International Journal of Diabetes in Developing Countries (January–March 2025) 45(1):47–54
https://doi.org/10.1007/s13410-024-01322-8

ORIGINAL ARTICLE

Cluster analysis based on fasting and postprandial plasma glucose and insulin concentrations

Miguel Altuve[1,2] · **Erika Severeyn**[1,2]

Received: 25 March 2023 / Accepted: 7 February 2024 / Published online: 7 March 2024

Abstract

Objective Plasma glucose and insulin concentrations are clinical markers used to diagnose metabolic diseases, particularly prediabetes and diabetes. In this paper, we conducted a cluster analysis using plasma glucose and insulin data collected during both fasting and 2-h postprandial periods.

Methods Different clustering experiments were performed by changing the attributes, from one (fasting glucose) to four (fasting and postprandial glucose and insulin) attributes input to a k-means clustering algorithm. Based on the elbow and silhouette methods, three clusters were chosen to perform the clustering experiments. The Pearson correlation coefficient was utilized to evaluate the association between the levels of glucose and insulin within each created cluster.

Results Results show that one cluster comprised individuals with prediabetes, another cluster consisted of individuals with diabetes, while subjects without prediabetes and diabetes were assigned to a separate cluster. Despite not being used as an attribute, we observed varying age ranges among subjects in the three clusters. Furthermore, significant correlations were found between fasting and postprandial insulin levels, as well as between fasting and postprandial glucose levels, suggesting a consistent relationship between these variables, and highlighting their interdependence in the context of glucose metabolism.

Conclusion The clustering analysis successfully differentiated individuals into distinct clusters based on their metabolic conditions, confirming that the approach effectively captured the underlying patterns in the plasma glucose and insulin data. Furthermore, despite not being a considered attribute, the varying age ranges observed within the clusters indicate that age may play a role in the development and progression of diabetes. Additionally, the fasting and postprandial associations in insulin and glucose levels exhibited greater strength in the cluster encompassing individuals with diabetes, where insulin production or action is compromised.

Keywords Cluster analysis · K-means · Glucose · Insulin · Statistical analysis

Introduction

Diabetes, a progressive disease characterized by elevated blood glucose levels resulting from impaired insulin production or action [1, 2], does not have a precise onset. However, pre-diabetic conditions such as impaired glucose tolerance (IGT) and impaired fasting glucose (IFG) predispose individuals to the development of diabetes [3], increasing the risk of cardiovascular diseases [4]. Diagnosis of diabetes, IGT, and IFG often involves the oral glucose tolerance test (OGTT) [5], which measures plasma glucose and insulin levels before and after the intake of glucose. IGT is diagnosed when the plasma glucose level at 120 min is greater than or equal to 140 and less than 200 mg/dL, and IFG is diagnosed when the fasting glucose level is greater than or equal to 100 and less than 126 mg/dL [1].

As urbanization and sedentary lifestyles proliferate, diabetes and prediabetes prevalence have surged globally. Urban areas harbor two-thirds of adults with diabetes, while a third reside in rural regions [6]. The prevalence of diabetes in adults in 2017 was estimated in 425 million cases (almost 9% of adults) but it was estimated that 629 million people (about 10% of adults) will suffer from diabetes in 2045. Similarly, the prevalence of IGT in adults in 2017 was estimated in

✉ Miguel Altuve
maltuve@usb.ve

Erika Severeyn
severeynerika@usb.ve

[1] Valencian International University, Valencia, Spain

[2] Applied Biophysics and Bioengineering Group, Simon Bolivar University, Caracas, Venezuela

374 million cases (almost 8% of adults), 69% of which live in low- and middle-income countries, but it was estimated that 587 million adults (8.4% of adults) will suffer from IGT in 2045 [6]. The North American and Caribbean region has the highest prevalence of IGT (13.6%) and the Southeast Asia region has the lowest one (3.4%). The prevalence of IFG is between 43.9% and 58% for Caucasians and between 29.2% and 48.1% for Asians [7].

