

# Retinopathy risk factors in type II diabetic patients using factor analysis and discriminant analysis

Mahdi Tazhibi, Sheida Sarrafzade, Masoud Amini<sup>1</sup>

Department of Biostatistics and Epidemiology, School of Health, <sup>1</sup>Endocrine and Metabolism Research Center, Isfahan University of Medical Sciences, Isfahan, Iran

## ABSTRACT

**Introduction:** Diabetes is one of the most common chronic diseases in the world. Incidence and prevalence of diabetes are increasing in developing countries as well as in Iran. Retinopathy is the most common chronic disorder in diabetic patients. **Materials and Methods:** In this study, we used the information of diabetic patients' reports that refer to endocrine and metabolism research center of Isfahan University of Medical Sciences to determine diabetic retinopathy risk factors. We used factor analysis to extract retinopathy's factors. Factor analysis is using to analyze multivariate data, in which a large number of dependent variables summarize into the fewer independent factors. Factor analysis is applied, in both diabetic and nondiabetic patients, separately. To investigate the efficacy of factor analysis, we used discriminant analysis. **Results:** We investigated 3535 diabetic patients whose prevalence of retinopathy was 53.4%. Six factors were extracted in each group (i.e. diabetic and nondiabetic groups). These six factors were explained 69.5% and 69.6% of total variance in diabetic and nondiabetic groups, respectively. Using original variables such as sex, weight, blood sugar control method, and some laboratory variables, the correct classification rate of discriminant analysis was identified as 67.4%. However, it decreased to 49.5% by using extracted factors. **Discussion:** Retinopathy is one of the important disorders in diabetic patients that involves a large number of variables and can affect its incidence. By the method of factor analysis, we summarize diabetic retinopathy risk factors. Factor analysis is applied separately, in two diabetic and nondiabetic group. In this way, 10 variables were summarized into the six factors. Discriminant analysis was used to investigate the efficacy of factor analysis. **Conclusion:** Although factor analysis is a powerful way to reduce the number of variables, in this study did not worked very well.

**Key words:** Factor analysis, retinopathy, type II diabetic

## INTRODUCTION

All over the world diabetes is one of the most common chronic diseases, which can affect people with each ethnicity. Specifically in all societies, developing countries, incidence and prevalence of diabetes are increasing.<sup>[1-3]</sup> Body's disability to produce and use insulin can cause diabetes mellitus. This syndrome occurs because of imbalance between demand and supply of insulin.<sup>[4]</sup> In fact, diabetics type II is the result of environmental factors, such as lifestyle and genetic conditions.<sup>[2]</sup> Nowadays, diabetes is known as an epidemic disease in many developed countries.<sup>[5]</sup> According to the WHO, based on epidemiologic studies in 75 regions in 32 UN member countries, which published in 1998, world's

**Address for correspondence:** Miss. Sheida Sarrafzade, Department of Biostatistics and Epidemiology, School of Health, Isfahan University of Medical Sciences, Isfahan, Iran.  
E-mail: shsarafzade@gmail.com

Access this article online	
Quick Response Code:	Website: www.jehp.net
	DOI: 10.4103/2277-9531.139251

Copyright: © 2014 Tazhibi M. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This article may be cited as: Tazhibi M, Sarrafzade S, Amini M. Retinopathy risk factors in type II diabetic patients using factor analysis and discriminant analysis. J Edu Health Promot 2014;3:85.

