apolipoprotein-B and lipoprotein-A in serum were determined by immunoturbidimetry. Serum urea nitrogen (BUN) was determined by glutamate dehydrogenase and urease kinetic methods. Serum creatinine (Scr) and uric acid (SUA) levels were determined by sarcosine oxidase method and uricase method respectively. These indicators were measured by an automatic biochemical analyser (AU5800, Beckman Coulter, California, USA). Prothrombin time (PT), activated partial thrombin time (APTT), thrombin time (TT), and fibrinogen level were measured by coagulation method. The concentration of antithrombin III was determined by chromogenic substrate method, and the levels of fibrinogen degradation product (FDP) and D dimer were determined by turbidimetric method. All fibrinolytic functions were measured using an automatic analyser (ACLOT TOP 700, Beckman Coulter, California, USA). The Laboratory Center of the University Hospital implements internal and external quality management procedures directed by the China Laboratory Quality Control Center. All blood samples were analysed by clinical laboratory professionals at ZhongDa Hospital.

Echocardiogram

All echocardiographic data were obtained using commercial echocardiographic systems (Vivid E9; Ge-vingmed, Horten, Norway). The criteria for echocardiographic measurements are based on current guidelines of the American Society of Echocardiography (ASE)/European Association of Cardiovascular Imaging (EACVI). That is, a 2.5–3.5 mhz phased array probe is used by an experienced registered sonographer. Subjects were placed in partial left lateral decubitus, and images were obtained by the observer along the long and short axis of the prosthernum and apical four compartment and two compartment long axis views. All records included at least five cardiac cycles and were stored digitally for offline analysis. Offline analysis Echocardiography was analysed without knowledge of the subjects’ characteristics by a registered ultrasound physician with more than 10 years of experience. EchoPac software (GE Vingmed) is used for post-processing of digitally stored images. Mean measurements were statistically analysed over at least three cardiac cycles. Measure and record the aortic root (AO) diameter, aorta ascendens (AAO) diameter, left atrium (LA) diameter, right ventricular (RV) end-diastolic internal diameter, interventricular septum (IVS) thickness, left ventricular (LV) end-diastolic internal diameter,
left ventricular posterior wall (LVPW) thickness, diameter of pulmonary artery (PA), the right atrium (RA) diameter, mitral valve blood flow velocity (MV) tricuspid, and aortic blood flow velocity (AV), pressure gradient (PG), pulmonary valve (PV) opening velocity, and ejection fraction (EF). Left ventricular diastolic dysfunction (LVDD) was calculated and assessed according to expert consensus guidelines of the American Society of Echocardiography and the European Society of Cardiovascular Imaging.\(^{16}\)

**Calculation of variable**

Calculation of the glomerular filtration rate (GFR) was based on the guideline consensus, following the simplified MDRD formula: \( \text{GFR} = 186 \times \left( \frac{\text{Scr}}{88.402} \right)^{-1.154} \times \text{age}^{0.203} \times \left( \text{female} \times 0.742 \right).\)\(^{17,18}\) Left ventricular mass (LVM) was calculated by Devereaux’s LVM correction formula.\(^{19}\) LVM (g) = 0.8 × 1.04 × \left( \left( \text{IVST} + \text{LVPWT} + \text{LVDD} \right)^3 - \text{LVDD}^3 \right) - 0.6. Formula for calculating body surface area (BSA): \( \text{BSA} \left( m^2 \right) = 0.0061 \times \text{height} \left( \text{cm} \right) + 0.0128 \times \text{weight} \left( \text{kg} \right) - 0.1529. \) Left ventricular mass index (LVMI) (g/m\(^2\)) = LVM/BSA. The formula for calculating the relative wall thickness (RWT): \( \text{RWT} = \left( \frac{\text{LVST} + \text{LVPWT}}{\text{LVDD}} \right) \). Echocardiographic criteria for the diagnosis of LV hypertrophy using the 2015 American and European Echocardiographic Association guidelines: \( \text{LVMI} \geq 115 \) g/m\(^2\) in men and \( \text{LVMI} \geq 95 \) g/m\(^2\) in women.\(^{20}\) Hospitalization rates for all cardiovascular diseases (including cardiovascular medicine and cardiovascular surgery) were obtained from medical records of ZhongDa Hospital.

**Cluster analysis**

The initial step analysis divided patients into four clinically defined groups based on admission records and diagnostic records for the presence of HTN and/or obesity. To assess differences in baseline data, laboratory biology, and echocardiographic features among patients with T2DM. Four clinically defined groups were thus constituted of: (i) patients with T2DM alone; (ii) T2DM mellitus with obesity; (iii) patients with T2DM mellitus and HTN; and (iv) patients with T2DM mellitus with obesity and HTN.

In the second step of the analysis, we first analysed the included data by removing five variables with missing values greater than 25% (PA, MV, TR, PG, and \( e'/a' \)) from the 51 included variables and interpolating the remaining 46 variables. Parameter estimation using EM (Expectation Maximization) for great likelihood was used to reduce the loss due to missing data. These datasets were combined as complete data according to Rubin’s standard rules. For cardiac ultrasound-related metrics the raw data were transformed by Z-score normalization to meet the data standardization requirements; data that conformed to a normal distribution were taken between \(-3\) standard deviations and \(+3\) standard deviations depending on the data type to exclude abnormally high and low values \((n = 406)\); Euclidean distance was used as a similarity measure and the Ward method was used. All clusters were performed by one to two researchers with no knowledge of the clinical data.

Various unsupervised analyses were performed in turn in order to more accurately identify associations between echocardiographic phenotypic features and other clinical indicators and characteristics. Firstly, based on the collated echocardiographic data, a principal component analysis was performed in order to understand the correlation between the ultrasound data and to construct a three-dimensional principal component analysis plot describing the categorical cardiac ultrasound characteristics of the patients. Based on the results of the subsequent cluster analysis, the clustering results were evaluated using principal component analysis double-labelled plots by assigning different colours to different groups of patients.

The best number of clusters was obtained by sequentially analysing nine combinations of different numbers of clusters, based on the consistency matrix and cluster consistency scores, according to the echocardiographic characteristics of the patients. The conclusions suggest that the clustering analysis method divided into two classes has the best consistency matrix and the highest cluster consistency score. It is suggested that cluster analysis into two categories has the least intra-group variability and the greatest inter-group heterogeneity.