Numerous methodologies have been used for the diagnosis of diabetes and prediabetes. Support vector machines, genetic algorithms, and k-means have been used for the diagnosis of diabetes [8] and gestational diabetes [9] using datasets that include plasma glucose, plasma insulin, and blood pressure, among others. Likewise, linear regression has been used for the diagnosis of IGT to correlate plasma metabolite values with insulin and OGTT glucose [10]. Additionally, machine learning algorithms have been explored to predict the evolution of diabetes in patients at high risk [11]. Studies that explored the categorization of subjects into more specific groups, taking into account their distinctive metabolic characteristics, emphasize that such stratification can significantly enhance the quality of life for patients by enabling the development of personalized treatments tailored to their specific metabolic conditions [12, 13].

Building upon these concepts, this study aims to cluster subjects based on fasting and postprandial plasma glucose and insulin levels using the k-means algorithm. This approach unveils patterns reflecting metabolic deterioration processes, enriching our understanding of these conditions. As simple lifestyle changes can reverse prediabetes, early detection is crucial. However, the traditional OGTT test's invasiveness and costliness warrant alternative diagnostic methodologies that optimize accuracy and efficiency. These alternatives, often incorporating non-invasive variables, offer promise, especially in resource-limited settings. Early diagnosis and prediction of diabetes onset are pivotal for improved patient outcomes, tailored treatments, extended life expectancy, symptom alleviation, and complication prevention. The novelty of this paper lies in the introduction of a novel approach to differentiate individuals with varying metabolic conditions using clustering analysis. This approach holds the potential to transform diabetes diagnosis and management. By utilizing clustering analysis to categorize individuals based on comprehensive metabolic markers, this methodology captures nuanced variations that traditional tests might overlook. The results provide insight into metabolic deviations even before blood sugar levels reach the diabetic range, enabling proactive interventions with personalized treatments aligned to individual metabolic dysregulations.

The significance of this study is rooted in its ability to uncover intricate patterns and relationships among metabolic markers, contributing to a more refined understanding of distinct metabolic conditions. Beyond its diagnostic value, the proposed approach stands as an alternative diagnostic tool, transcending the limitations of existing tests and ensuring individuals at risk receive appropriate attention. As we navigate the complex landscape of diabetes research, this novel approach to categorizing individuals based on metabolic conditions offers a beacon of hope, potentially enhancing diagnosis accuracy, treatment efficacy, and overall patient well-being.

Methods and Materials

Dataset

The age and levels of glucose and insulin in plasma in fasting from 2835 women, collected in a previous study at Caracas University Hospital, Venezuela, were employed in this work. Please refer to [13] for a more detailed explanation of the clinical protocol and the dataset in general. The dataset is openly accessible on the IEEE DataPort platform at https://dx.doi.org/10.21227/5g52-jc59 [14].

k-means clustering

In this work, our primary focus was to analyze the clustering outcomes across various attribute combinations. To accomplish this, we performed four distinct clustering experiments by changing the attributes as follows:

1st clustering

The first clustering experiment was performed using only fasting glucose levels (G0). This experiment is important since estimating the fasting glucose concentration is a feasible task to be performed by an inexperienced person using, for example, commercial devices, such as a glucometer or a smartphone application [15, 16], or wearable technology, such as a contact lens [17, 18]. Therefore, discovering physiological patterns or clinical indicators using only fasting blood glucose concentration could have a significant impact on society since people could obtain useful information without going to a specialized clinical laboratory.

2nd clustering

The second clustering experiment uses fasting and postprandial glucose levels (G0&G120). Although two blood samples are required to perform this experiment and ingest a certain glucose concentration, which indicates that it is a more expensive and complicated procedure than the previous one, most specialized laboratories perform this type of test to diagnose metabolic diseases. Therefore, discovering

physiological patterns or clinical indicators using fasting and postprandial blood glucose concentrations could be a useful tool for clinical laboratories.

3rd clustering

The third one uses fasting glucose and insulin levels (G0&I0). The measurement of the concentration of insulin in the blood is carried out with specialized equipment in a clinical laboratory, therefore, it is an expensive procedure that requires specialized personnel. This experiment, like the previous one, could yield useful information for clinical laboratories.

