

RESEARCH

Categorization of β -cell capacity in patients with obesity via OGTT using K-means clustering

Li Li¹, Qifa Song² and Xi Yang¹¹Department of Endocrinology and Metabolism, Ningbo First Hospital, Ningbo, Zhejiang, China²Department of Microbiology, Ningbo Municipal Centre for Disease Control and Prevention, Ningbo, Zhejiang, ChinaCorrespondence should be addressed to Q Song: qifasong@126.com

Abstract

Insufficient insulin release plays a crucial role in the development of unhealthy status in patients with obesity; the present study aimed to classify these patients by the indices for insulin resistance and insulin release. After the indices from OGTT were assessed to achieve high differentiability and low redundancy in classifying patients, HOMA-IR and IGI_{30min} were chosen to classify the patients using K-means clustering method. A total of 249 non-diabetic patients with obesity were classified into four groups. In Group 1, 19 patients were characteristic of high insulin resistance and high insulin release, as well as well-controlled glucose levels, the highest BMI, the youngest age, and the highest early phase release of insulin. In Group 2, 38 patients were unhealthiest in terms of high insulin resistance, reduced insulin release and IGT status. Group 3 consisted of 63 patients that were healthiest with low insulin resistance and high insulin release. In Group 4, 46 IGT patients and 14 IFG patients were identified among 129 patients that showed low insulin resistance, low insulin release, moderate obesity and older age. These concurrent impotent insulin release, older age, and moderate obesity indicated decreasing obesity with increasing age and reduced insulin release. The classification of patients with obesity using K-means clustering method by HOMA-IR and IGI_{30min} provides more information about the development of obesity and unhealthy status. The patients with distinct insulin resistance and insulin release should be followed up, especially for those with reduced or even absent insulin response to glucose stimulation.

Key Words

- ▶ patients with obesity
- ▶ oral glucose tolerance test
- ▶ insulin resistance
- ▶ pancreatic β -cell function
- ▶ K-means clustering

Endocrine Connections
(2020) **9**, 135–143

Introduction

In the past 40 years, a dramatic increase in the overweight and obese population has been seen globally. The potential adverse effects of obesity include insulin resistance and possible insufficient insulin secretion due to the exhaustion of pancreatic β -cells (1). This deficiency in insulin function is the primary cause for developing a metabolically unhealthy status in patients with obesity who usually display impaired glucose tolerance (IGT) and even type 2 diabetic mellitus (T2DM) (2). While obesity is a significant risk for T2DM, β -cells possess the compensatory capacity that prevents most patients with obesity and insulin resistance from developing T2DM (3). During the progression of obesity toward IGT and T2DM,

insulin resistance and insufficient insulin release due to the dysfunction of pancreatic β -cells are the key features prior to β -cell failure (4). So, accurately evaluating the β -cell compensatory capacity in patients with obesity is essential for understanding the progression of unhealthy state. However, how to score the β -cell compensatory capacity in patients with obesity is inconsistent, for example, LDL cholesterol levels were used as predictors of transition from metabolically healthy obesity to unhealthy obesity (5). A previous study illustrated that the trajectories of change in glucose tolerance in obese youths were highly reversible and early identification could provide important information for intervention (6).

The ‘gold standard’ for quantifying the insulin secretion by β-cells after glucose stimulation is hyperglycemic and euglycemic hyperinsulinemic glucose clamp. This method measures how well an individual metabolizes glucose and how sensitive an individual responds to insulin (7). Nevertheless, the time-consuming and labor-intensive feature of this technique limits its routine use. Instead, based on the indices from oral glucose tolerance test (OGTT), several mathematical models have been developed to generate indices that correlate well with the “gold standard” results in the assessment of insulin resistance and β-cell function (8). These indices can be largely divided into two types. One type of indices refers to insulin resistance, such as insulin resistance index (HOMA-IR) (9) and composite (Matsuda) insulin sensitivity index (ISI) (10). The other type refers to insulin release, such as homoeostasis model assessment of β-cell function (HOMA-β) (9), insulinogenic index (IGI_{30min}) (11), first (1st PH) and second phase (2nd PH) insulin secretion levels (12), and glucose disposition index (DI) (13). These indices are routinely accessible and are more applicable to the assessment of β-cell compensatory capacity among a large number of patients with obesity.

Insulin resistance and β-cell function or insulin release were individually evaluated for patients with obesity according to predetermined reference values (5). This type of traditional classification method belongs to a supervised technique and has the risk of missing naturally occurring disease patterns or subgroups (14). In contrast, K-means clustering method, a type of big data analysis techniques, applies unsupervised clustering algorithms to classify patients according to the distribution of naturally existing values of one variable or a combination of multiple variables (15). It has the advantage to uncover naturally occurring disease patterns or subgroups and is useful in precision medicine and epidemiological research. Recently, several researches have been conducted to diagnose pre-diabetes and T2DM, as well as classify patients with diabetes mellitus using big data and machine learning techniques (16).

