Correlation of risk factors for coronary heart disease in type-2 diabetes mellitus using the chi-square method

L Syafaah*, D Arisnanda, M Irfan, and R H Hendaryati
Department of Informatics Engineering, Faculty of Engineering, Universitas Muhammadiyah Malang, Malang, Indonesia

*Email: lailis@umm.ac.id

Abstract. The male sex is taken into account a crucial risk issue for coronary heart disease (CHD) within the general population, however, studies on diabetic patients counsel that diabetic girls have a better risk of developing CHD. This study was to determine whether there is correlation between risk factors and the incidence of CHD in patients with type 2 DM. Furthermore, this work also investigates the mathematical model work in calculating the magnitude of the risk of variables against the tendency of CHD complications for patients with type 2 diabetes. This research utilizes current information on the setup of Dr. Soetomo Hospital's medical records in 2015-2016. Samples were all patients with type 2 diabetes with CHD and without CHD reported in the inpatient medical records of Dr. Soetomo Hospital Surabaya. Data processing was carried out using the chi-square test technique. As a result, the system is proof to calculate the worth of the chi-square and confirm the results of conditions on every variable. There is a distinction in price between the calculations of the chi-square on the system with the calculation of the manual chi-square test. This is often caused by taking variety of digits behind completely different commas (.). The variables of gender, age, GDA, GDP, GD2PP, triglycerides, HDL, weight, smoking, and length failed to have a significant relationship on the incidence of CHD in patients with sort two DM. Sterol variables and lipoprotein levels have a highly vital relationship within the incidence of CHD for patients with sort two polygenic disease.

1. Introduction
Diabetes mellitus (DM) is currently one of the medical problems that affect efficiency and reduces the quality of human resources. DM patients worldwide in 2015 increased from 333 million people (5.4%). People with diabetes in Indonesia are continuing to occur. Considering the record of the world health organization in 2004, Indonesia ranked fourth with the highest number of diabetics after India, China, and USA [1]. It was realized that in 1995 there were around 5 million people with DM in Indonesia with a growth of around 230 thousand diabetics consistently so that by 2025 it was estimated that there would be 12 million people. The increase occurred due to expansion in older populations and changes in lifestyle, starting from eating habits, types of food, to decreased physical activity. This happens especially in the category of adults or more for all strata of society. In addition, developments in the number of DM cases occur due to the absence of workers, checking specific equipment and medicines, especially in remote zones, and there is no consistency in supervising DM patients.
The ten countries estimated to have the highest numbers of people with diabetes in 2000 and 2030 are reported in reference [1]. The “top five” countries are the same as those identified for 1995 (India, China, U.S., Indonesia, and Japan). Bangladesh, Brazil, Russia, and Pakistan also appear in the lists for both 2000 and 2030. The similarity of DM arises all over the world, especially in Asian countries. Calculations in 2020 of this disease refer to the death of 7 individuals from each of 10 individuals in developing countries. About 16 million people in America are diagnosed with diabetes, the dominant being 6% to 7% in individuals 45 to 65 years. About 90% of them experience the adverse effects of type 2 diabetes. About 9.7 million women in America have diabetes. Type 2 diabetes occurs at all ages, even at an early age and teenagers.

Type 2 diabetes is the most well-known type of diabetes found in DM patients than type 1 diabetes mellitus, gestational diabetes, and various types of diabetes [2]. Type 2 diabetes is a multifactorial disease with the same genetic and ecological parts as its effect to cause disease. The impact of heredity on this disease can be seen clearly with the high level of diabetics from fathers or mothers who have a profile of diabetes mellitus. Type 2 diabetes mellitus is also often called diabetes as a way of life, considering the causes other than heredity, ecological components including age, weight, insulin resistance, daily consumption, physical activity, and poor individual way of life also influence diabetes. Type 2 diabetes mellitus which is not monitored properly will cause a variety of chronic complications [3], for example, retinopathy and nephropathy and macroangiopathy, for example, heart disease, stroke, and further additional blood vessel infections.