4th clustering

The fourth and final clustering experiment uses fasting and postprandial glucose and insulin levels (G0&G120&I0&I120). Observing these four variables at the same time can help contribute to discovering patterns or clinical indicators in the data. Besides, we recall that the dataset was classified into 28 classes based on fasting and postprandial glucose and insulin values defined in the literature [13]. However, this test would require two blood samples, the intake of a certain amount of glucose, and a specialized instrument to determine the concentration of insulin in the blood (glucose concentration can also be obtained with a specialized instrument, however, as mentioned, a person can estimate its blood glucose concentration using, for example, a glucometer). This experiment would, therefore, have an impact on clinical laboratories.

Since our goal is to group the subjects based on blood glucose and insulin concentrations, the age variable was not considered as an attribute in the clustering process, however, it was used for analyzing the groups created. The number of subjects per group was determined for each clustering experiment.

The squared Euclidean distance was used as the distance measure in which, for each attribute, each centroid corresponds to the mean of the values of the attribute of the individuals assigned to that cluster.

Since hyperparameter k must be specified before the clustering process, the elbow and silhouette methods were used to estimate the number of clusters. In the case of the elbow method, for different values of k, we performed the clustering process and computed the percentage of variance explained. Then, the optimal number of clusters corresponds to the lowest k that gives 90% of the percentage of variance explained. The elbow method was also employed by considering the total within-cluster sum of squares (WCSS) as a function of the number of clusters. In the case of the Silhouette clustering evaluation criterion, again for different values of k, we performed the clustering process and computed the average Silhouette of observations. Then, the optimal number of clusters is the one that provides the highest average Silhouette value.

Statistical analysis

The statistical analysis applied in this study ensures a robust and objective evaluation of the data, allowing for the identification of significant differences, relationships, and patterns within the study variables.

The mean and standard deviation of age, fasting glucose, postprandial glucose, fasting insulin, and postprandial insulin levels were computed for the subjects assigned to the clusters to provide important descriptive information about these subjects and to summarize the central tendencies and variability of the variables under investigation.

The Kruskal-Wallis nonparametric test was utilized to identify significant differences between the clusters for age, fasting glucose, postprandial glucose, fasting insulin, and postprandial insulin levels. This test is appropriate for comparing independent samples and determining if there are statistically significant variations among the clusters. To further analyze and validate the findings, Tukey's honestly significant difference test was employed as a post hoc analysis. This test allows for pairwise comparisons between the clusters, revealing specific differences between groups and providing more detailed insights into the variations observed. Moreover, the Wilcoxon signed-rank test was utilized to assess the significant differences between fasting and postprandial glucose and insulin levels. This test is suitable for analyzing dependent samples and allows for a thorough investigation of the changes in these variables within individuals.

Lastly, the use of the Pearson correlation coefficient (ρ) facilitated the examination of the linear relationship between glucose and insulin levels within each cluster. This measure quantifies the strength and direction of the correlation, providing valuable insights into the interdependence of these variables.

Statistical significance was defined as a p-value of 5% or less.

Results

A capsule in Code Ocean to reproduce the results of this paper is available at [19]. Moreover, a GitHub repository was created with the MATLAB code [20].

Based on the results of applying the elbow and average Silhouette methods for deciding the number of clusters (See details in the supplementary material), we have opted for three clusters ($k = 3$) for all experiments.

Table 1 provides an overview of the subject distribution across clusters for the various clustering experiments con-

Table 1 Number of subjects assigned to each cluster based on different attribute combinations used by the clustering algorithm

Cluster	Attributes			
	G0	G0&G120	G0&I0	G0&I0&G120&I120
C1	2080	2013	2124	2256
C2	724	724	680	441
C3	31	98	31	138

ducted, considering different attribute combinations such as G0, G0&G120, G0&I0, and G0&G120&I0&I120. Cluster C1 comprises the largest number of subjects, while Cluster C3 has the lowest representation. The variation in subject distribution suggests the influence of different attribute combinations on the clustering outcomes.

Table 2 presents the results for the variables age, G0, G120, I0, and G120 of the subjects assigned to each cluster in the conducted clustering experiments. This table provides a comprehensive understanding of the distribution of insulin and glucose levels within the three clusters and reveals significant differences in variables between clusters and the cluster's association with glucose and insulin levels in fasting and postprandial states. Cluster C1 particularly stands out with the lowest average age and consistently lower levels of fasting and postprandial plasma glucose and insulin. The supplementary material contains box plots illustrating the distribution of these variables across clusters for each attribute combination involved in the clustering process.

