ORIGINAL ARTICLE

The performance of hemoglobin A1c against fasting plasma glucose and oral glucose tolerance test in detecting prediabetes and diabetes

Jale Karakaya, Safak Akin[1], Ergun Karagaoglu, Alper Gurlek[1]
Departments of Biostatistics and [1]Endocrinology and Metabolism, Faculty of Medicine, Hacettepe University, Ankara, Turkey

Background: In recent years, hemoglobin A1c (HbA1c) is accepted among the algorithms used for making diagnosis for diabetes and prediabetes since it does not require subjects to be prepared for giving a blood sample. The aim of this study is to assess the performance of HbA1c against fasting plasma glucose (FPG) and oral glucose tolerance test (OGTT) in detecting prediabetes and diabetes. **Materials and Methods:** A total of 315 subjects were included in this study. The success of HbA1c in distinguishing the three diagnostic classes was examined by three-way receiver operating characteristic (ROC) analysis. The best cut-off points for HbA1c were found for discriminating the three disease status. **Results:** The performance of HbA1c, measured by the volume under the ROC surface (VUS), is found to be statistically significant (VUS = 0.535, $P < 0.001$). The best cut-off points for discriminating between normal and prediabetes groups and between prediabetes and diabetes groups are $c_1 = 5.2\%$ and $c_2 = 6.4\%$ respectively. **Conclusion:** The performance of HbA1c in distinguishing between the prediabetes and diabetes groups was higher than its ability in distinguishing between healthy and prediabetes groups. This study provides enough information to understand what proportion of diabetes patients were skipped with the HbA1c especially when the test result is healthy or prediabetes. If a subject was diagnosed as healthy or prediabetes by HbA1c, it would be beneficial to verify the status of that subject by the gold standard test (OGTT and FPG).

Key words: Diabetes mellitus, hemoglobin A1c, receiver operating characteristic surface, three-way receiver operating characteristic, volume under receiver operating characteristic surface

How to cite this article: Karakaya J, Akin S, Karaagaoğlu E, Gurlek A. The performance of hemoglobin A1c against fasting plasma glucose and oral glucose tolerance test in detecting prediabetes and diabetes. J Res Med Sci 2014;19:1051-7.

INTRODUCTION

Diabetes mellitus (DM) is a group of metabolic diseases characterized by defects in insulin secretion and/or insulin action. Diabetes is rapidly becoming one of the most common diseases globally.[1] Type 2 DM remains a leading cause of end-stage renal failure, amputations, cardiovascular disorders, blindness, and hospitalizations. It is also associated with an increased risk of cancer and deadly complications.[2]

Plasma hemoglobin A1c (HbA1c) is a widely used marker of chronic glycemia, and reflects mean ambient fasting and postprandial glycemia over a 2-3 months period. It is a suitable and well-known biomarker in clinical practice. HbA1c is formed by the slow irreversible, nonenzymatic glycation of valine and lysine residues in the hemoglobin molecule. It is a useful tool for characterizing dysglycemia as it is easier to perform compared with the oral glucose tolerance test (OGTT).[3] Elevated HbA1c levels are likely the result of long-term insulin resistance; metabolic disturbances including hyperglycemia, dyslipidemia, hypercoagulability and inflammation might be the major pathogenic mechanism for the adverse impact of elevated HbA1c in the setting of cardiovascular and cerebrovascular disease.[4] However the discriminatory power of HbA1c is not well studied in its ability to diagnose patients as healthy, suffering from prediabetes, or suffering from diabetes.

The performance of a diagnostic test with continuous numerical outcome is often assessed by receiver operating characteristic (ROC) curve analysis. The ROC analysis is a method used to determine the accuracy of a diagnostic test for a two-class (i.e., diseased or healthy) situations determined through the gold standard test. In medicine, however, besides the existence of many instances where the results of two-class gold standard tests are used, there are also other quite frequent cases in which the diagnosis is related to >2 real situations. For example, the real situation may have three classes such as "diseased," "risky," "healthy;" "malign," "benign," "suspected" or "phase I," "phase II," "phase III."