Since insulin resistance and insufficient insulin release considerably account for the deficiency in insulin function and thereby IGT and T2DM, we hypothesized that the K-means clustering with a combination of variables representing insulin resistance and insulin release could be applied to categorize the pancreatic β-cell compensatory capacity of patients with obesity. This new classification way would generate subgroups of the patients with distinctive features regarding β-cell compensatory capacity and glucose metabolism.

In this study, we aimed to apply K-means clustering to classify Chinese patients with overweight and obesity using a combination of indices from OGTT. The study included the selection of appropriate variables and the comparison between different combinations of variables to achieve better and clinically more meaningful differentiation.

Materials and methods

Study design and study population

This study was approved by the Ethics Committee of Ningbo First Hospital and followed the Declaration of Helsinki. Written consent was acquired from all patients.

The study design included patient enrollment, variable selection for classifying patients, classification of patients, and result interpretation. As a lot of variables were related to insulin resistance and insulin release, the variables should be assessed to remove redundancy and obtain high differentiation of patients. Then, we compared the primary attributes among sub-groups of patients and clinically interpreted the results. The study enrolled patients with age of 18–75 years and body mass index (BMI) ≥ 24 kg/m² who sought weight loss treatment in the hospital from July 2015 to December 2018. Their medical history was reviewed by the medical staff. Exclusion criteria were diabetic mellitus and clinically diagnosed secondary obesity from underlining conditions including Cushing's syndrome, polycystic ovary syndrome, and insulin-producing tumors. Data from the patients were retrospectively analyzed.

Laboratory measurements

The patient physical examinations, including measurements of weight, height, waist circumference, hip circumference, and systolic and diastolic blood pressure, were carried out by trained medical staff. Patients were given a 2-h 75-g OGTT after 10–12 h overnight fasting. Concentrations of plasma glucose and insulin were obtained at 0, 30, and 120 min after ingestion of 75 g glucose. Fasting plasma glucose (FPG), insulin, and glycosylated hemoglobin (HbA1c) were measured after 12-h fasting. OGTT results were interpreted according to American Diabetes Association Criteria (2019) (17) as follows: normal glucose tolerance (NGT) was defined as FPG < 5.6 mmol/L and 2-h postprandial blood glucose (PBG_{2h}) < 7.8 mmol/L; IGT was defined as PBG_{2h} 7.8–11.0 mmol/L; impaired fasting glucose (IFG) was

Table 1 Calculated indices from OGTT results.

Usage	Index name	Equation	Units (glucose; insulin)	Reference
Insulin resistance	HOMA-IR	$FPG \times \text{fasting insulin} / 22.5$	mmol/L; mIU/L	(9)
	ISI	$10,000 / \sqrt{FBG \times \text{fasting insulin} \times \text{mean glucose} \times \text{mean insulin}}$	mg/dL; mIU/L	(10)
Insulin release (β-cell function)	HOMA-β	$20 \times \text{fasting insulin} / (FBG - 3.5)$	mmol/L; mIU/L	(9)
	IGI _{30min}	$(\text{Insulin}_{30min} - \text{insulin}_{0min}) / (\text{glucose}_{30min} - \text{glucose}_{0min})$	mmol/L; mIU/L	(11)
	1st PH	$1283 + 1.829 \times \text{Insulin}_{30m} - 138.7 \times \text{Glucose}_{30m} + 3.772 \times \text{Insulin}_{0m}$	mmol/L; pmol/L	(12)
Disposition index (DI)	2nd PH	$393 + 1.163 \times \text{Insulin}_{0m} - 40.72 \times \text{Glucose}_{120m} + 0.313 \times \text{Insulin}_{120m}$	mmol/L; pmol/L	(12)
		IGI ₃₀ /fasting insulin	mmol/L; mIU/L	(13)

defined as 5.6–7.0 mmol/L and PBG_{2h} <7.8 mmol/L; T2DM was defined as FPG >7 mmol/L and/or PBG_{2h} >11.0 mmol/L.

Selection of attributes representing insulin resistance and insulin release

A combination of one index for insulin resistance and one index for insulin release was used to classify the patients with obesity. The index representing insulin resistance was chosen from one of HOMA-IR and composite ISI. Another index representing insulin release and β-cell function was chosen from one of HOMA-β, IGI_{30min}, 1st PH, and 2nd PH (Table 1). As indices that were highly correlated with each other led to over-representation of the relevant features, correlation matrix of each pair of indices was reviewed to remove redundant indices that were highly inter-correlated.

Classification of the overweight and obese patients

K-means clustering classifies objects through calculating the least distance of indexing values of objects in a group and yields a center-based point for the group. This new center-based point is professionally called centroid and represents the means of attribute values in a group (18). Here, the Euclidean distance of two indices respectively representing insulin resistance and insulin release were jointly calculated.