Table 3 shows the Pearson correlation coefficients in each cluster (C1, C2, C3) between the variables of glucose and insulin in fasting and postprandial, in the four experiments carried out (with attributes G0, G0&G120, G0&I0, and G0&I0&G120&I120). We can see that the lowest correlation, $\rho = -0.008$, was obtained in the cluster C1 between the variables G0 and I120 using G0&G120 as attributes of the clustering process, whereas the highest correlation, $\rho = 0.714$, was obtained in the cluster C3 between the variables G0 and G120 both using G0 as attribute and using G0&I0 as attributes of the clustering process. These correlations provide insights into the relationships between variables within each cluster. The strong correlations between glucose levels at different time points (G0 and G120) suggest a consistent pattern of glucose behavior. Similarly, the positive correlations between insulin levels at different time points (I0 and I120) indicate a similar pattern.

Discussion

In all the experiments carried out with the different attribute combinations, subjects in cluster C1 have G0 and G120 lower than subjects in cluster C2, and they, in turn, have G0 and G120 lower than subjects in cluster C3. It is important to note

Table 2 Mean ± standard deviation of the variables per cluster for each attribute used in the clustering process

Variable	Cluster	Attributes			
		G0	G0&G120	G0&I0	G0&I0&G120&I120
Age	C1	39.83 ± 14.57 [1 2]	40.05± 14.74 [1 2 3]	40.07±14.61 [1 2]	41.87±14.81 [2 3]
	C2	50.11 ± 12.96	48.33 ±13.46	50.04±13.00	43.07±15.06
	C3	50.65 ± 10.04	52.00 ± 10.67	50.65±10.04	52.58±10.43
G0	C1	91.16±7.25 [1 2 3 4]	93.13 ±9.50 [1 2 3 4]	91.44±7.44 [1 2 3 4]	95.73±12.38 [1 2 3 4]
	C2	116.09±13.24 [4]	107.01 ± 16.52 [4]	116.82±13.35 [4]	99.37±12.83 [4]
	C3	234.48±64.54 [4]	163.22 ±61.80 [4]	234.48±64.54 [4]	153.25±55.44 [4]
G120	C1	107.54±23.62 [1 2 3]	100.52 ±13.53 [1 2 3]	107.88±23.95 [1 2 3]	107.52±21.75 [1 2 3]
	C2	145.30±50.34	149.25 ±22.89	146.70±50.96	136.33±29.80
	C3	313.55±112.95	287.78 ±66.85	313.55±112.95	260.34±70.39
I0	C1	8.01±6.70 [1 2 5]	8.03 ±6.74 [1 2 3 5]	7.81±6.48 [1 2 5]	7.35±5.82 [1 2 3 5]
	C2	11.01±8.79 [5]	10.21±8.00 [5]	11.82±9.13 [5]	15.13±9.80 [5]
	C3	12.39±8.87 [5]	14.99 ±11.45 [5]	12.39±8.87 [5]	12.82±9.96 [5]
I120	C1	62.94±46.03 [1 2 3]	55.08 ±38.74 [1 2 3]	62.08±44.95 [1 2 3]	47.79±22.06 [1 2 3]
	C2	73.27±51.83	91.35 ±56.91	76.61±54.37	153.11±48.34
	C3	38.45±28.57	83.03 ±60.84	38.45±28.57	71.15±41.02

[1] The variable shows significant difference between clusters C1 and C2
[2] The variable shows significant difference between clusters C1 and C3
[3] The variable shows significant difference between clusters C2 and C3
[4] The cluster shows significant difference between variables G0 and G120
[5] The cluster shows significant difference between variables I0 and I120