The main driver of mortality and morbidity in patients with type 2 DM is coronary heart disease (CHD) [4, 5]. As demonstrated by the American Heart Association in May 2012, at least 65% of people with diabetes die of heart disease or stroke. In addition, adults who are poor diabetics have a risk of coronary heart disease two to several times higher than individuals who are not diabetics [6]. A slow increase in type 2 diabetes mellitus often causes side effects and unclear signs. The component of type 2 CHD events is complicated and is associated with atherosclerosis which is influenced by various elements, including hypertension, hyperglycemia, dyslipidemia, smoking, CHD family offspring, and body weight.

Type 1 diabetes mellitus is a metabolic problem caused by an immune system response that makes pancreatic beta cells initiated by prolonged hyperglycemia due to the absence of substantial insulin [7]. During the range of DM disease can cause different complications, namely short-term and long-term complications. Short-term complications, namely hypoglycemia and ketoacidosis. Long-term complications occur due to microvascular changes such as retinopathy, nephropathy, and neuropathy. There are 2 types of DM, namely specific type 1 DM which is insulin subordinate diabetes, where the pancreas gives little insulin or does not create insulin at all. While in type 2 DM, the pancreas still gives insulin, but in some cases, the amount is higher than the normal limit, where this event will make the body's structure insensitive to its effect, which causes relative insulin deficiency.

Type 2 DM is a hyperglycemic disease due to a lack of insulin sensitivity of cells. Insulin levels can be reduced slightly or be within reasonable limits. Because insulin is still made by pancreatic beta cells, type 2 diabetes mellitus is considered a non-insulin subordinate diabetes mellitus [8]. Type 2 diabetes is not caused by the absence of insulin emissions, but because insulin target cells become flat or unable to react normally. This condition is called insulin resistance. Variables associated with diabetes risk are those who have polycystic ovary syndrome, patients with metabolic disorders, backgrounds that are characterized by inhibited glucose resistance or fasting blood glucose, and who have cardiovascular backgrounds. According to [9], there was a positive correlation obese patients. LITAF is a TNF-α description of the between obesity and blood glucose levels.

In previous study on the reference [10], employs the Spline Interpolation technique to convert discrete blood glucose data to continuous signal. The validation of interpolated signal is conducted by comparing the pattern of discrete data and continuous signal for both original and clustered data. The experiment showed that both scenarios depicted identical patterns. The Spline interpolation technique indicates that the method is the most appropriate in predicting the interpolated points on blood glucose data. This paper describes a correlation between risk factors of coronary heart disease in type 2 DM
patients using Chi-Square method. This method was applied to determine whether there is correlation between risk factors and the incidence of CHD in patients with type 2 DM. Furthermore, this work also investigate how does the mathematical model work in calculating the magnitude of the risk of variables against the tendency of CHD complications for patients with type 2 diabetes.

2. Methodology

This study uses existing data on the medical record installation of Dr. Soetomo Hospital in 2015-2016. Samples were all patients with type 2 diabetes with CHD and without CHD recorded in the medical records of the inpatient disease in Dr. Soetomo Hospital Surabaya. Data processing was performed using the chi-square test method. The purpose of processing data using the chi-square test is to find the chi-square test value of each CHD risk factor variable, and then the results will be in the form of a graph that will illustrate the high or low results of the chi-square test value. From the results obtained, it will be seen which risk factors are most influential on the occurrence of CHD in patients with type 2 DM at Dr. Soetomo Hospital in Surabaya in 2015-2016.

Data using patients totaling 50 people along with information from 12 CHD risk factors for type 2 diabetes. The risk factors in question are age, gender, GDA, GDP, GD2PP, triglycerides, cholesterol, HDL, LDL, body weight, smoking, and long DM.