Address for correspondence: Dr. Alper Gurlek, Department of Internal Medicine, Hacettepe University School of Medicine, Section of Endocrinology and Metabolism, 06100 Sihhiye, Ankara, Turkey. E-mail: agurlek2001@yahoo.com
Received: 23-10-2013; **Revised:** 29-06-2014; **Accepted:** 31-10-2014

A possible solution for situations where there are >2 classes is to transform the gold standard test into a two-class data structure. In such situations, one of the classes may be combined, with one or two additional diagnostic groups formed by discarding one-class. However, Obuchowski showed that transforming a gold standard test that has >2 classes artificially into a two-class data structure may lead to producing higher performance measures than should actually be the case and therefore to a significantly biased estimate of test performance.[5,6]

For assessing the performance of tests for three-class problems Mossman developed the "three-way (dimensional) ROC analysis"[7] in 1999. The performance of a diagnostic test is assessed through "ROC surface" and "volume under the ROC surface (VUS)."[7-10]

The OGTT and fasting plasma glucose (FPG) should be performed after at least 8 h fasting. Fasting is defined as no caloric intake for at least 8 h. The HbA1C has several advantages to the FPG and OGTT, including greater convenience, since fasting is not required. HbA1c provides information about the diabetes status of a subject easily with a blood sample taken from the subject at any time of day. The process is very easy. In recent years, HbA1c is accepted among the algorithms used for making diagnosis and its diagnostic performance needs to be examined in more detail. The performance of HbA1c in discriminating normal, prediabetes and diabetes has not been extensively studied in previous studies. Cut-off points of HbA1c will be determined in this study, so it is possible to determine the group in which the subject falls. But how reliable is this classification is not adequately known. This study will help us to understand what proportion of diabetes patients were skipped with the HbA1c especially when the test result is negative (such as indicating to healthy or prediabetes) and to decide whether or not OGTT test should be applied.

The objective of the present study is to assess the performance of HbA1c when used to diagnose the three categories of diabetes: "Diabetes," "prediabetes," and "healthy" by using three-way ROC analysis. Furthermore, another objective is to determine the best cut-off points for HbA1c at which three classes can be distinguished with the highest correct classification rates (CCR). One of these cut-off points is needed to distinguish between "healthy" and "prediabetes" groups and the other between "prediabetes" and "diabetes" groups.

MATERIALS AND METHODS

Study population and participants

As a total this cross-sectional study was conducted on 315 subjects who were evaluated between July and December 2011 in Hacettepe University Hospital, Department of Internal Medicine, Section of Endocrinology and Metabolism outpatient clinic. Disorders of carbohydrate metabolism in our patients were classified according to the following criteria: Prediabetes as exemplified by impaired fasting glucose (IFG) (FPG 100 mg/dl to 125 mg/dl) or impaired glucose tolerance (2-h values in the OGTT of 140 mg/dl to 199 mg/dl), and overt diabetes (FPG is ≥126 mg/dl, confirmed on a second sample; or when a random blood glucose of ≥200 mg/dl is accompanied by symptoms, or 2-h glucose in the OGTT ≥200 mg/dl; or HbA1c ≥6.5%).[11,12]

The indications for OGTT regardless of age were as follows: Having a first degree relative with type 2 diabetes, being overweight and/or obese (body mass index [BMI] ≥25), hypertension, dyslipidemia, history of gestational diabetes, cardiovascular disease, peripheral arterial disease, cerebrovascular disease and polycystic ovary syndrome. It was also performed in subjects with FPG ≥100 mg/dl and in all subjects over 45 years of age. Patients with known malignancy receiving chemotherapy or using immunosuppressive drugs, and subjects under 16 or over 65 were excluded.

Measurements

Hemoglobin A1c was assessed using high-pressure liquid chromatography using a Primus CLC330 Ghb analyzer (Primus, Kansas City, MO). Serum glucose level was measured with spectrophotometer assay (Human Gesellschaft, Wiesbaden, Germany).