Table 3 Correlations between variables within each cluster

Cluster	Variable	Attributes											
		G0			G0&G120			G0&I0			G0&I0&G120&I120		
		G120	I0	I120	G120	I0	I120	G120	I0	I120	G120	I0	I120
C1	G0	0.251	0.138	0.092	0.259	0.159	−0.008	0.262	0.088	0.060	0.438	0.163	0.027
	G120		0.142	0.448		0.089	0.338		0.121	0.432		0.016	0.289
	I0			**0.513**			0.492			**0.504**			0.411
C2	G0	**0.587**	0.112	0.038	0.196	0.117	−0.165	**0.588**	0.024	−0.023	0.430	0.155	0.049
	G120		0.095	0.267		−0.019	0.110		0.056	0.239		−0.049	0.151
	I0			0.492			0.484			0.492			0.282
C3	G0	**0.714**	0.319	−0.030	**0.701**	−0.058	−0.439	**0.714**	0.319	−0.030	**0.699**	0.093	−0.361
	G120		0.345	−0.089		−0.084	−0.377		0.345	−0.089		0.074	−0.231
	I0			**0.530**			**0.559**			**0.530**			0.460

Correlations that are statistically significant and greater than ±0.5 are highlighted in bold text

that we have named cluster C1 the one with the most subjects and cluster C3 the one with fewer subjects, as shown in Table 1. Moreover, the results of the clustering algorithm are in accordance with the diagnostic values of diabetes and prediabetes stated by the American Diabetes Association [1]: Cluster C2 contains prediabetics, cluster C3 contains diabetic, and cluster C1 contains subjects without any of the above pathologies.

Age is an important factor in the development of diabetes, which is reflected in the clustering performed. In all experiments, subjects in cluster C1 were younger than subjects in cluster C2, and they, in turn, were younger than subjects in cluster C3. Consequently, as we age, we are more likely to suffer from diabetes. Moreover, a chronological trend was observed: the youngest tend to be healthy (cluster C1), prediabetes tends to occur in middle-aged adults (cluster C2) and diabetes tends to appear at an old-age (cluster C3) [21].

By considering the insulin concentration, on the one hand, in all the experiments, cluster C1 grouped subjects with lower I0 and greater insulin sensitivity (mean HOMA-IR < 1.84). This is an expected result since these subjects tend to have normal fasting glucose and insulin levels. Clustering experiments with G0, G0&G120, and G0&I0 as attributes, grouped subjects with higher I0 levels in cluster C3. These subjects have lower insulin sensitivity (mean HOMA-IR> 6.04). It is known that low insulin sensitivity can cause the development of diabetes. This can be reflected by the fact that subjects in cluster C2 have a mean HOMA-IR> 2.6, indicating that these subjects (with glucose values consistent with prediabetes) have insulin resistance [22]. On the other hand, in all the experiments, cluster C2 contains subjects with elevated levels of I120, which could indicate that subjects with prediabetes tend to produce more insulin for glucose metabolism than normal. These findings are in line with the fact that prediabetes refers to a state in which blood glucose levels are higher than normal but not high enough to be diagnosed as type 2 diabetes. In this state, the body can produce more insulin to try to maintain blood glucose levels within the normal range. However, over time, the body can become less sensitive to insulin, which means that more insulin is needed to achieve the same effect on reducing blood glucose levels [23]. This decrease in insulin sensitivity is known as insulin resistance and can be a precursor to type 2 diabetes [24].

Cluster C3 contains the lowest level of I120 when the clustering is carried out with G0 and G0&I0 as attributes. Subjects in cluster C3 are mostly diabetic as indicated by their fasting and postprandial glucose levels, their insulin production is thus compromised due to the diabetic condition [25]. In the presence of diabetes, it is common to achieve low insulin levels in the body, this is because, over time, beta cells of the pancreas can wear out due to excessive insulin production. Therefore, diabetic subjects usually have low levels of insulin production, which can make it difficult to control blood glucose levels. It is important that people with diabetes work with their physician to develop a treatment plan that includes changes in diet, exercise, and medicines to help control diabetes and prevent long-term complications [26].

The influence of various factors on insulin utilization in our bodies is widely recognized and well-established. For instance, some genes modulate insulin production and insulin receptors, and it is believed that both insulin quality and insulin receptor endocytosis signaling may be associated with insulin resistance and the development of type 2 diabetes [27, 28]. Also, an increase in body fat results in a decrease in adiponectin production by adipose tissue. Adiponectin plays an important role in regulating insulin sensitivity and possesses anti-inflammatory properties. Consequently, when the body accumulates excess adipose tissue, adiponectin levels tend to decrease, potentially leading to insulin resistance and an elevated risk of developing type 2 diabetes and other metabolic disorders. Therefore, main-

taining a healthy weight and an appropriate amount of adipose tissue becomes crucial for achieving optimal levels of adiponectin and improving insulin sensitivity [29]. Numerous studies have demonstrated that individuals who engage in regular exercise exhibit enhanced insulin sensitivity compared to people with sedentary lifestyles. This effect stems from the exercise-induced signaling of glucose receptor exocytosis, which promotes glucose uptake and reduces the reliance on high insulin levels for facilitating glucose uptake [30].