Patients were clustered into pre-specified number of groups according to the Euclidean least distance by K-means clustering method that was based on R software package factextra (version 1.0.5) (<http://www.sthda.com/english/rpkgs/factextra>). The optimal number of groups was determined by reviewing the sum of squared errors (SSE) plotted against different numbers of groups and the calculated Silhouette value. A lower SSE indicated a clearer separation between groups (19). The clustering results were interpreted by comparing the means of primary attributes among groups.

Statistical analysis

Continuous values were expressed as means and s.d. Pearson correlation and parameters of each pair of indices were obtained via corr () function in R software.

Results

During the study period, 283 patients (178 women and 105 men) with BMI >24 kg/m² were enrolled. Except for 34 patients with >7.0 mmol/L FBG, the remaining 249 (161 women and 88 men) patients with <7.0 mmol/L FBG were given OGTT test.

OGTT results showed that the 249 patients consisted of 12 T2DM patients with PBG_{2h} >11 mmol/L, 99 IGT or IFG patients, and 138 NGT patients (Table 2). Of the variable values, IGI_{30min} was most distinct between NGT patients, IGT and IFG patients, and T2DM patients.

The correlation matrix was plotted to specify the variable that represented insulin resistance and the variable that represented β-cell function or insulin release (Fig. 1). The coefficient r for correlation between HOMA-IR and HOMA-β was 0.46 (95% CI: 0.35–0.60), indicating HOMA-IR and HOMA-β were highly correlated and inappropriate to be jointly used as indices for clustering. In contrast, the coefficient r for correlation between HOMA-IR and IGI_{30min} showed a much lower value of 0.04, indicating less redundancy with each other. HOMA-IR was also highly correlated to ISI, 1st PH, and 2nd PH. As mentioned above, the values of IGI_{30min} were most contrasting among NGT patients, IGT and IFG patients, and T2DM patients. Theoretically, HOMA-IR and IGI_{30min} could achieve much higher differentiability and lower redundancy when they were used as clustering indices. Therefore, HOMA-IR and IGI_{30min} were selected as clustering indices that respectively represented insulin resistance and insulin release.

The optimal number of subgroups was specified to be four after the sum of square error within groups was plotted against different numbers of subgroups.

Table 2 Characteristics of NGT, IGT, IFG, and T2DM patients regarding obesity levels, glucose metabolism, insulin resistance and β-cell function.

Categories	Variables	Mean ± s.d.				
		NGT	IGT	IFG	T2DM	Total
Number of patients		138	62	37	12	249
Physical examination	Age	27.9 ± 8.2	33.9 ± 10.6	31.2 ± 7.2	40.6 ± 13.9	30.5 ± 9.6
	Weight	91.6 ± 16.5	92.6 ± 20.0	101.0 ± 22.0	95.1 ± 18.8	93.3 ± 18.5
	BMI	32.8 ± 4.4	33.5 ± 4.8	35.6 ± 5.9	36.2 ± 6.5	33.5 ± 4.9
	WHR	0.92 ± 0.07	0.95 ± 0.07	0.95 ± 0.07	0.95 ± 0.04	0.93 ± 0.07
	SBP	129 ± 12	135 ± 15	132 ± 14	137 ± 12	131 ± 134
Laboratory measurement	HbA1C	5.3 ± 0.38	5.8 ± 0.5	5.6 ± 0.4	6.0 ± 0.5	5.5 ± 0.5
	FBG	5.1 ± 0.3	5.6 ± 0.7	5.9 ± 0.3	6.1 ± 0.7	5.4 ± 0.6
	PBG _{30min}	8.5 ± 1.2	9.6 ± 1.5	9.4 ± 1.8	10.9 ± 2.0	9.1 ± 1.6
	PBG _{2h}	6.0 ± 1.0	8.9 ± 0.8	6.3 ± 1.0	12.2 ± 1.3	7.1 ± 2.0
	FINS	25 ± 13	30 ± 22	32 ± 20	31 ± 16	27 ± 17
	INS _{30min}	185 ± 140	134 ± 99	159 ± 101	112 ± 93	164 ± 125
	INS _{2h}	97 ± 104	170 ± 108	97 ± 76	188 ± 95	120 ± 106
Insulin resistance	HOMA-IR	5.5 ± 3.0	7.6 ± 5.7	8.6 ± 5.3	8.6 ± 5.1	6.7 ± 4.4
	ISI	3.7 ± 1.8	2.2 ± 1.1	2.7 ± 1.5	1.7 ± 0.9	3.0 ± 1.7
Insulin release (β-cell function)	HOMA-β	327 ± 226	303 ± 220	273 ± 176	236 ± 93	307 ± 213
	IGI _{30min}	48.6 ± 45.8	24.6 ± 32.4	39.4 ± 25.6	15.3 ± 10.6	39.3 ± 40.4
	1st PH	3105 ± 1974	2442 ± 1725	2852 ± 1647	1999 ± 1331	2832 ± 1855
	2nd PH	560 ± 290	641 ± 390	609 ± 263	550 ± 313	587 ± 314
Disposition index (DI)		2.2 ± 1.9	0.9 ± 1.4	1.4 ± 1.0	0.5 ± 0.3	1.7 ± 1.7

Glucose is expressed in mmol/L; insulin is expressed in mIU/L.