adult population who are greater than 20 years old from 1995 until 2025 will increase by 64%. Diabetic's prevalence which was 4% in 1995 will increase to 5.4% in 2025, and the total number of diabetic patients will increase by 122%. This means that the number of patients from 135 million in 1995 will be 300 million in 2025.<sup>[4]</sup> In Iran, diabetes is one of the most common diseases too, and its incidence rate is increasing day-by-day. In 1994, the total number of diabetic patients was almost 1,240,000 in Iran, which was 1.5 million in 1998. In 2007, about 4 million diabetic patients existed in Iran.<sup>[4]</sup> Retinopathy is one of the diabetic disorders.<sup>[3,6]</sup> Diabetic retinopathy is one of the most common and the most important disorders in diabetic patients, and risk of blindness in diabetics is about 25 times higher than nondiabetics.<sup>[3,7,8]</sup> In developed countries, the main cause of blindness in adults under 75 years is diabetic-related eye disease.<sup>[9]</sup> Retinopathy is the main cause of blindness in adults between 20 and 64, and more than 77% of diabetic patients will be affected by retinopathy 20 years after the disease onset.<sup>[10,11]</sup> In a study in Japan, diabetic mellitus was the main cause of blindness.<sup>[12]</sup> In addition, it was the most common cause of blindness in adults between 25 and 74 in America.<sup>[3,7,8]</sup> Diabetic retinopathy is a common disorder in Iran as well.<sup>[8]</sup>

Variety of factors can affect diabetic retinopathy promotion such as: Type of diabetic, duration of the disease, age, sex, blood sugar control status, high blood pressure (hypertension), cigarette smoking, high serum lipids and microalbuminuria,<sup>[9]</sup> fast blood sugar (FBS), cholesterol, triglyceride, urea, creatinine, and HBA1c.<sup>[8]</sup> The purpose of this study is to determine the retinopathy prevalence in diabetic patients and its relationships between risk factors of diabetic retinopathy and also summarize those factors in fewer number of factors by use of factor analysis.

## MATERIALS AND METHODS

In this retrospective study, we extracted the information from diabetic patients' files were referred to Isfahan Metabolism and Endocrine Research Center (MERC), to determine diabetic retinopathy risk factors. We collected the information of all diabetic patients who were referred to this center between 1992 and 2003. Among the 3535 registered diabetic patients, 1887 patients had diabetic retinopathy. All measurements were performed by experts, and the diagnosis of retinopathy was done by ophthalmologist in regular eye examinations and was recorded in patient's file. In addition to patient characteristics, time of diagnosis, and cigarette smoking, weight and height were also measured and recorded. Weight was measured in light clothing and without shoes, and height was measured in standing position without shoes. Since this center was supervised by Isfahan University of Medical Sciences (IUMS), and also to reduce the measurement error, the calibration of the scales was examined all the time. To decrease measurement error of recording, those people who were involved in filling questionnaires were trained.

However, as this was a retrospective study, we were not sure about scales calibration and unique performance among all

measurements. Patients' blood pressure after resting, 15 min sitting, from right arm had been measured by the standard mercury sphygmomanometer. It had been measured two times with 5 min delay and the average of these two measures had been recorded as patients' blood pressure. All patients had been investigated for FBS, blood sugar after meals, glycoside hemoglobin, blood lipids, and albuminuria. Each case which has not at least one of these measures was omitted.

## Statistical methods

In this study, first we used descriptive analysis to determine retinopathy prevalence in type II diabetic patients. Then, we used factor analysis to summarize diabetic retinopathy risk factors. Factor analysis is a statistical method for multivariate data. This method summarizes a large number of dependent variables into fewer independent factors. Factors obtained in this method are hidden, and they are not directly observable.<sup>[11,13,14]</sup> In studies which contain a large number of correlated variables, factor analysis can be used to reduce variable dimension.<sup>[11,13-15]</sup>

To perform factor analysis, diabetic patients were divided into two groups, depending on whether or not they had retinopathy. Then, each group was analyzed individually by using factor analysis. Factors were extracted by the principal component method. Factors with eigenvalues larger than 1 were accepted. To increase factors interpretation, varimax rotation was implemented. LDL, HDL, and HBA1C had not measured for two-third of cases, so these three variables were excluded from this study. In addition, logarithmic transformation was implemented for nonnormal variables.

To determine factor efficiency, Discriminant analysis was implemented. Discriminant analysis is widely used for case classification.<sup>[16,17]</sup> The main purpose of discriminant analysis is to determine the relative relationship between variables in group separation, and finding a role to classify new cases to one group with least error, according to the observed cases.<sup>[12,14,15]</sup>

SPSS software (version 13) was used for all analysis.