**Statistical analysis**

Numerical variables are expressed as mean ± standard deviation (SD), median (interquartile range), respectively, according to the presence or absence of a normal distribution. Qualitative variables are expressed as numbers (percentages). Numerical variables with and without normal distribution are expressed as mean ± standard deviation (SD), median (interquartile range), respectively. Qualitative variables are expressed as quantities (percentages). Comparisons between two groups were made using Student’s t-test for normally distributed quantitative variables, Mann–Whitney U-test for asymmetrically distributed quantitative variables, and \( \chi^2 \) test for qualitative variables. Multiple comparisons of continuous data were performed using one-way ANOVA or Kruskal–Wallis tests, with post hoc tests for paired comparisons if overall significance existed, and Sidak correction for multiplicity of tests. A threshold of statistical significance was set at a corrected two-sided \( P < 0.05 \). Missing data were interpolated using the EM algorithm, standardized using the Z-score method, analysed for extreme values and cleaned for data with \( \geq 3 \) standard deviations and \( \leq 3 \) standard deviations.

DOI: 10.1002/ehf2.14062
Data were processed, descriptively analysed and compared using SPSS 26.0 software (SPSS, Inc., Chicago, IL, USA). R statistical software (4.0.5) for cluster analysis and visualization (ConsensusClusterPlus, limma, heatmap, ggradar, scatterplot3d, and other packages), Microsoft excel (16.43.1) for case screening and Prism (8.2.1) for plotting violin plots and forest plots. The new results after clustering were data-completed using multiple interpolation. One-way analysis of variance was followed, using multi-factor regression analysis using binary unconditional logistic regression. Prevalence 95% confidence intervals (CI) were calculated based on a normal approximation of the binomial distribution. A corrected two-sided P < 0.05 was the threshold for a statistically significant difference.

Results

Traditional classification based on risk factors (presence or absence of hypertension and/or obesity)

A total of 7112 patients with T2DM were analysed. Baseline data were summarized for patients with type 2 diabetes only (n = 1788 (25.1%)), obese patients (n = 279 (3.9%)), hypertensive patients (n = 3769 (53.0%)), and obese patients with hypertension (n = 1276 (17.9%)), the characteristics of which are shown in Table 1. Compared with the other groups: the T2DM alone group had the lowest systolic and diastolic blood pressure, lowest Apo-B levels, lowest triglyceride levels, lowest creatinine levels and highest glycated haemoglobin levels; the obese group was younger, had a shorter duration of diabetes, had the highest Apo-B levels and higher LDL-C levels; the hypertensive group was older; the obese group with HTN had higher LDL-C levels, the highest triglyceride levels and the highest creatinine levels. There were no significant between-group differences in the use of antiplatelet agents, statin lipid-regulating drugs, insulin drugs and metformin in the four groups. Little variability between groups can be seen for most variables, with similar patient characteristics between groups.

Compare the echocardiographic data (Table 2) and calculated data (Table 3) among the four groups. Compared with the T2DM group alone and the obese with HTN group, the obese and hypertensive groups showed significant differences in aorta, aorta ascendens, left atrium, right ventricle, IVS, left ventricle, LVPW, pulmonary artery, right atrium mitral valve, TR, PG, pulmonary valve (m/s), E/A, e′/e′ were significantly different (P < 0.05). While the obese group and the hypertensive group, including aorta, aorta ascendens, left atrium, IVS, left ventricle, pulmonary artery, right atrium, aortic valve, PG, pulmonary valve, ejection fraction had significant overlap and were not significantly different (P > 0.05). LVMi and RWT levels were lower in T2DM mellitus group and obesity group, but higher in HTN group and obesity group with HTN, and the difference was statistically significant (P < 0.05). EGFR levels were higher in the simple T2DM group and the obesity group, but lower in the HTN group and the obesity group with HTN, and the difference was statistically significant (P < 0.05). The incidence of Left ventricular hypertrophy (LVH) was lower in T2DM mellitus group and obesity group, but higher in HTN group and obesity with HTN group (P < 0.05). Despite significant differences between mean values, an important overlap of individual values of echocardiographic parameters was observed among the four groups (Figure 1).

Cluster analysis of echocardiographic phenotypes

Echocardiographic variables identified three clusters of cardiac phenotypes (Figures 2A, 2B, 3, and 4).

Analysis of cardiac phenotypes and clinical phenotypes according to patient clusters

After clustering the two groups identified according to the echocardiographic phenotype data, the echocardiographic data (Table 4) and clinical characteristics data (Table 5) of the two clusters were further compared, and it was evident that there was significant heterogeneity between the echocardiographic and clinical characteristics groups of the two clustered patients (Figure 5).

Figure 6 shows and describes the profile of echocardiographic phenotypes and clinical indicators using radar plots.

Cluster 1 (female and diastolic dysfunction more) with smaller aortic internal diameter (2.51 ± 0.43), ascending aortic internal diameter (3.16 ± 0.33), pulmonary artery internal diameter (2.25 ± 0.28), left atrial internal diameter (3.43 ± 0.34), left intraventricular diameter (4.31 ± 0.35), right atrial internal diameter (3.50 (3.30, 3.70]), right intraventricular diameter (2.22 ± 0.21), lower aortic orifice flow velocity [1.22 (1.10, 1.40)], pulmonary orifice flow velocity (0.95 ± 0.17); there was a smaller septal thickness (0.97 ± 0.12), posterior left ventricular wall thickness (0.94 ± 0.10), left ventricular mass index (82.93 ± 14.64), relative ventricular wall thickness (0.45 ± 0.06), and a high ejection fraction (0.69 ± 0.06), presenting a ‘small, thin, slow heart’. The clinical features of T2DM mellitus without obesity and HTN are usually lean with a low prevalence of HTN, low prevalence of smoking and alcohol consumption, low levels of urea nitrogen, creatinine and uric acid, and high glomerular filtration rate; however, total cholesterol, HDL-C, LDL-C, Apo-A1, and Apo-B levels are high, and glycated haemoglobin levels are high (P < 0.05).
Cluster 2 (more men, higher LVMi) with larger aortic internal diameter (2.74 ± 0.45), ascending aortic internal diameter (3.41 ± 0.33), pulmonary artery internal diameter (2.45 ± 0.31), left atrial internal diameter (3.98 ± 0.36), left intraventricular diameter (4.78 ± 0.39), right atrial internal diameter (3.90±3.70, 4.13), right intraventricular diameter (2.48 ± 0.21), higher aortic orifice flow velocity [1.27(1.10, 1.43)], and pulmonary orifice flow velocity (0.97 ± 0.18); and there was greater septal thickness (1.11 ± 0.12), posterior left ventricular wall thickness (1.08 ± 0.11), left ventricular mass index (108.31 ± 20.56), and relative ventricular wall thickness (0.46 ± 0.06), and a low ejection fraction (0.67 ± 0.06), presenting a ‘large, thick, urgent’ heart. The clinical features of T2DM mellitus are more common in combination with obesity and HTN, resulting in a high prevalence of obesity and HTN, a high prevalence of smoking and alcohol consumption, high levels of urea nitrogen, creatinine and uric acid, and a low glomerular filtration rate, but low levels of total cholesterol, LDL-C, low-density lipoprotein cholesterol; PT, prothrombin time; SBP, systolic blood pressure; Scr, serum creatinine; SUA, serum uric acid; T-CHO, total cholesterol; T2DM, type 2 diabetes mellitus; TT, thrombin time.