It is interesting to note the behavior of the correlations between the variables as the metabolic disease progresses (from cluster C1 to C3). For instance, clustering experiments with G0 and G0&I0 show significant correlations above 0.5 between fasting and postprandial insulin in cluster C1 and between fasting and postprandial glucose in cluster C2. The association between fasting and postprandial glucose appears in cluster C2 (mostly composed of prediabetics) but it is then strengthened in cluster C3 (mostly composed of diabetics), with correlations around 0.7, for all the clustering experiments. This suggests that the association between fasting and postprandial plasma glucose concentration is marked in subjects with problems with insulin production or action [31]. Two things could be therefore highlighted: (i) fasting and postprandial glucose values become highly correlated when subjects develop diabetes and (ii) the clustering method can group diabetic subjects using any combination of attributes, even just using G0. This does not apply to the other two groups, particularly prediabetics, who would be the most important group because, with the right treatment, diabetes could be prevented.

The findings of our study carry significant implications for the understanding and management of diabetes and related metabolic conditions. Through a comprehensive analysis of correlations between fasting and postprandial glucose and insulin levels within distinct clusters, we have illuminated previously unexplored temporal relationships between these key metabolic markers.

Temporal relationships and mechanistic insights

Our analysis revealed compelling temporal associations that deepen our understanding of glucose metabolism. For instance, individuals identified within the cluster exhibiting elevated fasting glucose levels also demonstrated correspondingly high postprandial glucose levels. This observation suggests potential difficulties in postprandial glucose clearance, hinting at an underlying mechanism of impaired glucose tolerance. This insight provides a foundation for the development of targeted interventions aimed at ameliorating glucose intolerance mechanisms, thereby advancing our ability to tailor treatments to specific metabolic dysregulations.

Personalized treatment strategies

The categorization of individuals into distinct clusters based on their metabolic profiles holds promise for the development of personalized treatment strategies. The stratification of clusters allows for the identification of optimal interventions that align with individual metabolic conditions. For instance, individuals within the cluster characterized by elevated fasting glucose levels might benefit from interventions aimed at directly managing blood sugar levels, while those within the cluster with the highest postprandial glucose levels could benefit from interventions designed to enhance insulin sensitivity. Such tailored treatments align with the evolving paradigm of precision medicine and have the potential to enhance treatment outcomes and patient well-being significantly.

Alternative diagnostic approach

The utilization of clustering algorithms to analyze fasting and postprandial glucose and insulin levels introduces an innovative diagnostic approach with notable advantages over conventional methods. Traditional diagnostic tests for diabetes, such as the OGTT, are not only resource-intensive but also time-consuming. Our proposed cluster analysis methodology offers a streamlined alternative by effectively identifying individuals at risk for diabetes or prediabetes, even if they fall below the diagnostic thresholds of formal criteria. This approach has the potential to facilitate early diagnosis and intervention, contributing to improved outcomes for individuals at risk.

Conclusion

This work has shown that the k-means clustering algorithm with three clusters groups the subjects into healthy, prediabetic, and diabetic, by using fasting and postprandial glucose and insulin levels. In addition, it also groups subjects with low insulin sensitivity. Age can also be considered an important factor in the development of diabetes since subjects with diabetes and prediabetes (clusters C3 and C2, respectively) were among the oldest subjects. Low insulin sensitivity could be another risk factor in the development of diabetes. Clusters C2 and C3 contain subjects with HOMA-IR$>$ 2.5, being the subjects of cluster C3 (diabetics) who presented a greater detriment in insulin sensitivity.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13410-024-01322-8.

Author contribution Miguel Altuve carried out the experiments and contributed to the conception of the study and the writing and revising

of the manuscript. Erika Severeyn contributed to the conception of the study, the analysis of the results, and the writing and revising of the manuscript.