Statistical analysis

Three-way ROC analysis is a method developed on the basis of generalizing binary ROC analysis. As the performance indicator of the test, ROC surface is drawn and the VUS is calculated.[7-10,13-15] While the two-dimensional ROC curve can be drawn with two CCR (sensitivity and specificity), the ROC surface is drawn according to three CCR obtained for each class. The coordinates of the graph ($x = CCR_1$, $y = CCR_2$, $z = CCR_3$) represent three CCR. A test is considered successful when this surface comes closer to the point where three CCR reach their maximum values, or when the volume under the surface becomes larger. The ideal point on the ROC surface is the coordinate (1, 1, 1). This point is where three CCR are 100% and where the test achieves its maximum discriminating value. After having obtained the ROC surface, VUS is calculated as an indicator of the test's performance. This indicator is used as a common measure of the discriminative power of the test. In case this value gives the probability of correctly ranking (Y_1, Y_2, Y_3), the test results of individuals randomly selected from each group in three-class diagnostic problems and assuming that larger test results indicate a more positive test, it is shown as $P(Y_3 > Y_2 > Y_1)$. A diagnostic test is considered "poor" where the

VUS is close to 1/3! = 0.17 (no more than chance) and "good" whereby it approaches to unity.[7-10,13,14] The drawing of the ROC surface and calculation of volume under this surface can be realized through either a parametric or nonparametric method as is the case with two-dimensional situations. Relevant sources can be visited for detailed information about obtaining VUS, its standard error (SE) and confidence interval (CI).[8-10,16] Two-way (dimensional) ROC analysis is used to assess how successful a diagnostic test is in terms of differentiating pairs of diagnostic classes. However, these procedures can be considered *post-hoc* tests following the application of three-dimensional ROC analysis.[9]

The diagnostic performance of HbA1c in determining three classes according to the gold standard is assessed through three-way ROC analysis. The ROC surface for HbA1c is obtained and VUS and its SE are estimated. In order to distinguish three classes from each other, the best cut-off points to maximize the sum of CCR are determined. Of these two cut-off points, one is necessary to distinguish between "healthy" and "prediabetes" and the other for "prediabetes" and "diabetes." Pairwise test was performed to examine how well HbA1c can distinguish between each pair of diagnostic classes.

The MATLAB (matrix laboratory) R2009b program (MathWorks Natick, MA) and R 2.13.1 software (R Foundation for Statistical Computing, Vienna, Austria) were used for three-dimensional ROC analysis. The code developed in Matlab was used for drawing the ROC surface and determining the best cut-off points. By using the package DiagTest3grp developed by Luo and Xiong[17] in the R program, the volume under the surface, SE and CIs were calculated for three-way ROC analysis. Two-way (dimensional) ROC analyses were performed as a *post-hoc* test using Stata Statistical Software release 10 (College Station TX: StataCorp LP) software.

Descriptive statistics were expressed as mean ± standard deviation for quantitative variables and count (percent) for qualitative variables. For group comparison of covariates such as age, sex, BMI etc., the one-way ANOVA test was used for numerical variables and Chi-square test for categorical variables. $P < 0.05$ was selected as the criterion of statistical significance.

The study protocol was approved by the Ethics Committee of Hacettepe University School of Medicine. This study was supported by Hacettepe Unversity, Scientific Research Projects Coordination Unit (010 T02 101 001).

RESULTS

Data related to 315 persons in total were assessed. Taking into account the American Diabetes Association (ADA) criteria, three groups were formed according to FPG and OGTT, defined as "healthy," "prediabetes" and "diabetes." According to FPG and OGTT outcomes, 104 were classified as healthy, 161 as with prediabetes and 50 as with diabetes.

Mean ± standard deviation values of HbA1c are calculated as 5.25 ± 0.48, 5.63 ± 0.56, 7.24 ± 1.89 for healthy, prediabetes and diabetes groups, respectively. Groups are compared in terms of age, sex, BMI, family history of DM, waist and hip circumference variables that may be associated with DM. A statistically significant difference is found between groups in terms of sex, age, waist circumference ($P < 0.05$). Descriptive statistics are presented in Table 1.

By using three-way ROC analysis, the performance of HbA1c in diagnosing "diabetes," "prediabetes," and "healthy" is examined. The ROC surfaces of HbA1c from two different perspectives are given in Figure 1.