### Analysis of risk factors for each cluster

DCM is characterized by a decrease in myocardial compliance and diastolic function in the early stages of the disease, with further development of myocardial hypertrophy and even systolic dysfunction.\textsuperscript{21,22} To further explore the risk factors

---

**Table 1 Clinical and biological characteristics**

|                | All patients | Isolated T2DM |
|----------------|--------------|---------------|
|                | (N = 7112)   | (n = 1788)    |
| **Clinical characteristics** |              |               |
| Female, n (%)  | 7112         | 3511/7112 (49.4) | 845/1788 (47.3) |
| Age (year)     | 7112         | 67.5 ± 10.23, n = 7112 | 66.05 ± 10.54, n = 1788 |
| Diabetes duration (year) | 6014 8 (6, 9), n = 6014 | 8 (7, 9), n = 1484 |
| Hypertension, n (%) | 7112         | 5045/7112 (70.9) | - |
| SBP (mmHg)     | 6382         | 136.3 ± 19.4, n = 6382 | 127.2 ± 16.1, n = 1633 |
| DBP (mmHg)     | 3880         | 78.0 ± 11.6, n = 3880 | 74.6 ± 10.1, n = 916 |
| Smoking, n (%) | 7028         | 1558/7028 (22.2) | 390/1766 (22.1) |
| Drinking, n (%) | 7026         | 922/7026 (13.1) | 217/1765 (12.3) |
| BMI (kg/m\(^2\)) | 7112         | 25.24 ± 3.76, n = 7112 | 23.17 ± 2.77, n = 1788 |
| Obesity (≥28.0 kg/m\(^2\)) | 7112         | 1555/7112 (21.9) | - |
| **Metabolic factor** |              |               |
| Apolipoprotein-A1 (g/L) | 6564         | 1.09 ± 0.28, n = 6564 | 1.1 ± 0.31, n = 1619 |
| Apolipoprotein-B (g/L) | 6564         | 0.84 ± 0.25, n = 6564 | 0.82 ± 0.25, n = 1619 |
| T-CHO (mmol/L) | 6710         | 4.6 ± 1.27, n = 6710 | 4.54 ± 1.26, n = 1656 |
| HDL-C (mmol/L) | 6564         | 1.15 ± 0.3, n = 6564 | 1.17 ± 0.34, n = 1619 |
| LDL-C (mmol/L) | 6564         | 2.84 ± 0.96, n = 6564 | 2.79 ± 0.95, n = 1619 |
| Triglyceride (mmol/L) | 6709         | 1.52 (1.06, 2.22), n = 6709 | 1.20 (0.82, 1.70), n = 1656 |
| Lipoprotein-a (mg/L) | 6564        | 187 (101, 334), n = 6564 | 187 (194, 299), n = 1619 |
| BUN (mmol/L) | 6940         | 5.6 (4.5, 7.1), n = 6940 | 5.5 (4.3, 6.9), n = 1734 |
| SCR (μmol/L) | 7034         | 76 (63, 93), n = 7034 | 72 (58, 86), n = 1763 |
| SUA (μmol/L) | 7034         | 309.92 ± 106.98, n = 7034 | 281.74 ± 99.88, n = 1763 |
| FBG (mmol/L) | 7034         | 8.25 (6.47, 11.49), n = 7034 | 8.23 (6.15, 11.10), n = 1763 |
| HbA1c (%) | 5749         | 8.03 ± 2.07, n = 5749 | 8.42 ± 2.31, n = 1439 |
| **Blood clotting index** |              |               |
| PT (s) | 6415         | 11.1 (10.4, 11.9), n = 6415 | 11.2 (10.5, 12.1), n = 1610 |
| APTT (s) | 6415         | 27.2 (29.7, 32.3), n = 6415 | 30.6 (28.3, 32.7), n = 1610 |
| TT (s) | 6411         | 15.31 ± 3.17, n = 6411 | 15.49 ± 2.9, n = 1609 |
| INR | 6406 | 1.05 ± 0.19, n = 6406 | 1.06 ± 0.21, n = 1606 |
| ATIII (%) | 6411 | 96 (87, 110), n = 6411 | 97 (82, 109), n = 1608 |
| FDP (mg/L) | 6392         | 1.6 (0.9, 2.97), n = 6392 | 1.5 (0.82, 3.1), n = 1603 |
| Fibrinogen (g/L) | 6394 | 3.76 ± 0.92, n = 6394 | 3.69 ± 0.95, n = 1605 |
| D dimer (μg/L) | 6072 | 128 (38, 340), n = 6072 | 121 (11,363), n = 1541 |
| **Medications** | | | |
| Antiplatelet user, n (%) | 6222 | 2599/6222 (41.1) | 609/1594 (38.2) |
| Statin user, n (%) | 6323 | 2237/6323 (35.4) | 539/1592 (33.9) |
| Insulin user, n (%) | 6337 | 4291/6337 (67.7) | 1126/1596 (70.6) |
| Metformin user, n (%) | 6329 | 1640/6329 (25.9) | 418/1595 (26.2) |

Values are mean ± SD, median (interquartile range), or n/N (% of non-missing data). Bold results are statistically significant at the \(P < 0.05\) level. Sidak correction: \(†P < 0.05\) compared with T2DM; \(‡P < 0.05\) compared with T2DM and obesity; \§P < 0.05\) compared with T2DM and HTN.

APTT, activated partial thromboplastin time; ATIII, Antithrombin III; BUN, blood urea nitrogen; DBP, diastolic blood pressure; FBG, fasting blood-glucose; FDP, fibrinogen degradation product; HbA1c, glycosylated haemoglobin; HDL-C, high-density lipoprotein cholesterol; HTN, hypertension; INR, international normalized ratio; LDL-C, low-density lipoprotein cholesterol; PT, prothrombin time; SBP, systolic blood pressure; Scr, serum creatinine; SUA, serum uric acid; T-CHO, total cholesterol; T2DM, type 2 diabetes mellitus; TT, thrombin time.