## RESULTS

In this study, 3535 patients were investigated, where 1887 cases (53.4%) had retinopathy. A total of 1261 cases (35.7%) and 2271 cases (64.2%) were men and women, respectively. Further, 696 (19.7%), 2245 (63.5%), and 531 (15%) of patients were controlled their blood sugar by using diet, oral medication, and insulin injections, respectively. Sixty-three cases (1.8%) did not answer to this question. Of them, 385 subjects (10.9%) were smokers. In this study, retinopathy prevalence in diabetic patients was 53.4%. According to the contingency table and chi square statistic, gender, smoking, and the blood sugar control method had significant effect on retinopathy development.

According to an independent *t*-test, the mean age, BMI, systolic blood pressure, disease duration, weight, BUN, and FBS were significantly different between two retinopathy and

nonretinopathy groups. However, the average of cholesterol, triglyceride, diastolic blood pressure, and the patient's age of at the beginning of the disease did not significantly different between the two groups. Patient's characteristics are shown in Table 1.

First of all, normality of all variables has been checked. To investigate variable's normality, the Q-Q normal plots of all variables were plotted, logarithmic transformation was used for skewed variables. Then, to determine the efficiency of logarithmic transformation, the Q-Q normal plots of transformed variables were plotted and it was observed that all skewed variables were normalized by this transformation. Factor analysis was investigated in two retinopathy and nonretinopathy groups separately. The principal components method was used to factors extraction. Further, varimax orthogonal rotation has been used to increase factors interpretability. Six factors were extracted from factor analysis for each group. These six factors explained 69.5% and 69.6% of total variance in retinopathy and nonretinopathy groups, respectively. In the retinopathy group, age and age at illness onset were in factor 1, weight and BMI were in factor 2, systolic and diastolic blood pressure were in factor 3, sex and smoking were in factor 4, the blood sugar control method and disease duration were in factor 5, and cholesterol and triglyceride were in factor 6. In the nonretinopathy group, age and age at illness onset were in factor 1, weight and BMI were in factor 2, systolic and diastolic blood pressure were in factor 3, sex and smoking were in factor 4 as well as retinopathy group, but cholesterol and triglyceride were in factor 5, and blood sugar control method, disease duration, and FBS were in factor 6. These results are briefly shown in Table 2.

Discriminant analysis was used to investigate the efficiency of extracted factors. In order to do this, discriminant analysis was first performed for original variables. According to the Wilks lambda test, discriminant function was significant,  $P < 0.001$ . The percentage of correct classification was 67.4%. Then, discriminant analysis was performed for extracted factors. The  $P$  value for the Wilks lambda test was 1 for the latest analysis. It means that factors did not have a significant role to group separation. The percentage of correct classification was 49.5% for this analysis, which is not a suitable rate.

**Table 1: Participants demographic information**

Variable	Mean	Standard deviation
Age	52.57	10.28
Age at illness onset	45.46	10.45
Disease duration	7.11	6.68
BMI	27.38	4.46
Weight	68.68	11.65
FBS	203.70	73.99
BUN	25.49	9.74
Cholesterol	226.74	52.19
Triglycerides	228.49	155.60
Systolic blood pressure	123.52	19.41
Diastolic blood pressure	76.20	10.75

FBS = Fast blood sugar, BUN = Blood urea nitrogen, BMI = Body mass index

## DISCUSSION

Retinopathy is one of the chronic and common disorders of diabetic and is the main cause of blindness in adult population. Many factors affect the development of diabetic retinopathy. In this study, in addition to determine retinopathy prevalence in diabetic patients, factor analysis has been done to summarize risk factors of retinopathy.

The purpose of factor analysis is to extract hidden variables. When data set includes a lot of variables, factor analysis can classify variables for easier interpretation. In this study, exploratory factor analysis can be implemented to select retinopathy risk factors in diabetic patients. Using factor analysis, additional variables can be considered, and then by extracting less number of factors, the risk of developing retinopathy in diabetics can be evaluated with precision. Here, instead of entering 12 variables, six obtained factors can be used in model.