Data availability The dataset used in this study is freely available at https://ieee-dataport.org/documents/fasting-and-postprandial-glucose-and-insulin-dataset.

Code availability The code used in this study is freely available at https://doi.org/10.24433/CO.7408455.v1 and at https://github.com/miguelaltuve/ClusterAnalysisBasedFastingPostprandialGlucoseInsulin.

Declarations

Ethics approval All the procedures performed in the study were adjusted to the ethical standards of the University Hospital of Caracas and the Declaration of Helsinki of 1964 and its subsequent amendments or comparable ethical standards.

Consent to participate Informed consent was obtained from all participants included in the study.

Conflict of interest The authors declare no competing interests.

References

1. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, Collins BS, Hilliard ME, Isaacs D, Johnson EL, Kahan S, Khunti K, Leon J, Lyons SK, Perry ML, Prahalad P, Pratley RE, Seley JJ, Stanton RC, Gabbay RA. 2. Classification and diagnosis of diabetes: standards of care in diabetes–2023. Diabetes Care. 2023;46:19–40. https://doi.org/10.2337/dc23-S002.
2. Fonseca VA. Defining and characterizing the progression of type 2 diabetes. Diabetes Care. 2009;32(suppl 2):151–6. https://doi.org/10.2337/dc09-S301.
3. Bansal N. Prediabetes diagnosis and treatment: a review. World J Diabetes. 2015;6(2):296–303. https://doi.org/10.4239/wjd.v6.i2.296.
4. Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. BMJ. 2016;355:5953. https://doi.org/10.1136/bmj.i5953.
5. Hulman A, Vistisen D, Glümer C, Bergman M, Witte DR, Færch K. Glucose patterns during an oral glucose tolerance test and associations with future diabetes, cardiovascular disease and all-cause mortality rate. Diabetologia. 2018;61(1):101–7. https://doi.org/10.1007/s00125-017-4468-z.
6. Cho N, Shaw J, Karuranga S, Huang Y, da Rocha Fernandes J, Ohlrogge A, Malanda B. IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract. 2018;138:271–81. https://doi.org/10.1016/j.diabres.2018.02.023.
7. Yip W, Sequeira I, Plank L, Poppitt S. Prevalence of pre-diabetes across ethnicities: a review of impaired fasting glucose (IFG) and impaired glucose tolerance (IGT) for classification of dysglycaemia. Nutrients. 2017;9(11):1273. https://doi.org/10.3390/nu9111273.
8. Shakeel PM, Baskar S, Dhulipala VS, Jaber MM. Cloud based framework for diagnosis of diabetes mellitus using K-means clustering. Health Inf Sci Syst. 2018;6(1):16. https://doi.org/10.1007/s13755-018-0054-0.
9. Santhanam T, Padmavathi M. Application of K-means and genetic algorithms for dimension reduction by integrating SVM for diabetes diagnosis. Procedia Comput Sci. 2015;47:76–83. https://doi.org/10.1016/j.procs.2015.03.185.
10. Wildberg C, Masuch A, Budde K, Kastenmüller G, Artati A, Rathmann W, Adamski J, Kocher T, Völzke H, Nauck M, et al. Plasma metabolomics to identify and stratify patients with impaired glucose tolerance. J Clin Endocrinol Metab. 2019. https://doi.org/10.1210/jc.2019-01104.
11. Cahn A, Shoshan A, Sagiv T, Yesharim R, Raz I, Goshen R. Use of a machine learning algorithm improves prediction of progression to diabetes. American Diabetes Association. 2018. https://doi.org/10.2337/db18-1286-P.
12. Ahlqvist E, Storm P, Käräjämäki A, Martinell M, Dorkhan M, Carlsson A, Vikman P, Prasad RB, Aly DM, Almgren P, 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(5):361–9. https://doi.org/10.1016/S2213-8587(18)30051-2.
13. Altuve M, Severeyn E. Joint analysis of fasting and postprandial plasma glucose and insulin concentrations in Venezuelan women. Diabetes Metab Syndr: Clin Res Rev. 2019;13(3):2242–8. https://doi.org/10.1016/j.dsx.2019.05.029.
14. Altuve M, Severeyn E. Fasting and postprandial glucose and insulin dataset. IEEE Dataport. 2019. https://doi.org/10.21227/5g52-jc59
15. Bandodkar AJ, Imani S, Nunez-Flores R, Kumar R, Wang C, Mohan AV, Wang J, Mercier PP. Re-usable electrochemical glucose sensors integrated into a smartphone platform. Biosens Bioelectron. 2018;101:181–7. https://doi.org/10.1016/j.bios.2017.10.019.
16. Yang J. Blood glucose monitoring with smartphone as glucometer. Electrophoresis. 2019;40(8):1144–7. https://doi.org/10.1002/elps.201800295.
17. Farandos NM, Yetisen AK, Monteiro MJ, Lowe CR, Yun SH. Contact lens sensors in ocular diagnostics. Adv Healthc Mater. 2015;4(6):792–810. https://doi.org/10.1002/adhm.201400504.
18. Badugu R, Reece EA, Lakowicz JR. Glucose-sensitive silicone hydrogel contact lens toward tear glucose monitoring. J Biomed Opt. 2018;23(5):057005. https://doi.org/10.1117/1.JBO.23.5.057005.
19. Altuve M. Cluster analysis based on fasting and postprandial plasma glucose and insulin concentrations [Source Code]. 2024. https://doi.org/10.24433/CO.7408455.v1
20. Altuve M. Cluster analysis based on fasting and postprandial plasma glucose and insulin concentrations [Source Code]. 2024. https://github.com/miguelaltuve/ClusterAnalysisBasedFastingPostprandialGlucoseInsulin
21. Ha KH, Kim DJ. Trends in the diabetes epidemic in Korea. Endocrinol Metab. 2015;30(2):142–6. https://doi.org/10.3803/EnM.2015.30.2.142.
22. Czech MP. Insulin action and resistance in obesity and type 2 diabetes. Nat Med. 2017;23(7):804–14. https://doi.org/10.1038/nm.4350.
23. Muller D, Elahi D, Tobin J, Andres R. The effect of age on insulin resistance and secretion: a review. Semin Nephrol. 1996;16(4):289–98.
24. Johnson JD. On the causal relationships between hyperinsulinaemia, insulin resistance, obesity and dysglycaemia in type 2 diabetes. Diabetologia. 2021;64:2138–46. https://doi.org/10.1007/s00125-021-05505-4.
25. DeFronzo RA, Ferrannini E, Groop L, Henry RR, Herman WH, Holst JJ, Hu FB, Kahn CR, Raz I, Shulman GI, et al. Type 2 diabetes mellitus. Nat Rev Dis Primers. 2015;1:15019. https://doi.org/10.1038/nrdp.2015.19.
26. Padhi S, Nayak AK, Behera A. Type II diabetes mellitus: a review on recent drug based therapeutics. Biomed Pharmacother. 2020;131:110708. https://doi.org/10.1016/j.biopha.2020.110708.