ESC Heart Failure 2022; 9: 3327–3344
DOI: 10.1002/ehf2.14062
in the two clustered groups of patients, we analysed patients in Cluster 1 with reduced left ventricular diastolic function as the independent variable and patients in Cluster 2 with LVMi as the independent variable.

Data from patients in Cluster 1 were further processed for analysis of covariance to remove variables with covariate associations. After the univariate analysis (Supporting Information, Table S2), the independent variable was taken as the test of <0.1 and the LVDD was used as the dependent variable to perform a binary logistic regression analysis. Figure 7 combined with forest plots yielded age, systolic blood pressure, uric acid level, and ATIII as risk factors in the Cluster 1 (female and diastolic dysfunction more) group (P < 0.05).

Patient data from Cluster 2 were further processed for analysis of covariance to remove variables with covariate correlation. After conducting a univariate analysis, the independent variables were taken as the test of <0.1 and the LVMi was used as the dependent variable to perform a multiple linear regression analysis (Supporting Information, Table S2). Table 6 suggests that in the Cluster 2 (more men, higher LVMi) group, the factors associated with higher LVMi were: higher systolic blood pressure and creatinine levels, and lower diastolic blood pressure and BMI levels, respectively (P < 0.05).

### Discussion

Computer learning methods are divided into supervised and unsupervised learning methods. Supervised learning is the
Table 2  Echocardiographic variables

|                      | All patients | Isolated T2DM |
|----------------------|--------------|---------------|
|                      | (N = 7112)   | (n = 1788)    |
| Aorta (cm)           | 6054         | 6054          |
| Aorta ascendens (cm) | 5,986        | 5,986         |
| Left atrium (cm)     | 6126         | 6126          |
| Right ventricle (cm) | 5980         | 5980          |
| IVS (cm)             | 6323         | 6323          |
| Left ventricle (cm)  | 6156         | 6156          |
| LV PW (cm)           | 6329         | 6329          |
| Pulmonary artery (cm)| 4842         | 4842          |
| Right atrium (cm)    | 6016         | 6016          |
| Mitral valve (m/s)   | 755          | 755           |
| Aortic valve (m/s)   | 6071         | 6071          |
| TR                   | 2104         | 2104          |
| PG                   | 2217         | 2217          |
| Pulmonary valve (m/s)| 6059         | 6059          |
| Ejection fraction (%)| 5591         | 5591          |
| LVPW (cm)            | 2217         | 2217          |
| Right ventricle (cm) | 6016         | 6016          |
| Left atrium (cm)     | 6071         | 6071          |
| Mitral valve (m/s)   | 5784         | 5784          |
| Aorta (cm)           | 3749         | 3749          |

Values are mean ± SD, median (interquartile range), or n/N (% of non-missing data). Bold results are statistically significant at the P < 0.05 level. Sidak correction: †P < 0.05 compared with T2DM; ‡P < 0.05 compared with T2DM and obesity; §P < 0.05 compared with T2DM and HTN.

Table 2 (continued)

|                      | T2DM and obesity | T2DM and HTN | T2DM, Obesity and HTN | P-value* |
|----------------------|-------------------|--------------|------------------------|---------|
|                      | (n = 279)         | (n = 3769)   | (n = 1276)             |         |
| Aorta (cm)           | 2.69 ± 0.43†, n = 239 | 2.63 ± 0.48†, n = 3188 | 2.69 ± 0.49§, n = 1095 | <0.0001 |
| Aorta ascendens (cm) | 3.28 ± 0.42†, n = 225 | 3.31 ± 0.38†, n = 3176 | 3.4 ± 0.39†§, n = 1088 | <0.0001 |
| Left atrium (cm)     | 3.75 ± 0.44†, n = 241 | 3.73 ± 0.48†, n = 3240 | 3.94 ± 0.47†§, n = 1113 | <0.0001 |
| Right ventricle (cm) | 2.41 ± 0.35†, n = 238 | 2.35 ± 0.27†‡, n = 3184 | 2.44 ± 0.26†§, n = 1066 | <0.0001 |
| IVS (cm)             | 1.03 ± 0.23†, n = 244 | 1.06 ± 0.15†‡, n = 3353 | 1.1 ± 0.16†§, n = 1137 | <0.0001 |
| Left ventricle (cm)  | 4.63 ± 0.43†, n = 235 | 4.55 ± 0.47†, n = 3291 | 4.71 ± 0.49†§, n = 1086 | <0.0001 |
| LV PW (cm)           | 1.00 ± 0.12†, n = 247 | 1.02 ± 0.14†‡, n = 3343 | 1.07 ± 0.15†§, n = 1143 | <0.0001 |
| Pulmonary artery (cm)| 2.41 ± 0.31†, n = 177 | 2.36 ± 0.32†, n = 2576 | 2.47 ± 0.34†§, n = 885 | <0.0001 |
| Right atrium (cm)    | 3.83 (3.49, 4.27)†§, n = 246 | 3.80 (3.54, 4.20)†‡, n = 3186 | 3.95 (3.66, 4.25)†§§, n = 1080 | <0.0001 |
| Mitral valve (m/s)   | 0.89 ± 0.54, n = 21 | 0.71 ± 0.21, n = 418 | 0.73 ± 0.37, n = 131 | 0.001   |
| Aortic valve (m/s)   | 1.20 (1.08, 1.37), n = 230 | 1.27 (1.12, 1.43)‡‡, n = 3209 | 1.28 (1.10, 1.43)‡‡, n = 1102 | <0.0001 |
| TR                   | 2.67 (2.38, 2.80), n = 70 | 2.60 (2.44, 2.90)†‡, n = 1156 | 2.61 (2.40, 2.84)†‡, n = 345 | <0.0001 |
| PG                   | 26.91 ± 7.24, n = 70 | 26.93 ± 9.66, n = 1230 | 28.21 ± 7.92, n = 361 | 0.003   |
| Pulmonary valve (m/s)| 0.95 ± 0.16, n = 225 | 0.98 ± 0.26, n = 3202 | 0.98 ± 0.21†, n = 1094 | <0.0001 |
| Ejection fraction (%)| 0.68 ± 0.06, n = 228 | 0.68 ± 0.07, n = 2946 | 0.68 ± 0.06, n = 1022 | 0.313   |
| E/A ≤ 1, n (%)       | 186/213 (87.3)†  | 295/3132 (94.4)‡‡ | 969/1042 (93.0)§‡ | <0.0001 |
| e/a<1, n (%)         | 116/135 (85.9)† | 1965/2031 (96.8)‡‡ | 642/666 (96.4)‡‡ | <0.0001 |

Values are mean ± SD, median (interquartile range), or n/N (% of non-missing data). Bold results are statistically significant at the P < 0.05 level. Sidak correction: †P < 0.05 compared with T2DM; ‡P < 0.05 compared with T2DM and obesity; §P < 0.05 compared with T2DM and HTN.