Prevalence of retinopathy in diabetic patients which referred to MERC was about 53.4%. Askarishahi and colleagues in their research mentioned that prevalence of retinopathy in diabetic patients was about 40.7%.<sup>[6]</sup> In other study, Rafati *et al.* declared that on 634 diabetic patients in Tehran, 240 patients (37.9%) had retinopathy.<sup>[9]</sup> However, in a research by Malek-madani, on 84 diabetic patients, retinopathy prevalence was obtained 77.4%.<sup>[8]</sup> It is possible that the high prevalence observed in Malek-madani's study is because of its investigation of diabetic patients, which had been involved all patients hospitalized for complications due to diabetes and expect retinopathy.

In our study, variables such as sex, age, BMI, systolic blood pressure, cigarette smoking, blood sugar control method, disease duration, weight, BUN, and FBS had significant differences in the two retinopathy and nonretinopathy groups.

In UK Prospective Diabetes Study (UKPDS) study, the strong relationship had been observed between retinopathy and hypertension, FBS, and smoking.<sup>[18]</sup> In the study by Chatziralli *et al.*, age, period of disease, and hypertension history had positive relation with diabetic patients.<sup>[10]</sup>

In the survey conducted by Askarishahi *et al.*, significant association had been observed between duration of diabetes and blood sugar control method with retinopathy.<sup>[6]</sup> On the basis of the results declared by Rafati and their colleagues, significant association between retinopathy and sex, period of disease, and the blood sugar control method had been observed.<sup>[9]</sup>

In this study, the principal components method and Varimax orthogonal rotation have been used to extract factors. All subjects divided into two separate, retinopathy or nonretinopathy groups and factor analysis have been done separately for each group. Shmulewitz *et al.*, implemented

**Table 2: Extracted factors separated by groups**

Variable	Retinopathy group						Nonretinopathy group					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Age	0.91						0.90					
Age at disease onset	0.93						0.96					
Log (weight)		0.88						0.91				
Log (BMI)		0.89						0.90				
Systolic BP			0.91						0.90			
Diastolic BP			0.87						0.86			
Sex				-0.83						-0.84		
smoking				0.77						0.78		
Triglyceride						0.73					0.80	
Cholesterol						0.78					0.80	
Blood sugar control method					-0.75							-0.77
Log (disease duration)					0.83							0.71
Log (FBS)												0.50
Percentage of variance proportion	13.09	12.56	12.18	10.82	10.66	10.23	13.39	12.52	12.16	10.84	10.63	10.11
Cumulative percentage of variance proportion	13.09	25.65	37.83	48.65	59.31	69.54	13.39	25.91	38.07	48.91	59.54	69.65

BP = Blood pressure, FBS = Fast blood sugar, BMI = Body mass index

the principal component method to extract factors with the method of Varimax rotation.<sup>[19]</sup> In other study in China by Wang, factors extracted by the principal component method and Varimax orthogonal rotation used to classify risk factors of metabolic syndrome. They divided subjects into two diabetic and nondiabetic groups, and factor analysis was conducted separately for two groups. In addition, Wang *et al.* have been implemented logarithmic transformation to skewed variables.<sup>[11]</sup>

In our study, extracted factors in each retinopathy and nonretinopathy group explained almost 70% of total variation. Hanley *et al.* reported that the principal component method and Varimax rotation had been used to extract factors. Two extracted factors were explained almost 40% of total variation.<sup>[13]</sup>

Correct classification percent obtained by discriminant analysis on original variables and extracted factors were 67.4% and 49.5%, respectively. Discriminant function based on extracted factors was not significant.

## CONCLUSION

Here we used factor analysis to reduce number variables. Although factor analysis is a powerful way to reduce the number of variables, in this study did not worked very well. It is possible that extracted factors were not suitable for discriminant analysis.