The curve became slowly down at four groups, indicating a lower sum of square error within groups and thereby good separation quality (Supplementary Fig. 1, see section on [supplementary materials](#) given at the end of this article). This good quality was also proved by the highest Silhouette value that was 0.43 when patients were divided into four groups (Fig. 2). Patients in Group 1 showed both high insulin resistance and high insulin release, showing a potent β-cell function (Table 3).

The main features of Group 1 consisted of well-controlled glucose levels, the highest BMI, the youngest patients, and the highest early phase secretion of insulin. Patients in Group 2 showed high insulin resistance and low insulin secretion, indicating an impotent β-cell function and the unhealthiest status in terms of reduced insulin release and obvious IGT status. Group 3 was the healthiest with low insulin resistance and high insulin release. Patients in Group 4 showed both low insulin

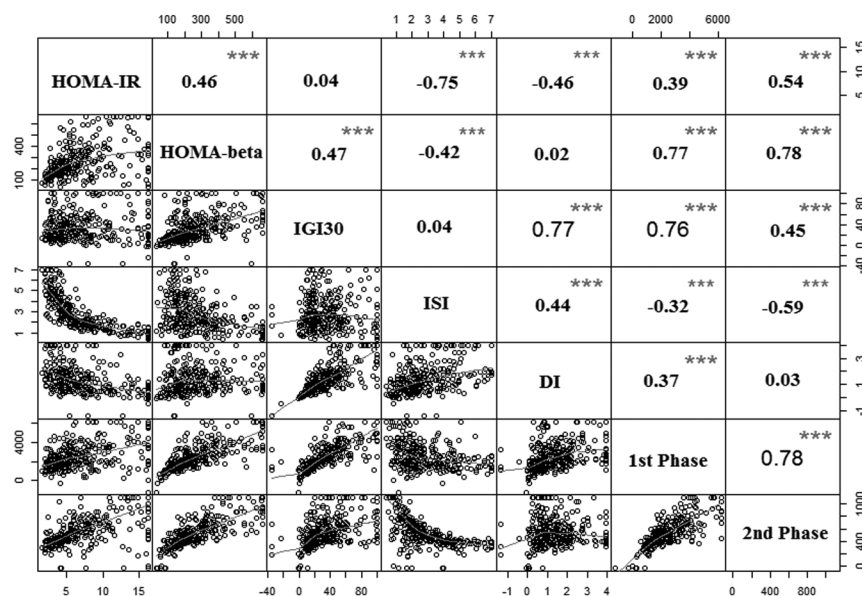


Figure 1 Correlation matrix of seven attributes representing insulin resistance and insulin release. IGI_{30min} that represents insulin release has low-level correlation with HOMA-IR ($r = 0.04$) and ISI ($r = 0.04$).

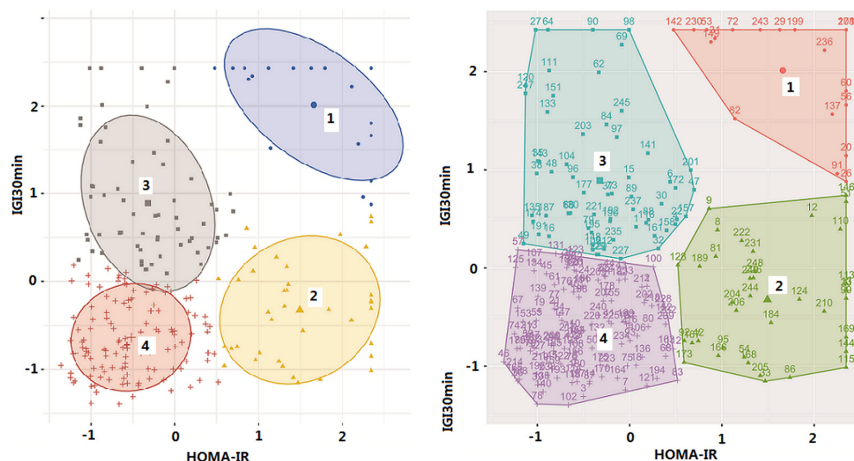


Figure 2 Four groups generated according to HOMA-IR and IGI_{30min} by the K-means clustering method. Left: The larger point at the center of each circle stands for the centroid whose means of variables are listed in Table 3. Right: The distribution of encircled patients.

resistance and low insulin release, which was contrary to Group 1 in insulin resistance and insulin release. As to the composition of patients in Group 4, 46 IGT patients and 14 IFG patients were identified when 129 patients were further classified according to glucose metabolism (see In Group 4 in Table 3). These patients were characteristic of impotent insulin release, older age, and moderate obesity. These features were also contrary to features of Group 1.