Chi-Square is also called Kai Squared. Chi-square is a type of non-parametric similar test carried out on two factors, where the size of the information from both factors is in the form of values [11]. In parametric insight, the proportion of relationships that are normally used is Pearson’s product moment relationship. Among the nonparametric relationships that can be compared with these standard connection coefficients and those commonly used are the Spearman R, Kendal Tau and gamma coefficients. Besides these three methods, chi-square depending on the cross table is also quite well known in estimating the relationship between factors.

Chi-square test technique depends on this cross table is organizing variables in the classification and testing of hypotheses that the observed iteration is the same as the expected iteration [6]. The integrity of the chi-square compatibility test is about the observed repetition and the expected repetition in each class, it must be tested that all groups contain the same value or a test that each group contains a certain value.

The assumption used is an example derived from a random sample. The expected frequency for each group must be more than 1. The expected frequency of values below 5 must not be more than 20% of the group. The chi-square calculation can be expressed in Equation (1).

\[ x^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \]  

where \( r \) is number of rows, \( c \) is number of column, \( i \) is indicates of line to- \( i \) and \( j \) is column to- \( j \). While \( O_{ij} \) is observed frequencies in line to \( i \) and column \( j \). And \( E_{ij} \) is expected frequencies in line to \( i \) and column \( j \).

3. Result and discussions

Based on the results on Table 1, it can be drawn condition decisions based on the results of the chi-square test value of the DF value of each risk factor. If the chi-square test value < DF value, then the H0 logic is accepted, which means there is no relationship between these risk factors with the occurrence of CHD in patients with type 2 DM. If the chi-square test value > DF value, then the H1 logic is accepted, which means there is a significant relationship between these risk factors with the occurrence of CHD in patients with type 2 DM. From the value of the chi-square test in Table 1, A Chi-square test indicated that cholesterol and LDL have a significant relationship in the incidence of CHD in patients with type 2 DM, while gender, age, GDA, GDP, GD2PP, triglycerides, HDL, obesity, smoking, and for a long time DM did not have a significant relationship in the incidence of CHD in patients with type 2 DM in the hospital, Dr. Soetomo Surabaya in 2015-2016.
To find out the level of accuracy of the system that has been made, can calculate with a simple accuracy formula in Equation (2), which is to add up the total value of the chi-square test system, divided by the sum of all calculated chi-square test values. Table 1 shows the comparison result of Chi-Square Test and Measured Chi-Square.

\[
\text{Accuracy} = \frac{\text{number of Chi Square test}}{\text{number of measured Chi Square}} \times 100\% \\
= \frac{30.247}{30.753} \times 100\% = 98.354\% 
\]

Table 1. Comparison result of Chi Square Test and measured Chi Square

| Risk Factor Variables | Chi Square Measured | Chi Square Test | DoF value |
|----------------------|---------------------|----------------|-----------|
| Genders              | 0.673               | 0.673          | 3.841     |
| Ages                 | 3.165               | 3.171          | 7.814     |
| GDA                  | 1.188               | 0.855          | 3.841     |
| GDP                  | 0.608               | 0.421          | 3.841     |
| GD2PP                | 1.235               | 1.236          | 3.841     |
| TRIGLISERIDA         | 0.178               | 0.17           | 3.841     |
| Cholesterol          | 3.846               | 3.848          | 3.841     |
| HDL                  | 0.169               | 0.17           | 3.841     |
| LDL                  | 11.455              | 11.459         | 3.841     |
| Obesities            | 0.478               | 0.48           | 3.841     |
| Smoking              | 2.802               | 2.805          | 3.841     |
| Time of DM           | 4.956               | 4.959          | 5.991     |

Based on data testing that has been done using 50 patient data, it can be drawn a hypothesis by comparing the value of the chi-square test with the DF value of each risk factor. The conclusions that can be drawn from the mentioned experiments in which risk factors that influence the incidence of CHD in patients with type 2 DM are LDL and cholesterol because the value of the calculated chi-square test and chi-square test system is greater than the DF value each. Therefore the H1 condition applies. LDL risk factors have the highest value among other risk factors. Risk factors for gender, age, GDA, GDP, GD2PP, triglycerides, HDL, obesity, smoking, and duration of DM indicate that they do not have a significant relationship to the incidence of CHD in patients with type 2 DM, because of the calculated chi-square test or chi test -square- system is smaller than each DF value. Therefore the H0 condition applies. There is a difference in value for several risk factors between the calculated chi-square test and the chi-square test system. This is because the retrieval value behind the comma (,) in the system is more than the manual calculation.