The VUS is obtained as 0.535 (CI: 95%: 0.46-0.61, $P < 0.001$). This value is statistically significant and points to a distinguishing power that can be considered to be above the level of chance. By using the bootstrap technique the SE for VUS is estimated to be 0.041. Assuming that CCR of HbA1c for all three classes are of equal importance, the

Table 1: Distribution of covariates according to disease status

Variables	Healthy (n_1 = 104) Count (%)	Prediabetes (n_2 = 161) Count (%)	Diabetes (n_3 = 50) Count (%)	*P*
Sex				
Male	7 (10)	31 (60)	15 (30)	<0.001
Female	97 (40)	130 (50)	35 (10)	
Family history of DM				
No	56 (38)	71 (49)	19 (13)	0.129
Yes	48 (29)	90 (53)	31 (18)	
	Mean±SD	Mean±SD	Mean±SD	
Age	45.90±12.17	48.76±11.59	51.5±9.60	0.014
BMI (kg/m^2)	30.35±6.77	30.64±5.59	32.13±5.66	0.208
Waist circumference (cm)	98.83±7.53	98.63±14.30	103.2±12.31	0.002
Hip circumference (cm)	110.24±12.76	109.94±11.37	112.48±12.1	0.416

SD = Standard deviation; DM = Diabetes mellitus; BMI = Body mass index

best cut-off points are found to be c_1 = 5.2% and c_2 = 6.4%, respectively. When an HbA1c value smaller than 5.2% is diagnosed as "healthy", value from 5.2% to 6.4% is defined as "prediabetes" and finally a value >6.4% as "diabetes." Table 2 gives the level of conformity with HbA1c results classified by the gold standard. Accordingly, CCR corresponding to three classes are obtained, respectively, as CCR_1 = 57%, CCR_2 = 65%, CCR_3 = 64%. Subjects who have diabetes are incorrectly classified as healthy (6%) or prediabetes (30%) according to HbA1c. A total of 104 subjects are classified as healthy by HbA1c test. However, 40% (42/104) and 3% (3/104) of these subjects are actually prediabetes and diabetes, respectively. When the subjects' status are classified as diabetes by HbA1c (n = 48), 29% (14/48) of them are prediabetes and 4% (2/48) are healthy.

In order to better understand how well HbA1c can distinguish between each pair of diagnostic classes, *post-hoc* pairwise comparisons are conducted. ROC curves showing the discriminative ability of HbA1c between two diagnostic classes are presented in Figure 2a-c. The diagnostic performance of HbA1c in distinguishing between healthy and prediabetes groups, between healthy and diabetes groups, and between prediabetes and diabetes groups yielded areas under curve ± SE as 0.697 ± 0.032 [Figure 2a], 0.909 ± 0.029 [Figure 2b], and 0.817 ± 0.039 [Figure 2c], respectively (P < 0.001). Given these results, it can be concluded that the performance of HbA1c in distinguishing between the prediabetes and diabetes groups is higher than when distinguishing between the healthy and prediabetes groups. As expected HbA1c has a much higher performance when distinguishing between healthy and diabetes groups.

DISCUSSION

Actual disease status may not always turn out as a two-class problem, such as diseased and nondiseased. In some cases disease status may have a three-class nature. Clinicians may need to distinguish three classes of illness by applying a single diagnostic test. In such cases, trying to