A, peak late diastolic velocity; e′, mitral annular early diastolic velocity; E, peak early diastolic velocity; HTN, hypertension; IVS, interventricular septum; LV PW, left ventricular posterior wall; PG, pressure gradient; T2DM, type 2 diabetes mellitus; TR, tricuspid regurgitation.

classification or prediction of a given learning or outcome. Another type of unsupervised learning analyses the internal structure and connections in the data.23 The cluster analysis is an unsupervised learning process. Different from common classification methods, the cluster analysis does not rely on pre-classification in accordance with clinically defined categories but requires machine learning algorithms to determine markers on the basis of non-prior data sets and classify data into different clusters automatically.24 The cluster analysis based on machine learning has been validated in clinical practice and has profound importance in identifying clinical phenotypes and differentiating heterogeneity.25–26 Previous relevant studies have also suggested the clinical value of clustering studies of echocardiographic correlates, leading to different clustering results and clinical significance.10 For example, unsupervised cluster analysis was used to propose a unique grouping model for evaluating left ventricular diastolic dysfunction. These cluster models can better identify patient groups with similar risk, and their inclusion in clinical practice may help eliminate uncertain results and improve
Table 3. Computed variables

|                  | All patients | Isolated T2DM | T2DM and obesity | T2DM and HTN | T2DM, obesity, and HTN |
|------------------|--------------|---------------|------------------|--------------|------------------------|
| N                | (n = 7112)   | (n = 1788)    | (n = 279)        | (n = 3769)   | (n = 1276)             |
| LVMi (g/m²)      | 5467         | 88.62 ± 19.72 | 89.34 ± 24.50    | 90.85 ± 25.72| 89.95 ± 33.99          |
| RWT              | 5467         | 0.44 ± 0.08   | 0.44 ± 0.08      | 0.46 ± 0.13  | 0.46 ± 0.07            |
| LVH, n (%)       | 5467         | 78.12 ± 31.65 | 89.81 ± 31.65    | 89.81 ± 31.65| 89.81 ± 31.65          |
| eGFR (mL/min/1.73 m²) | 5467        | 143.12 ± 33.77 | 143.12 ± 33.77  | 143.12 ± 33.77| 143.12 ± 33.77         |
| LVH, eGFR (%)    | 5467         | 141.23 ± 19.42| 141.23 ± 19.42   | 141.23 ± 19.42| 141.23 ± 19.42         |
| Values are mean ± SD or N/ (% of non-missing data). Bold results are statistically significant at the P < 0.05 level. Sidak correction: IP < 0.05 compared with T2DM, IP < 0.05 compared with T2DM and obesity, IP < 0.05 compared with T2DM and HTN. LVMi, left ventricular mass index; RWT, relative wall thickness; LVM, left ventricular hypertrophy; eGFR, estimated glomerular filtration rate; LVH, left ventricular mass index; RWT, relative wall thickness; eGFR, estimated glomerular filtration rate; LVH, left ventricular mass index; LVM, left ventricular hypertrophy.

DCM is the widespread focal necrosis of the myocardium caused by metabolic disorders and microvascular lesions, which gradually lead to diastolic and contractile dysfunction. This phenomenon eventually leads to heart failure or even sudden death. Studies showed that DCM has diastolic function and structural abnormalities caused by metabolic disorders. The earliest features of the disease are impaired cardiac compliance, early decreased diastolic function, and increased atrial filling. Diastolic function is further impaired by metabolic changes, neurohumoral activation, and development of myocardial fibrosis in the advanced and late stages of DCM. Systolic dysfunction and heart failure with reduced ejection fraction are finally present. Thus, diastolic dysfunction, myocardial hypertrophy, and systolic dysfunction are successively established as pathological features of DCM. However, heterogeneity is observed in related cardiovascular risk factors, such as age, duration of diabetes, obesity, and HTN, in patients with T2DM. These confounding factors influence the echocardiographic data in the daily clinical practice of patients with T2DM. T2DM with HTN, T2DM with obesity, or both with HTN further affect the occurrence and development of DCM, and whether these factors accelerate the occurrence and development of DCM due to the synergistic effect and interaction caused by the disorder of glucose and lipid metabolism. Therefore, distinguishing the risk factors of DCM in the early and middle stages and carrying out clinical intervention strategies are challenging. This research is the first study to apply cluster analysis in the identification of different cardiac phenotypes and determine their correlation with clinical characteristics in a Chinese population with T2DM. Interestingly, the cardiac characteristics of the two groups of patients in our echocardiographic cluster analysis correspond to the early and mid-
Figure 1 Echocardiographic morphological and functional parameters. The distribution of each index parameter in the echocardiogram shows significant overlap in the groups of values when plotted according to the presence or absence of obesity and/or hypertension in the patient. The violin plot shows the median (horizontal midline) and interquartile distance (upper and lower horizontal lines), with the width indicating the amount of data. 

†P < 0.05 compared with T2DM; ‡P < 0.05 compared with T2DM and obesity; §P < 0.05 compared with T2DM and HTN. DM, diabetes mellitus; HTN, hypertension; IVS, interventricular septum; LVMi, left ventricular mass index; LVPW, left ventricular posterior wall.
Figure 2  (A) Consistency matrix plot for cluster analysis. Consistency matrix plots are shown for the cluster analysis using categories 1 to 9, respectively. A tree diagram is also used to illustrate the combinations between the different variables and the distances between the groups. The size of the matrix plot area represents the sample size. (B) Consistency scores for cluster analysis. The horizontal coordinate is the number of clusters and the vertical coordinate is the cluster consistency score.
dle changes in DCM. Guiding clinical identification and intervention strategies is important.