When the patients were more finely grouped into six groups, Group 1 in the six-group classification was similar to Group 1 in the four-group classification in terms of the features of groups (Supplementary Fig. 2 and Supplementary Table 1) mentioned above. Both Group 2 and Group 3 were at typically unhealthy state as they showed high insulin resistance and low insulin secretion, with Group 2 severer than Group 3. Group 4, Group 5, and Group 6 showed low or normal insulin resistance and three distinct levels of insulin release. Although the patients in Group 6 were not insulin-resistant, their β-cell function was worst as improved by the lowest insulin release values (mean of $IGI_{30min}=16.5$). This condition led to a typical IGT (mean of $PBG_{30min}=9.2$ mmol/L) despite the absence of insulin resistance.

Discussion

The present study applied K-means clustering method with a combination of HOMA-IR and IGI_{30min} to classify 249 patients with obesity. The obtained subgroups of the patients showed distinct features in glucose metabolism, BMI, age, insulin resistance and insulin release. The features could provide information for precision medicine.

The classification of ever-increasing patients with obesity is a prerequisite for the precision intervention of

this condition (20). Clinically, the primary unhealthy state related to obesity is the development of IGT and T2DM (21). In patients with insulin resistance, hyperglycemia usually arises when β cells can not timely secrete sufficient insulin in response to various stimuli such as glucose (21). For that reason, an accurate description of β-cell function is critical to tell the metabolic status of patients with obesity. Nonetheless, a lack of widely accepted criteria for such classification led to inconsistent results. Previous studies have categorized the patients into metabolically healthy and unhealthy subgroups by sets of arbitrarily selected clinical features (22, 23), for instance, a batch of FBG, systolic blood pressure, and low-density lipoprotein cholesterol (2). Regarding the indices in classification, the previous researches failed to include both indices of insulin resistance and insulin release as a whole to explore the natural subgroups of such patients.

As to the classification method, the traditional method usually made use of a supervised classification method that involved pre-specified features (24). For example, the patients were classified into NGT, IGT, and T2DM subgroups according to glucose reference ranges in FBG and PBG_{2h} . However, two deficits were present in the traditional classification method. One limitation was due to the supervision feature that might fail to uncover dominant naturally occurring subgroups. The other was the failure to include combined indices for insulin resistance and insulin release concurrently. The present study took use of unsupervised K-means clustering. Its unsupervised traits resulted in the absence of pre-specified reference ranges of grouping variables and had the potentiality to uncover the naturally occurring patterns of numerous clinical features (25). Concerning the indices for insulin resistance and insulin release, the hyperglycemic and euglycemic hyperinsulinemic glucose

Table 3 Characteristics of patients in four groups regarding obesity levels, glucose metabolism, insulin resistance and β-cell function.

Categories	Variables	Mean ± s.d.			
		Group 1	Group 2	Group 3	Group 4
Number of objects		19	38	63	129
Physical examinations	Age	24.7 ± 4.6	31.4 ± 9.1	28.4 ± 8.8	32.2 ± 10.3
	Weight	114.2 ± 19.2	102 ± 18.5	91.1 ± 16.1	88.1 ± 15.9
	BMI	39.3 ± 6.1	35.6 ± 5.3	32.7 ± 4.5	32.3 ± 3.8
	WHR	0.95 ± 0.06	0.96 ± 0.06	0.93 ± 0.07	0.92 ± 0.07
	SBP	134 ± 12	138 ± 15	127 ± 14	131 ± 12
Laboratory measurements	HbA1C	5.4 ± 0.4	5.9 ± 0.6	5.4 ± 0.4	5.5 ± 0.4
	FBG	5.3 ± 0.6	6.0 ± 0.6	5.1 ± 0.4	5.4 ± 0.5
	PBG _{30min}	8.5 ± 1.1	10.7 ± 1.9	8.1 ± 1.2	9.1 ± 1.3
	PBG _{2h}	6.5 ± 1.7	8.2 ± 2.4	6.0 ± 1.1	7.3 ± 2.0
	FINS	61 ± 30	45 ± 12	23 ± 7	19 ± 6
	INS _{30min}	423 ± 233	183 ± 78	206 ± 86	102 ± 48
	INS _{2h}	257 ± 237	171 ± 97	105 ± 81	92 ± 60
	HOMA-IR	12.0 ± 2.3	11.4 ± 2.1	5.3 ± 1.8	4.6 ± 1.6
Insulin resistance	ISI	1.4 ± 0.6	1.4 ± 0.5	3.4 ± 1.5	3.6 ± 1.5
	HOMA-β	555 ± 126	388 ± 152	296 ± 119	222 ± 106
	IGI _{30min}	90 ± 14.2	30.3 ± 14.0	61.4 ± 16.8	21.4 ± 10.4
Insulin release (β-cell function)	1st phase	5354 ± 792	3321 ± 1207	3340 ± 920	1817 ± 657
	2nd phase	932 ± 181	784 ± 185	564 ± 171	452 ± 143
Disposition index (DI)		2.0 ± 1	0.7 ± 0.3	2.7 ± 1	1.3 ± 0.7
	Interpretation	High insulin resistance, high insulin release	High insulin resistance, low insulin release	Low insulin resistance, high insulin release	Low insulin resistance, low insulin release