## REFERENCES

- Larejani B, Zahedi F. Epidemiology of diabetes mellitus in Iran. *Iran J Diabetes Lipid Disord* 2001;1:1-8.
- Emini-Sadiku M, Car N, Metelko E, Bajraktari G, Morina N, Devolli D. Prevention or delay of type 2 diabetes by pharmacological or lifestyle interventions. *Diabetologia Croatica* 2008;37:23-30.
- Manaviat M, Afkhami-Ardakani M, Rashidi M. Study of four year incidence rate of diabetic retinopathy and Its associated factors in type II diabetes patients. *Journal of Shahid Sadoughi University of Medical Sciences* 2006;14:11-9.
- Haririan H, Moghadasian S, Aghajani A. Quality of life and it's aspects in diabetic patients referred to the Diabetes Center of Tabriz University of Medical Sciences in 2007. *Iranian Journal of Diabetes and Lipid Disorders*. Winter 2009;92:152-60.
- Aghamollaei T, Eftekhari H, Shojaeizadeh D, Mohammad K, Nakhjavani M, Ghofrani Pour F. Behavior, metabolic control and health-related quality of life in diabetic patients at Bandar Abbas diabetic clinic. *Iran J Public Health* 2003;32:54-9.
- Askarishahi M, Hajizadeh E, Afkhami-Ardakani M. Factors affecting retinopathy in patients with type 2 diabetes by analyzing the current status data. *Tehran Univ Med J* 2011;68:674-80.
- Akbarzadeh S, Mani-kashani Kh. Evaluation of types of diabetic retinopathies and associated factors in diabetic patients. *Medical journal of Tabriz university of medical sciences and health services*. Summer 2006;28:15-8.
- Malek-madani M, Lashiyi A, Behjati-ardakani J, Ganji R. Determination of frequency and severity of diabetic retinopathy in patients hospitalized due to complications of diabetes Ghynchshmy. *Tehran University Meridal Journal* 2007;65:58-63.
- Rafati N, Dehghan M, Javadi M, Soori H, Zaeri F, Gooya M, *et al.* The prevalence of diabetic retinopathy in Tehran Province. *Journal of Ophthalmology Bina*. Winter 2007;13:130.
- Chatziralli IP, Sergentanis TN, Keryttopoulos P, Vatakis N, Agorastos A, Papazisis L. Risk factors associated with diabetic retinopathy in patients with diabetes mellitus type 2. *BMC Res Notes* 2010;3:153.
- Wang JJ, Qiao Q, Miettinen ME, Lappalainen J, Hu G, Tuomilehto J. The metabolic syndrome defined by factor analysis and incident type 2 diabetes in a Chinese population with high postprandial glucose. *Diabetes Care* 2004;27:2429-37.
- Lix LM, Sajobi TT. Discriminant analysis for repeated measures data: A review. *Front Psychol* 2010;1:146.
- Hanley AJ, Karter AJ, Festa A, D'Agostino R, Wagenknecht LE, Savage P, *et al.* Factor Analysis of Metabolic Syndrome Using Directly Measured Insulin Sensitivity The Insulin Resistance Atherosclerosis Study. *Diabetes* 2002;51:2642-7.
- Raykov T, Marcoulides GA. An introduction to applied multivariate analysis. London: Routledge/Psych press; 2008.

15. Rencher AC. *Methods of multivariate analysis*. New York: Publisher Wiley; 1995.
16. Balakrishnama S, Ganapathiraju A. *Linear discriminant analysis-a brief tutorial*. Institute for Signal and information Processing. Department of Electrical and Computer Engineering Mississippi State University 1998.
17. Guo Y, Hastie T, Tibshirani R. Regularized linear discriminant analysis and its application in microarrays. *Biostatistics* 2007;8:86-100.
18. Stratton I, Kohner E, Aldington S, Turner R, Holman R, Manley S, *et al.* UKPDS 50: Risk factors for incidence and progression of retinopathy in Type II diabetes over 6 years from diagnosis. *Diabetologia* 2001;44:156-63.
19. Shmulewitz D, Auerbach SB, Lehner T, Blundell ML, Winick JD, Youngman LD, *et al.* Epidemiology and factor analysis of obesity, type II diabetes, hypertension, and dyslipidemia (syndrome X) on the Island of Kosrae, Federated States of Micronesia. *Hum Hered* 2001;51:8-19.

**Source of Support:** This study was supported by Vice-chancellery for Research, Isfahan University of Medical Sciences, Isfahan, Iran. Project Number 290272, **Conflict of Interest:** None declared