In our study, patients in Cluster 1 showed a high rate of LVDD and a low decline rate of hypertrophic systolic function, and these findings were consistent with the echocardiographic characteristics of the early identification of DCM. Moreover, the cardiac ultrasound indicators of these patients were all better than those in Cluster 2 and were in the low-risk group of DCM. In addition, the clinical characteristics of this type of patients were evidently heterogeneous and were more common in simple patients with T2DM but no obesity and HTN. Gender analysis indicated that the majority of patients were female, and the BMI analysis indicated thin body; low rates of smoking and alcohol drinking; and low levels of urea nitrogen, creatinine, and uric acid. In addition, we found that Cluster 1 had inferior glucose and lipid metabolism control and higher levels of HbA1c, apolipoprotein-A1, total cholesterol, and low-density lipoprotein compared with Cluster 2, leading to increased prevalence of hypercholesterolemia, low-density to high-density lipoproteinemia, and low-density lipoproteinemia. At the same time, the low utilization rate of statins and insulin in Cluster 1 might be related to the poor control of glucolipid metabolism indicators. Therefore, for this group of patients, our clinical treatment strategy should emphasize the early intervention of blood glucose and blood lipids. By further analysing the decline in diastolic function in this group of patients, we found that blood pressure (especially SBP), uric acid levels, and testing

Figure 3  Heat map of a scheme based on echocardiographic variables for two classification. Red indicates an increase in standardized values expressed in SD; blue indicates a decrease in standardized values expressed in SD. The horizontal coordinates represent different patients and the vertical coordinates represent different echocardiographic variables. The bar attached at the top of the figure identifies and colours two clusters: Cluster 1 (blue), Cluster 2 (red). IVS, interventricular septum; LVPW, left ventricular posterior wall.
Figure 4  Three-dimensional principal component analysis graph. Visual representations were used to construct relationships between cluster variables while displaying patients (points) based on their individual echocardiographic characteristics. The results were projected onto a three-dimensional principal component analysis plot. The colours observed correspond to the two-group solutions from the cluster analysis, suggesting a very clear delineation between individuals of the clustering solutions based on the two classifications of the patient’s echocardiographic characteristics.

| Table 4  Cardiac phenotypes according to patient clusters |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|                            | N              | Cluster 1                 | Cluster 2                 | P-value*                     |
| Aorta (cm)                 | 6054           | 2.51 ± 0.43, n = 3008     | 2.74 ± 0.45, n = 2690     | <0.0001                     |
| Aorta ascendens (cm)       | 5986           | 3.16 ± 0.33, n = 2978     | 3.41 ± 0.33, n = 2653     | <0.0001                     |
| Left atrium (cm)           | 6126           | 3.43 ± 0.34, n = 3036     | 3.98 ± 0.36, n = 2731     | <0.0001                     |
| Right ventricle (cm)       | 5980           | 2.22 ± 0.21, n = 2977     | 2.48 ± 0.21, n = 2647     | <0.0001                     |
| IVS (cm)                   | 6323           | 0.97 ± 0.12, n = 3166     | 1.11 ± 0.12, n = 2789     | <0.0001                     |
| Left ventricle (cm)        | 6156           | 4.31 ± 0.35, n = 3071     | 4.78 ± 0.39, n = 2722     | <0.0001                     |
| LVPW (cm)                  | 6329           | 0.94 ± 0.1, n = 3144      | 1.08 ± 0.11, n = 2812     | <0.0001                     |
| Pulmonary artery (cm)      | 4842           | 2.25 ± 0.28, n = 2344     | 2.45 ± 0.31, n = 2186     | <0.0001                     |
| Right atrium (cm)          | 6016           | 3.50 (3.30, 3.70), n = 2999| 3.90 (3.70, 4.13), n = 2674| <0.0001                     |
| Aortic valve (m/s)         | 6071           | 1.22 (1.10, 1.40), n = 3001| 1.27 (1.10, 1.43), n = 2732| <0.0001                     |
| Pulmonary valve (m/s)      | 6059           | 0.95 ± 0.17, n = 2984     | 0.97 ± 0.18, n = 2731     | <0.0001                     |
| Ejection fraction (%)      | 5591           | 0.69 ± 0.06, n = 2770     | 0.67 ± 0.06, n = 2494     | <0.0001                     |
| E/A ≤ 1, n (%)             | 5784           | 2697/2942 (91.7)          | 2387/2580 (92.5)          | 0.245                       |
| LVMi (g/m²)                | 5467           | 82.93 ± 14.64, n = 2736   | 108.31 ± 20.56, n = 2406  | <0.0001                     |
| RWT                        | 5467           | 0.45 ± 0.06, n = 2736     | 0.46 ± 0.06, n = 2406     | <0.0001                     |

Values are mean ± SD, median (interquartile range). Bold results are statistically significant at the \( P < 0.05 \) level.

A, peak late diastolic velocity; E, peak early diastolic velocity; e', mitral annular early diastolic velocity; eGFR, estimated glomerular filtration rate; HTN, hypertension; IVS, interventricular septum; LVH, left ventricular hypertrophy; LVMi, left ventricular mass index; LVPW, left ventricular posterior wall; PG, pressure gradient; RWT, relative wall thickness; T2DM, type 2 diabetes mellitus; TR, tricuspid regurgitation.

ESC Heart Failure 2022; 9: 3327–3344
DOI: 10.1002/ehf2.14062
For antithrombin III might be risk factors for the decline in diastolic function.

Patients in Cluster 2 showed a high rate of hypertrophic systolic function decline. According to the occurrence and development of DCM, patients with reorganization might have been in the progression of DCM and were in the high-risk group of DCM. In comparison with those of patients in Cluster 1, the clinical characteristics of patients in Cluster 2 were
**Figure 5** Distribution of echocardiographic morphological and functional parameters after clustering. When patients are plotted according to the two-category clustering scheme, there are significant differences in individual values. Violin plots show median (horizontal midline) and interquartile spacing (upper and lower horizontal lines), with widths indicating the amount of data. ****\( p < 0.0001 \) Comparison between the two groups. DM, diabetes mellitus; HTN, hypertension; IVS, interventricular septum; LVMi, left ventricular mass index; LVPW, left ventricular posterior wall.

**Legend:**
- LVMi
- Left atrium
- Aorta
- Aorta ascendens
- Right ventricle
- IVS
- Left ventricle
- LVPW
- Right atrium
- Aortic valve

**Variables:**
- Left atrium (cm)
- Aorta (cm)
- Aorta ascendens (cm)
- Right ventricle (cm)
- IVS (cm)
- Left ventricle (cm)
- LVPW (cm)
- Right atrium (cm)
- Aortic valve (cm)
also evidently heterogeneous with more patients complicated with obesity and HTN. Gender analysis suggested that most patients were male, and the BMI analysis suggested that these patients were overweight and had high rates of smoking and alcohol drinking and high levels of urea nitrogen, creatinine, and uric acid. These findings indicated the necessity of lowering blood pressure, reducing weight, and improving bad living habits in such patients. In addition, we found that patients in Cluster 2 had lower levels of HbA1c, apolipoprotein-A1, total cholesterol, and low-density lipoprotein. Thus, the prevalence of hypercholesterolemia, low-high-density lipoproteinemia, and low-density lipoproteinemia in patients in Cluster 2 was lower than that in Cluster 1, indicating that the control of glucose and lipid metabolism of patients in Cluster 2 was better than that in Cluster 1. Combined with our findings, this result might be related to the clinical use of statins, lipid-lowering drugs, and insulin. Therefore, for this group of patients, our clinical treatment should emphasize intensive treatment of weight and blood pressure, smoking cessation, and alcohol restriction, and the control of urea nitrogen, creatinine, and uric acid levels might have some improvement in DCM with reduced hypertrophic systolic function. The multiple linear regression of the LVMI in this group of patients also suggested that blood pressure (SBP and DBP), BMI, and creatinine levels might be important in the course of DCM with hypertrophic systolic dysfunction. Thus, our data suggested that clinicians should pay attention to the treatment of female patients with decreased diastolic function in the early stages of DCM and treatment of male patients with HTN, obesity, and hypertrophic systolic dysfunction in the progressive stages of DCM (Figure 8).