Glucose is expressed in mmol/L; insulin is expressed in mIU/L.

clamp is not suitable for a large number of patients with obesity. Therefore, several calculated indices such as ISI_{30min} , 1st PH, 2nd PH, and HOMA-IR from OGTT are more practicable (7). In this study, after checking for redundancy, differentiability and representativeness, as well as taking into account the professional knowledge, we used a combination of HOMA-IR for insulin resistance and IGI_{30min} for β-cell function to classify the patients. Additionally, a professional interpretation of obtained groups was performed.

In the present study, the patients were first classified into NGT, IFG and IGT, as well as T2DM according to pre-specified FBG and PBG_{2h} levels (Table 2). This commonly used classification provided information about glucose metabolism outcomes. When the patients were divided into four groups according to insulin resistance represented by HOMA-IR and β-cell function represented by IGI_{30min} , more information could be obtained about the progression of obesity, glucose metabolism, as well as other risk factors such as age (Table 3). Group 1 was characteristic of well-controlled FBG and PBG_{2h} , young patients, high insulin resistance, and dramatically higher insulin release, as illustrated by INS_{30min} . Group 2 consisted of patients with the most abnormal PBG_{2h} . This finding could be explained by the high insulin resistance and low insulin release. In Group 3, because of insulin sensitivity and high insulin release, the patients were the healthiest in terms of low obesity indices, and normal glucose metabolism. As to Group 4, the interpretation was complex. We could observe the primary features of this group, i.e. both low insulin resistance and low insulin secretion, which was contrary to the corresponding features of Group 1. After further classifying the patients in Group 4 according to the levels of FBG and PBG_{2h} , we identified 46 IGT patients and 14 IFG patients. These patients were characteristic of older age, moderate obesity, and impotent insulin secretion (see In Group 4 in Table 3). These results suggested more attention be required for this subgroup of patients at the treatment of weight loss because they were not clinically healthy due to their obviously failed β-cell response to glucose stimulation despite mild obesity. As the remaining patients that showed NGT in Group 4, more identification assays were needed to make out whether these patients were of low insulin release because of β-cell impotence or slow digestive uptake of glucose.

Regarding obesity development, the highest weight and BMI values were observed in Group 1. The lowest weight and BMI values were observed in Group 4, whose patients were almost eldest and of lowest insulin release. It was reasonable to hypothesize that young age and high

insulin release are contributors to obesity. In this study, we could also observe an obvious relationship between increasing age and decreasing insulin release illustrated by IGI_{30min} . In consideration of the young age and the highest insulin release in the patients in Group 1, these patients had a short duration of obesity and were healthy in terms of glucose metabolism. How they would elapse into IGT and even T2DM and the benefit from the precision intervention were worthy being traced. All of the above conclusions were important for the assessment of weight loss treatment.

After increasing the number of classification groups to six, the conclusions mentioned above, including the reduced insulin release causing IGT and T2DM, as well as young age and high insulin release both related to high BMI, still existed (Supplementary Fig. 2 and Supplementary Table 1). These results indicated the stability of the clustering method. The six-group method generated more finely categorized patients whose insulin resistance and insulin release levels were divided into high and moderate subgroups, as compared with Group 2 and Group 3 in the four-group method (Table 3).

The present unsupervised K-means classification method reflected the overall effects of insulin resistance and insulin release on glucose metabolism. After oral administration of glucose, β-cells increase insulin secretion in response to increased glucose concentrations (26). Concerning that obesity is generally characterized by insulin resistance and hyperinsulinemia (27, 28), this reduced level of serum insulin may be indicative of a severer pancreatic state when β-cells have gradually lost the compensative capacity to meet a metabolic need (29). It is also noticeable that the final effect of insulin on glucose metabolism is affected by many other factors such as digestive absorption of glucose, regulation by gluconeogenesis in liver and kidney, transportation of glucose into muscle and adipose cells. Therefore, it is advocated to consider the implication of the clustering results as a way to provide clues in understanding the mechanism of unhealthy status and the overall trend in obesity development. As for an individual patient, more risk factors should be considered to clarify the exact illness state.