4. Conclusions

After designing and testing this web application, it can be concluded that the system is able to calculate the value of the chi-square test and determine the results of conditions on each variable. There is a difference in value between the calculation of the chi-square test system with the calculation of the manual chi-square test, this is caused by taking a number of digits behind different commas (,). The variables of gender, age, GDA, GDP, GD2PP, triglycerides, HDL, body weight, smoking, and duration did not have a significant relationship on the incidence of CHD in patients with type 2 DM. Cholesterol variables and LDL levels have a very significant relationship in the incidence of CHD for patients with type 2 diabetes.
Acknowledgement
The authors wish to express their gratitude to the Faculty of Engineering, Universitas Muhammadiyah Malang for supporting this study. To the Dr. Soetomo hospital for providing the research data, the cooperation is highly appreciated.

References
[1] Wild S, Roglic G, Green A, Sicree R, and King H 2004 Global prevalence of diabetes estimates estimates for the year 2000 and projections for 2030 Diabetes Care vol 27 pp 1047–53.
[2] Islam R and Rahman O 2012 The risk factors of type 2 diabetic patients attending Rsajshahi Diabetes Association, Rajshahi, Bangladesh and its primary prevention Food Public Heal. vol 2 no 2 pp 5–11.
[3] Li Y, Li H, and Yao H 2018 Analysis and study of diabetes follow-up data using a data-mining-based approach in new urban area of Urumqi, Xinjiang, China, 2016-2017 Comput. Math. Methods Med. vol 2018.
[4] Yuliani F, Oenzil F, and Iryani D 2014 Hubungan berbagai faktor risiko terhadap kejadian penyakit jantung koroner pada penderita diabetes mellitus tipe 2 J. Kesehat. Andalas vol 3 no 1 pp 37–40.
[5] Chiha M, Njeim M, and Chedrawy E G 2012 Diabetes and coronary heart disease: A risk factor for the global epidemic Int. J. Hypertens. vol 2012.
[6] Langari S H, Bahar A, Asadian L, Abediankenai S, Namazi S S, and Kashi Z 2019 Coronary heart disease and ABO blood group in diabetic women: A case-control study Sci. Rep. vol 9 no 1 pp 5–10.
[7] Halilovic E A, Ljaljevic S, Alimanovic I, Mavija M, Oros A, and Nisic F 2015 Analysis of the influence of type of diabetes mellitus on the development and type of glaucoma Med. Arch. (Sarajevo, Bosnia Herzegovina) vol 69 no 1 pp 34–7.
[8] Betteng R, Pangemanan D, and Mayulu N 2014 Analisis faktor resiko penyebab terjadinya diabetes melitus tipe 2 pada wanita usia produktif Dipuskesmas Wawonasa J. e-Biomedik vol 2 no 2.
[9] Silitonga H A, Siahaan J M, and Anto E J 2019 Correlation between obesity and lipid profile in type 2 diabetes mellitus patients at the endocrine and metabolic polyclinic in general hospital pirngadi Medan Open Access Maced. J. Med. Sci. vol 7 no 8 pp 1309–13.
[10] Syaf’a’ah L, Purnomo M H, and Basuki S 2018 Discrete mean amplitude of glycemic excursion (MAGE) measurement on diabetics with spline interpolation method Int. J. Electr. Eng. Informatics vol 10 no 2 pp 259–70.
[11] Jousheghany F, Phelps J, Crook T, and Hakkak R 2016 Relationship between level of HbA1C and breast cancer BBA Clin. vol 6 pp 45–8.