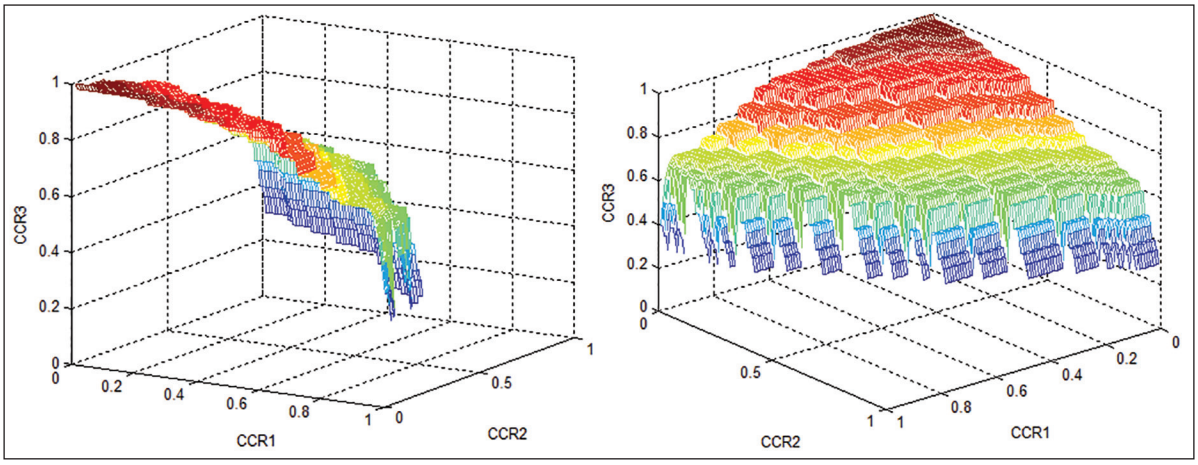

Figure 1: The receiver operating characteristic surfaces of hemoglobin A1c from two different perspectives

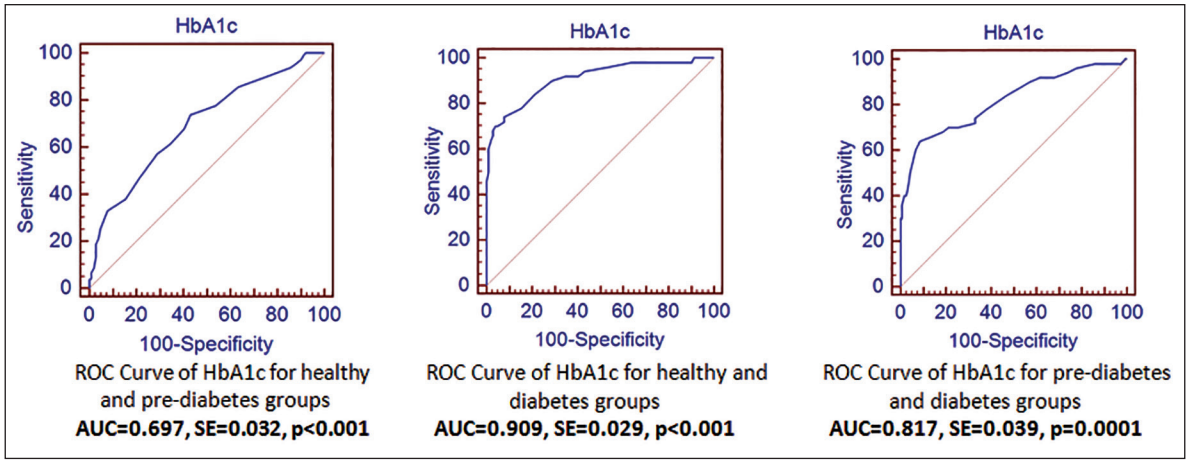

Figure 2: Receiver operating characteristic (ROC) curves of hemoglobin A1c (HbA1c) for pairwise comparisons. (a) ROC curve of HbA1c for healthy and prediabetes groups (areas under curve [AUC] = 0.697, standard error [SE] = 0.032, P < 0.001). (b) ROC curve of HbA1c for healthy and diabetes groups (AUC = 0.909, SE = 0.029, P < 0.001). (c) ROC curve of HbA1c for prediabetes and diabetes groups (AUC = 0.817, SE = 0.039, P = 0.001)

Table 2: 3 × 3 contingency table describing the performance of HbA1c according to the gold standard test

HbA1c	The results of gold standard test			Total
	Healthy *n* (%)	Prediabetes *n* (%)	Diabetes *n* (%)	
Healthy	59 (57)	42 (26)	3 (6)	104
Prediabetes	43 (41)	105 (65)	15 (30)	163
Diabetes	2 (2)	14 (9)	32 (64)	48
Total	104 (100)	161 (100)	50 (100)	315

HbA1c = Hemoglobin A1c

reduce three classes to two is the most frequently observed solution. However, this method may prove insufficient for determining the performance of the test, which may lead to both loss of information and biased estimation of the performance of the diagnostic test. Since one of the groups is often omitted in such assessments, this method should not be used for overall comparison. Instead, three-way ROC analysis capable of assessing the performance of the test simultaneously in all groups should be performed. At the later stage, two-class comparisons can be made by two-way ROC analysis. Furthermore, three-way ROC analysis also provides a solution for determining the best cut-off points used when distinguishing three-class cases. This method provides the opportunity to find two cut-off points to distinguish three classes from each other.