The present study had two major limitations. One was the small sample size. However, the two indices that were chosen for classifying patients varied dramatically, ensuring a high and reliable differentiation. The primary attributes of patients in different sub-groups also showed a clear separation. If indices varied among patients at a less level, a careful assessment of sample size should be

performed to ensure an enough differentiation of patients. Another limitation was the overlapping proportion of patients between sub-groups. This issue is often met in most classification methods, including categorizing patients by reference values. Therefore, an interpretation of results by clinical knowledge and more confirmative tests should be taken.

To summarize, through obesity via K-means clustering method with HOMA-IR and IGI_{30min} provides, we have successfully classified the patients with obesity into groups with distinctive features of β-cell function and glucose metabolism. Combined with different classification methods, we conclude that young age and the high response of insulin release are contributors to obesity, and a large proportion of patients with obesity show reduced pancreatic β-cell function despite low insulin resistance, which is causative to glucose dysmetabolism. In the future, the patients with distinct insulin resistance and insulin release should be followed up, especially for the patients with reduced or even absent insulin response to glucose stimulation. These results demonstrate that classification by the calculated indices from OGTT can provide more information about the development of obesity.

Supplementary materials

This is linked to the online version of the paper at <https://doi.org/10.1530/EC-19-0476>.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This study was supported by the Natural Science Foundation of Ningbo, China (Grant No. 2016A610169) and the Commonwealth Research Project on Application Technology in Zhejiang Province, China (Grant No. 2017C35006).

Data availability

All data used in this study can be obtained from the corresponding author via email address.

References

- Kahn SE, Hull RL & Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* 2006 **444** 840–846. (<https://doi.org/10.1038/nature05482>)
- Iacobini C, Pugliese G, Blasetti Fantauzzi C, Federici M & Menini S. Metabolically healthy versus metabolically unhealthy obesity. *Metabolism: Clinical and Experimental* 2019 **92** 51–60. (<https://doi.org/10.1016/j.metabol.2018.11.009>)
- Meigs JB, Wilson PW, Fox CS, Vasan RS, Nathan DM, Sullivan LM & D'Agostino RB. Body mass index, metabolic syndrome, and risk of type 2 diabetes or cardiovascular disease. *Journal of Clinical Endocrinology and Metabolism* 2006 **91** 2906–2912. (<https://doi.org/10.1210/jc.2006-0594>)
- Kahn SE. The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of type 2 diabetes. *Diabetologia* 2003 **46** 3–19. (<https://doi.org/10.1007/s00125-002-1009-0>)
- Gilardini L, Zambon A, Soranna D, Croci M & Invitti C. Predictors of the transition from metabolically healthy obesity to unhealthy obesity. *Eating and Weight Disorders* 2018 **23** 739–744. (<https://doi.org/10.1007/s40519-018-0600-4>)
- Galderisi A, Giannini C, Weiss R, Kim G, Shabanova V, Santoro N, Pierpont B, Savoye M & Caprio S. Trajectories of changes in glucose tolerance in a multiethnic cohort of obese youths: an observational prospective analysis. *Lancet: Child and Adolescent Health* 2018 **2** 726–735. ([https://doi.org/10.1016/S2352-4642\(18\)30235-9](https://doi.org/10.1016/S2352-4642(18)30235-9))
- Kuroda A, Taniguchi S, Akehi Y, Mori H, Tamaki M, Suzuki R, Otsuka Y & Matsuhisa M. Accuracy and time delay of glucose measurements of continuous glucose monitoring and bedside artificial pancreas during hyperglycemic and euglycemic hyperinsulinemic glucose clamp study. *Journal of Diabetes Science and Technology* 2017 **11** 1096–1100. (<https://doi.org/10.1177/1932296817735122>)
- World Health Organization. *WHO Expert Committee on Diabetes Mellitus. Second Report* (Technical Report Series no. 646-1). Geneva, Switzerland: WHO, 1980.
- Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF & Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 1985 **28** 412–419. (<https://doi.org/10.1007/bf00280883>)
- Matsuda M & DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. *Diabetes Care* 1999 **22** 1462–1470. (<https://doi.org/10.2337/diacare.22.9.1462>)
- Tura A, Kautzky-Willer A & Pacini G. Insulinogenic indices from insulin and C-peptide: comparison of beta-cell function from OGTT and IVGTT. *Diabetes Research and Clinical Practice* 2006 **72** 298–301. (<https://doi.org/10.1016/j.diabres.2005.10.005>)
- Stumvoll M, Mitrakou A, Pimenta W, Jenssen T, Yki-Järvinen H, Van Haeften T, Renn W & Gerich J. Use of the oral glucose tolerance test to assess insulin release and insulin sensitivity. *Diabetes Care* 2000 **23** 295–301. (<https://doi.org/10.2337/diacare.23.3.295>)
- Utzschneider KM, Prigeon RL, Faulenbach MV, Tong J, Carr DB, Boyko EJ, Leonetti DL, McNeely MJ, Fujimoto WY & Kahn SE. Oral disposition index predicts the development of future diabetes above and beyond fasting and 2-h glucose levels. *Diabetes Care* 2009 **32** 335–341. (<https://doi.org/10.2337/dc08-1478>)
- Deo RC. Machine learning in medicine. *Circulation* 2015 **132** 1920–1930. (<https://doi.org/10.1161/CIRCULATIONAHA.115.001593>)
- Li L, Song Q & Yang X. K-means clustering of overweight and obese population using quantile-transformed metabolic data. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 2019 **12** 1573–1582. (<https://doi.org/10.2147/DMSO.S206640>)
- Alamsyah M, Nafisah Z, Prayitno E, Afida AM & Imah EM. The classification of diabetes mellitus using Kernel k-means. *Journal of Physics: Conference Series* 2018 **947** 012003. (<https://doi.org/10.1088/1742-6596/947/1/012003>)
- American Diabetes Association. Classification and diagnosis of diabetes: standards of medical care in Diabetesd 2019. *Diabetes Care* 2019 **42** (Supplement 1) S13–S28. (<https://doi.org/10.2337/dc19-S002>)
- Tan P-N, Steinbach M, Karpatne A & Kumar V. *Introduction to Data Mining*, 2nd ed., Chapter 8. London, UK: Pearson, 2018.