In conclusion, our findings suggested that the new clustering of altered cardiac function and structure in patients with T2DM is superior to traditional classification based on risk factors because the former allows for the improved identification of early clinical features in patients with DCM and
provides information about risk factors. This study demonstrates the potential clinical value of a powerful algorithm incorporating artificial intelligence to improve the identification and use of alterations in the cardiac phenotype of patients with T2DM for the determination of different clinical features at different stages of DCM. Moreover, clustering can be easily applied to existing cohorts of patients with T2DM, and this new classification may ultimately help to bring benefits for targeted and tailored early treatment of cardiomyopathy in patients with T2DM, thus representing an important step towards precision medicine in diabetes.

Study limitations

First, this was a retrospective analysis with a statistical study of a large number of patients. There is a lack of further prospective studies on changes in different clinical indicators and cardiovascular endpoints in the two categories of patients screened by the cluster analysis, and the value of the cluster analysis in providing prognostic information lacks further validation. Secondly, this is a single-centre study, with the majority of patients originating from a specific region, and information on patients from other regions needs to be added to further enrich the study area. Thirdly, during the 5-year clinical data collection process, there was a lack of inclusion of cardiac-related indicators such as patients’ ambulatory ECG and BNP, making the results of the cluster analysis lacking clinical analysis for comparison of these two indicators on cardiac function. Fourth, the random absence of data during the clinical data collection process left a proportion of patients without sufficiently complete data. For example, the TR signal used to calculate TR velocity may not be available for many patients. For the small number of randomly missing data, we used the technique of interpolating the missing data. Finally, despite our strict enforcement of exclusion.

Figure 7 Forest plot of risk factors for reduced left ventricular diastolic function in the cluster two groups of patients. All covariables listed were included in the model simultaneously. *Numbers for continuous variables are values of 1 standard deviation. T2DM, type 2 diabetes mellitus; HTN, hypertension; SBP, systolic blood pressure; DBP, diastolic blood pressure; T-CHO, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; BUN, blood urea nitrogen; Scr, serum creatinine; SUA, serum uric acid; FBG, fasting blood-glucose; HbA1c, glycosylated haemoglobin; PT, prothrombin time; APTT, activated partial thromboplastin time; TT, thrombin time; INR, international normalized ratio; ATIII, Antithrombin III; FDP, fibrinogen degradation product.

Table 6 Multiple linear regression for LVMi in cluster 2

| Risk factors            | β    | 95% CI          | P     |
|-------------------------|------|-----------------|-------|
| Constant                | 99.999 | 84.500 – 115.498 | 0.000 |
| Age (year)              | –0.004 | –0.104 – 0.097 | 0.944 |
| Diabetes duration (year)| 0.103  | –0.229 – 0.435 | 0.541 |
| SBP (mmHg)              | 0.246  | 0.192 – 0.301 | 0.000 |
| DBP (mmHg)              | –0.137 | –0.248 – 0.027 | 0.016 |
| BMI (kg/m²)             | –0.577 | –0.788 – 0.366 | 0.000 |
| T-CHO (mmol/L)          | 0.612  | –1.283 – 2.507 | 0.524 |
| HDL-C (mmol/L)          | –3.388 | –7.490 – 0.714 | 0.104 |
| Triglyceride (mmol/L)   | –0.227 | –0.781 – 0.327 | 0.421 |
| Lipoprotein-a (mg/L)    | 0.001  | –0.002 – 0.005 | 0.397 |
| BUN (mmol/L)            | 0.149  | –0.114 – 0.411 | 0.267 |
| Scr (µmol/L)            | 0.031  | 0.023 – 0.039 | 0.000 |
| SUA (µmol/L)            | –0.005 | –0.013 – 0.004 | 0.300 |
| FBG (mmol/L)            | 0.098  | –0.073 – 0.268 | 0.259 |
| TT (s)                  | 0.131  | –0.199 – 0.462 | 0.425 |
| FDP (mg/L)              | 0.022  | –0.067 – 0.112 | 0.618 |
| D dimer (µg/L)          | 0.001  | –0.001 – 0.002 | 0.556 |
| Drinking                | –0.892 | –3.219 – 1.435 | 0.449 |

Bold results are statistically significant at the P < 0.05 level.

T2DM, type 2 diabetes mellitus; HTN, hypertension; SBP, systolic blood pressure; DBP, diastolic blood pressure; T-CHO, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; Scr, serum creatinine; SUA, serum uric acid; FBG, fasting blood-glucose; HbA1c, glycosylated haemoglobin; PT, prothrombin time; APTT, activated partial thromboplastin time; TT, thrombin time; INR, international normalized ratio; ATIII, Antithrombin III; FDP, fibrinogen degradation product.
criteria, there are still many unknown confounders and co-
morbidities that may not have been followed up, and it is
possible that future stratification could be further refined
by including other clustering variables for cardiac ultrasound.

Conflict of interest
The authors declare that they have no competing interests.

Funding
This work was supported by the Natural Science Foundation
of China (No. 81870568, Shaohua Wang).

Acknowledgements
We would like to express our heartfelt thanks to all the par-
ticipants in this study.