The most commonly used screening tests for type 2 diabetes include measurement of FPG, OGTT and HbA1c. The Expert Committee on Diagnosis and Classification of DM issued a consensus report in June 2009, recommending that a HbA1c level ≥6.5% be used to diagnose diabetes, and the ADA acknowledged this decision on 2010. HbA1c values may also be used to predict the incidence of type 2 diabetes. However the diagnostic characteristics of HbA1c for the assessment of type 2 diabetes have been investigated by many authors. Many countries have not combined its use for this purpose yet and there is no consensus on a suitable cut-off point of HbA1c for the diagnosis of diabetes. The inclusion of HbA1c increases the feasibility and dissemination of DM screening because it eliminates the need for fasting.[18,19] However, the HbA1c criteria are relatively new, and there are still concerns, a cut-off point, including the standardized HbA1c measurement, different ethnic distributions and discordance with the glucose criteria.[20,21] However, prior to now the three-class nature of HbA1c has been ignored and evaluations based on a two-class case where the groups were classified either as a person with diabetes or a person who was normal. Since no studies have been conducted so far by taking into account the three-class nature of HbA1c, our study would probably the first in this area, which we believe makes contribution towards better interpreting the results.

In the present study, the effectiveness of HbA1c in diagnosing diabetes is investigated with respect to the outcome of the gold standard (i.e., previously accepted and validated) test obtained from FPG and OGTT.

The sensitivity and specificity of FPG, OGTT and HbA1c as screening tests vary according to the population tested and the threshold used to define type 2 DM. Using the same reference standard, the specificity and sensitivity of an HbA1c cut-off point of 6.5% were 79% and 44%, respectively.[22] Tatsch *et al*.[23] evaluated the diagnostic characteristics of HbA1c and FPG for the assessment of type 2 DM. Their results demonstrate the association between HbA1c and type 2 DM to be independent of age, gender, hypertension, smoking, and BMI. They recommended HbA1c as a suitable tool for the diagnosis of type 2 diabetes.[23] In a prospective cohort study of 26,563 women followed for 10 years, with a baseline HbA1c in the highest quintile (HbA1c >5.22%), the adjusted relative risk of diabetes was 8.2, 95% CI: 6.0-11.1. Zhang *et al*. estimate that a total of 13.1 million adult inhabitants have DM and that 40.0 million adult inhabitants are at risk of developing diabetes and cardiovascular disease. Their results suggest that the glucose and the HbA1c approaches should be used together for the screening and diagnosis of diabetes and prediabetes.[19] HbA1c criteria for identifying patients with impaired glucose regulation were derived using data from the National Health and Nutrition Survey, from 2005 to 2006. An HbA1c cut-off point of 5.7% had the best sensitivity (39%) and specificity (91%) for identifying cases of IFG. The ADA selected 5.7% to 6.4% to reflect an increased risk for diabetes.[24]

Hemoglobin A1c levels provide a simple, single-sample determination that is reflective of both fasting and postprandial glycemia. HbA1c testing with the proposed cut-off points displays a promising performance as indicated in this study. OGTT can be helpful in some clinical settings for diagnosis; however, it has some drawbacks in terms of its applicability and patient tolerability. Yet the determination of HbA1c is simple, inexpensive, can be performed on one single occasion, and does not require fasting. Blood intake is adequate, making it a more appropriate test for diagnosing diabetes.

The outcome of a diagnostic test may be influenced by some factors.[25-28] In other words, the outcome of a diagnostic test may differ depending on various characteristics of the patient including age, gender or a specific phase of the illness concerned. For example, in our study, three groups of subjects differed in some patient characteristics [Table 1]. This may be influential on test outcomes and it might be investigated whether or not this influence is significant. If this influence is significant, then it will be much more useful to conduct analyses with different sub-groups instead of giving a single performance measure and a set of cut-off points for the whole group. This approach would supply

more information when diagnosing subjects. It may be examined as to whether HbA1c is associated with patient characteristics and if so, what kind of association exists. Sub-groups may be identified where HbA1c is much more or less successful in determining the patient's status. Further studies on this issue will make a significant contribution to clinicians and researchers towards reaching a more elaborate understanding of the impact of common variables on the success of HbA1c in diagnosing diabetes. However, there is a need to work on many more observations for such assessments since higher numbers are required in sub-groups to conduct similar analyses in patient sub-groups. Fewer numbers of people in the group with diabetes may impose additional restrictions on sub-group analyses.