- 19 Wilkin GA & Huang X. A practical comparison of two K-Means clustering algorithms. *BMC Bioinformatics* 2008 **9** (Supplement 6) S19. (<https://doi.org/10.1186/1471-2105-9-S6-S19>)
- 20 Frühbeck G, Kiortsis DN & Catalán V. Precision medicine: diagnosis and management of obesity. *Lancet Diabetes and Endocrinology* 2017 **6** 164–166. ([https://doi.org/10.1016/S2213-8587\(17\)30312-1](https://doi.org/10.1016/S2213-8587(17)30312-1))
- 21 Seino S, Shibasaki T & Minami K. Dynamics of insulin secretion and the clinical implications for obesity and diabetes. *Journal of Clinical Investigation* 2011 **121** 2118–2125. (<https://doi.org/10.1172/JCI45680>)
- 22 De LA, Da CLC, Lessa R & Asb M. ‘Metabolically healthy’ obesity: fact or threat? *Current Diabetes Reviews* 2017 **13** 405–423.
- 23 Stefan N, Häring HU, Hu FB & Schulze MB. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet: Diabetes and Endocrinology* 2013 **1** 152–162. ([https://doi.org/10.1016/S2213-8587\(13\)70062-7](https://doi.org/10.1016/S2213-8587(13)70062-7))
- 24 Paoletti M, Camiciottoli G, Meoni E, Bigazzi F, Cestelli L, Pistolesi M & Marchesi C. Explorative data analysis techniques and unsupervised clustering methods to support clinical assessment of chronic obstructive pulmonary disease (COPD) phenotypes. *Journal of Biomedical Informatics* 2009 **42** 1013–1021. (<https://doi.org/10.1016/j.jbi.2009.05.008>)
- 25 Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiadu M, Bonow RO, Huang CC & Deo RC. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 2015 **131** 269–279. (<https://doi.org/10.1161/CIRCULATIONAHA.114.010637>)
- 26 Tillil H, Shapiro ET, Miller MA, Karrison T, Frank BH, Galloway JA, Rubenstein AH & Polonsky KS. Dose-dependent effects of oral and intravenous glucose on insulin secretion and clearance in normal humans. *American Journal of Physiology* 1988 **254** E349–E357. (<https://doi.org/10.1152/ajpendo.1988.254.3.E349>)
- 27 Polonsky KS, Given BD, Hirsch L, Shapiro ET, Tillil H, Beebe C, Galloway JA, Frank BH, Karrison T & Van Cauter E. Quantitative study of insulin secretion and clearance in normal and obese subjects. *Journal of Clinical Investigation* 1988 **81** 435–441. (<https://doi.org/10.1172/JCI113338>)
- 28 Jones CN, Abbasi F, Carantoni M, Polonsky KS & Reaven GM. Roles of insulin resistance and obesity in regulation of plasma insulin concentrations. *American Journal of Physiology: Endocrinology and Metabolism* 2000 **278** E501–E508. (<https://doi.org/10.1152/ajpendo.2000.278.3.E501>)
- 29 Porte Jr D. Clinical importance of insulin secretion and its interaction with insulin resistance in the treatment of type 2 diabetes mellitus and its complications. *Diabetes/Metabolism Research and Reviews* 2001 **17** 181–188. ([https://doi.org/10.1002/1520-7560\(200105/06\)17:3<181::aid-dmrr197>3.0.co;2-1](https://doi.org/10.1002/1520-7560(200105/06)17:3<181::aid-dmrr197>3.0.co;2-1))

Received in final form 23 December 2019

Accepted 7 January 2020

Accepted Manuscript published online 7 January 2020