References
1. Li Y, Teng D, Shi X, Qin G, Qin Y, Quan H, et al. Prevalence of diabetes recorded
in mainland China using 2018 diagnostic
criteria from the American Diabetes As-
sociation: national cross sectional study.
Bmj. 2020; 369: m997.
2. Colosia AD, Palencia R, Khan S. Preva-
lence of hypertension and obesity in pa-
tients with type 2 diabetes mellitus in
observational studies: a systematic liter-
ature review. Diabetes Metab Syndr Obes.
2013; 6: 327–338.
3. Li Y, Teng D, Shi X, Teng X, Teng W,
Shan Z, et al. Changes in the prevalence
of obesity and hypertension and demo-
graphic risk factor profiles in China over
10 years: two national cross-sectional
surveys. Lancet Reg Health West Pac.
2021; 15: 100227.
4. Cameli M, Lembo M, Sciaccaluga C,
Bandera F, Ciccone MM, D’Andrea A,
et al. Identification of cardiac organ
damage in arterial hypertension: in-
sights by echocardiography for a com-
prehensive assessment. J Hypertens.
2020; 38: 588–598.
5. Wong CY, O’Moore-Sullivan T, Leano R,
Byrne N, Beller E, Marwick TH. Alterations
of left ventricular myocardial charac-
teristics associated with obesity. Circulation.
2004; 110: 3081–3087.
6. Gozdzik A, Marwick TH, Przewlocka-
Kosmala M, Jankowska EA, Ponikowski
P, Kosmala W. Comparison of left ven-
tricular longitudinal systolic function pa-
rameters in the prediction of adverse
outcome in heart failure with preserved
ejection fraction. ESC Heart Fail. 2021;
8: 1531–1540.
7. Cheng S, Xanthakis V, Sullivan LM, Lieb
W, Massaro J, Aragam J, et al. Correlates
of echocardiographic indices of cardiac
remodeling over the adult life course:
longitudinal observations from the Fra-
mingham Heart Study. Circulation.
2010; 122: 570–578.
8. Howden EJ, Carrick-Ranson G, Sarma S,
Hieda M, Fujimoto N, Levine BD. Effects
of Sedentary Aging and Lifelong Exer-
cise on Left Ventricular Systolic Func-
tion. Med Sci Sports Exerc. 2018; 50:
494–501.
9. Volzke H, Schmidt CO, Baumeister SE,
Ittermann T, Fung G, Krafczyk-Korth J,
et al. Personalized cardiovascular medi-
cine: concepts and methodological con-
10. Ernande L, Audureau E, Jellis CL, Bergerot C, Henegar C, Sawaki D, et al. Clinical Implications of Echocardiographic Phenotypes of Patients With Diabetes Mellitus. J Am Coll Cardiol. 2017; 70: 1704–1716.
11. Lancaster MC, Salem Omar AM, Narula S, Kulkarni H, Narula J, Sengupta PP. Phenotypic Clustering of Left Ventricular Diastolic Function Parameters: Patterns and Prognostic Relevance. JACC Cardiovasc Imaging. 2019; 12: 1149–1161.
12. Whelton PK, Carey RM, Aronow WS, Casey DE Jr, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/ABC/ACP/AGS/AHA/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2018; 138: e426–e483.
13. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension. J Hypertens. 2018; 36: 1953–2041.
14. Consultation WHOE. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004; 363: 157–163.
15. Gao M, Wei YX, Lyu J, Yu CQ, Guo Y, Bian Z, et al. The cut-off points of body mass index and waist circumference for predicting metabolic risk factors in Chinese adults. Zhonghua Liu Xing Bing Xue Za Zhi. 2019; 40: 1533–1540.
16. Naguse SF, Smiseth OA, Appleton CP, Byrd BF 3rd, Dokainish H, Edvardsen T, et al. Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr. 2016; 29: 277–314.
17. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med. 2006; 145: 247–254.
18. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. Ann Intern Med. 1999; 130: 461–470.
19. Devereux RB, Reichek N. Echocardiographic determination of left ventricular mass in man. Anatomic validation of the method. Circulation. 1977; 55: 613–618.
20. Iwashima Y, Horio T, Kuroda S, Takishita S, Kawano Y. Influence of plasma aldosterone on left ventricular geometry and diastolic function in treated essential hypertension. Hypertens Res. 2002; 25: 49–56.
21. Tan Y, Zhang Z, Zheng C, Wintergerst KA, Keller BB, Cai L. Mechanisms of diabetic cardiomyopathy and potential therapeutic strategies: preclinical and clinical evidence. Nat Rev Cardiol. 2020; 17: 585–607.
22. Jia G, DeMarco VG, Sowers JR. Insulin resistance and hyperinsulinaemia in diabetic cardiomyopathy. Nat Rev Endocrinol. 2016; 12: 144–153.
23. Wiwie C, Baumbach J, Rottger R. Comparing the performance of biomedical clustering methods. Nat Methods. 2015; 12: 1033–1038.
24. Karwath A, Bunting KV, Gill SK, Tica O, Pendleton S, Aziz F, et al. Redefining beta-blocker response in heart failure patients with sinus rhythm and atrial fibrillation: a machine learning cluster analysis. Lancet. 2021.
25. Ahmad T, Lund LH, Rao P, Ghosh R, Warier P, Vaccaro B, et al. Machine Learning Methods Improve Prognostication, Identify Clinically Distinct Phenotypes, and Detect Heterogeneity in Response to Therapy in a Large Cohort of Heart Failure Patients. J Am Heart Assoc. 2018; 7.
26. Ahlqvist E, Storm P, Karajamaki A, Martinell M, Dorkhan M, Carlsson A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol. 2018; 6: 361–369.
27. Falcao-Pires I, Leite-Moreira AF. Diabetic cardiomyopathy: understanding the molecular and cellular basis to progress in diagnosis and treatment. Heart Fail Rev. 2012; 17: 325–344.
28. Adeghate E, Singh J. Structural changes in the myocardium during diabetes-induced cardiomyopathy. Heart Fail Rev. 2014; 19: 15–23.
29. Packer M. Longevity genes, cardiac aging, and the pathogenesis of cardiomyopathy: implications for understanding the effects of current and future treatments for heart failure. Eur Heart J. 2020; 41: 3856–3861.
30. Maisch B, Alter P, Pankuweit S. Diabetic cardiomyopathy—fact or fiction? Herz. 2011; 36: 102–115.
31. Jia G, Whaley-Connell A, Sowers JR. Diabetic cardiomyopathy: a hyperglycaemia- and insulin-resistance-induced heart disease. Diabetologia. 2018; 61: 21–28.
32. Samson WK, Yosten GLC, Remme CA. Obesity and Cardiomyopathy. Physiol Rev. 2021.
33. Ng ACT, Delgado V, Borlaug BA, Bax JJ. Diabesity: the combined burden of obesity and diabetes on heart disease and the role of imaging. Nat Rev Cardiol. 2021; 18: 291–304.
34. Yang Q, Vijayakumar A, Kahn BB. Metabolites as regulators of insulin sensitivity and metabolism. Nat Rev Mol Cell Biol. 2018; 19: 654–672.