This study has some limitations, such as the number of patients with diabetes. Patients were evaluated according to the OGTT results by considering the risk groups, which resulted in a lower number of patients with diabetes. However, being the first study in this field to demonstrate the use of three-way ROC analysis, it offers valuable information for the evaluation of HbA1c. A high BMI in the healthy group led this group to be treated as being at risk for diabetes, and OGTT was ordered for such patients. These patients were included in the healthy group if their OGTT results were within the normal range.

CONCLUSION

Hemoglobin A1c has come to play an important role in the diagnosis of diabetes. We found that patients with HbA1c values lower than 5.2% can be diagnosed as "healthy", values from 5.2% to 6.4% as having "prediabetes," and a value >6.4% as having "diabetes". Since handling patients with different disease categories may involve some variation, it is important to determine with high confidence at which disease state a patient is. Our single center experience showed that the performance of HbA1c in distinguishing between prediabetes and diabetes groups is stronger than in distinguishing between healthy and prediabetes groups [Table 2].

According to the HbA1c test, misclassification of the subject with diabetes as prediabetes is more important than the misclassification of the healthy subject as prediabetes in terms of changing the frequency of follow-up and treatment strategy. Furthermore, if the subject with prediabetes was diagnosed as healthy, his clinical follow-up might be skipped. Therefore, if a subject was diagnosed as healthy or prediabetes by HbA1c, it would be beneficial to verify the status of that subject by the gold standard test (OGTT and FPG).

These findings need to be confirmed in larger studies. Validation of this study should be investigated by further research.

ACKNOWLEDGMENT

We acknowledge the assistance of Metin Odevci during the performance of HbA1c assays.

This study was supported by Hacettepe Unversity, Scientific Research Projects Coordination Unit (010 T02 101 001).

AUTHOR'S CONTRIBUTION

JK contributed in the conception of the work, statistical analysis of data, statistical interpretation of data for the work, revised the draft, approved the final version of the manuscript, agreed for all aspects of the work. SA clinical interpretation of data for the work, revised the draft, approved the final version of the manuscript, agreed for all aspects of the work. EK contributed in the conception of the work, statistical analysis of data, statistical interpretation of data for the work, revised the draft, approved the final version of the manuscript, agreed for all aspects of the work. AG clinical interpretation of the data for the work, revised the draft, approved the final version of the manuscript, agreed for all aspects of the work.

REFERENCES

1. Li H, Oldenburg B, Chamberlain C, O'Neil A, Xue B, Jolley D, *et al.* Diabetes prevalence and determinants in adults in China mainland from 2000 to 2010: A systematic review. Diabetes Res Clin Pract 2012;98:226-35.
2. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, *et al.* Management of hyperglycaemia in type 2 diabetes: A patient-centered approach. Position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia 2012;55:1577-96.
3. O'Sullivan CJ, Hynes N, Mahendran B, Andrews EJ, Avalos G, Tawfik S, *et al.* Haemoglobin A1c (HbA1C) in non-diabetic and diabetic vascular patients. Is HbA1C an independent risk factor and predictor of adverse outcome? Eur J Vasc Endovasc Surg 2006;32:188-97.
4. Bansilal S, Farkouh ME, Fuster V. Role of insulin resistance and hyperglycemia in the development of atherosclerosis. Am J Cardiol 2007;99:6B-14.
5. Obuchowski NA. Estimating and comparing diagnostic tests' accuracy when the gold standard is not binary. Acad Radiol 2005; 12:1198-204.
6. Li J, Zhou X. Nonparametric and semiparametric estimation of the three way receiver operating characteristic surface. J Stat Plan Inference 2009;139:4133-42.
7. Mossman D. Three-way ROCs. Med Decis Making 1999;19:78-89.
8. Dreiseitl S, Ohno-Machado L, Binder M. Comparing three-class diagnostic tests by three-way ROC analysis. Med Decis Making 2000;20:323-31.
9. Nakas CT, Yiannoutsos CT. Ordered multiple-class ROC analysis with continuous measurements. Stat Med 2004;23:3437-49.
10. Xiong C, van Belle G, Miller JP, Morris JC. Measuring and estimating diagnostic accuracy when there are three ordinal diagnostic groups. Stat Med 2006;25:1251-73.
11. Genuth S, Alberti KG, Bennett P, Buse J, Defronzo R, Kahn R, *et al.* Follow-up report on the diagnosis of diabetes mellitus. Diabetes Care 2003;26:3160-7.

12. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2011;4:S62-9.
13. Heckerling PS. Parametric three-way receiver operating characteristic surface analysis using mathematica. Med Decis Making 2001;21:409-17.
14. Nakas CT, Alonzo TA. ROC graphs for assessing the ability of a diagnostic marker to detect three disease classes with an umbrella ordering. Biometrics 2007;63:603-9.
15. Yang H, Carlin D. ROC surface: A generalization of ROC curve analysis. J Biopharm Stat 2000;10:183-96.
16. Dong T, Tian L, Hutson A, Xiong C. Parametric and non-parametric confidence intervals of the probability of identifying early disease stage given sensitivity to full disease and specificity with three ordinal diagnostic groups. Stat Med 2011;30:3532-45.
17. Luo J, Xiong C. DiagTest3Grp: An R Package for analyzing diagnostic tests with three ordinal groups. Journal of Statistical Software 2012;51:1-24. URL http://www.jstatsoft.org/v51/i03/.
18. Mann DM, Carson AP, Shimbo D, Fonseca V, Fox CS, Muntner P. Impact of A1C screening criterion on the diagnosis of pre-diabetes among U.S. adults. Diabetes Care 2010;33:2190-5.
19. Zhang YH, Ma WJ, Thomas GN, Xu YJ, Lao XQ, Xu XJ, *et al.* Diabetes and pre-diabetes as determined by glycated haemoglobin A1c and glucose levels in a developing southern Chinese population. PLoS One 2012;7:e37260.
20. Bang H, Edwards AM, Bomback AS, Ballantyne CM, Brillon D, Callahan MA, *et al.* Development and validation of a patient self-assessment score for diabetes risk. Ann Intern Med 2009;151:775-83.
21. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes Risk Calculator: A simple tool for detecting undiagnosed diabetes and pre-diabetes. Diabetes Care 2008;31:1040-5.
22. Kramer CK, Araneta MR, Barrett-Connor E. A1C and diabetes diagnosis: The Rancho Bernardo Study. Diabetes Care 2010;33:101-3.
23. Tatsch E, Bochi GV, Piva SJ, Pereira RS, Kober H, De Carvalho JA, *et al.* Hba(1c) as a tool for the diagnosis of type 2 diabetes: Comparison with fasting glucose. Clin Lab 2012;58:347-50.
24. Basevi V, Di Mario S, Morciano C, Nonino F, Magrini N. Comment on: American Diabetes Association. Standards of medical care in diabetes-2011. Diabetes Care 2011;34 Suppl 1:S11-61.
25. Goehring C, Perrier A, Morabia A. Spectrum bias: A quantitative and graphical analysis of the variability of medical diagnostic test performance. Stat Med 2004;23:125-35.
26. Hlatky MA, Pryor DB, Harrell FE Jr, Califf RM, Mark DB, Rosati RA. Factors affecting sensitivity and specificity of exercise electrocardiography. Multivariable analysis. Am J Med 1984;77:64-71.
27. Karakaya J, Aksoy DY, Harmanci A, Karaagaoglu E, Gurlek A. Predictive ability of fasting plasma glucose for a diabetic 2-h postload glucose value in oral glucose tolerance test: Spectrum effect. J Diabetes Complications 2007;21:300-5.
28. Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. Ann Intern Med 2002;137:598-602.

Source of Support: Hacettepe Unversity, Scientific Research Projects Coordination Unit (010 T02 101 001). **Conflict of Interest:** None declared.