27. Halban PA, Polonsky KS, Bowden DW, Hawkins MA, Ling C, Mather KJ, Powers AC, Rhodes CJ, Sussel L, Weir GC. β-cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. J Clin Endocrinol Metab. 2014;99(6):1983–92. https://doi.org/10.1210/jc.2014-1425.
28. Hall C, Yu H, Choi E. Insulin receptor endocytosis in the pathophysiology of insulin resistance. Exp Mol Med. 2020;52(6):911–20. https://doi.org/10.1038/s12276-020-0456-3.
29. Balsan GA, Vieira JLdC, Oliveira AMd, Portal VL. Relationship between adiponectin, obesity and insulin resistance. Rev Assoc Med Bras. 2015;61:72–80. https://doi.org/10.1590/1806-9282.61.01.072.
30. Yaribeygi H, Atkin SL, Simental-Mendía LE, Sahebkar A. Molecular mechanisms by which aerobic exercise induces insulin sensitivity. J Cell Physiol. 2019;234(8):12385–92. https://doi.org/10.1002/jcp.28066.
31. Ketema EB, Kibret KT. Correlation of fasting and postprandial plasma glucose with hba1c in assessing glycemic control; systematic review and meta-analysis. Arch Pub Health. 2015;73(1):43. https://doi.org/10.1186/s13690-015-0088-6.