A Mathematical Model for The Epidemiology of Diabetes Mellitus with Lifestyle and Genetic Factors

Purnami Widyaningsih¹, Rifqi Choiril Affan², Dewi Reno Sari Saputro³

¹Department of Mathematics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia
²Department of Mathematics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia
³Department of Mathematics, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

Abstract: Mathematical model to elaborate the prevalence of diabetics has been determined by diabetes complication (DC) model. In the DC model, people with diabetes were classified into two compartments, uncomplicated diabetics (D) and complicated diabetics (C). Diabetes is known as a disease caused by lifestyle and genetic factors. A bad lifestyle leads a susceptible individual to become a diabetic. Bad lifestyle is strongly influenced by risky social interaction. In the other side, a genetic factor is the main cause of the diabetes genetic disorder birth. Consider these both factors, the DC model was developed into a susceptible diabetes complication (SDC) model. Susceptible individuals were involved in the calculation of risky interactions. The SDC model is a first order nonlinear differential equation. The number and the change of individuals in each compartment can be determined from the solution of this model. In this paper, the SDC model is applied to predict changes of diabetics prevalence in the United States. As a result, the SDC model is good enough to predict the prevalence.

1. Introduction
Epidemiology is study of diseases spread in a living organism within the context of its environment[1]. Mathematical modeling can be used to study the epidemiology of a disease. Historically, researchers have used mathematical models to identify the spread of infectious diseases such as measles[2-4], rubella[5], HIV [6][7], dengue fever[8][9], TB[10,11], to what just happened such as ebola[12] and Zika virus[13].

Along with the development of science, mathematical modeling is not only used to study the spread of infectious diseases, but also non-communicable diseases. Even public illnesses such as drugs can also be modeled[14]. This can be done because there are similarities characteristic of its spread, that is through social interaction as media spread.

Diabetes is a non-communicable disease which has some various "spread" characteristic, one of them through social interactions that lead to lifestyle changes. Diabetes is a chronic disease that is indicated by high blood sugar levels. A person is diagnosed positive for diabetes when its fasting blood sugar level are more than 126 mg/dL or its blood sugar level 2 hours post meal are more than 200 mg/dL. Pancreatic organs that are supposed to produce the hormone insulin can not work optimally. Without insulin, the body cells can not absorb and process glucose into energy, so that the
blood sugar levels increasing (Diabetes Mellitus type I). Diabetes also occurs due to the body’s inability to utilize the available insulin (Diabetes Mellitus type II).

Based on 2015 International Diabetes Federation[15] data, the prevalence of diabetes is around 9.1% of the world’s population. World Health Organization[16] announced that diabetes was the sixth leading cause of death in the world. According to Deshpande et al.[17], diabetics death generally occur in complicated patients. In addition, complications of diabetes are also a major cause of amputations that result in permanent disability in the sufferer. Therefore, complications are an important component in the discussion of diabetes epidemiology.

In 2004, Boutayeb et al.[18] introduced the diabetes complication (DC) model to find out how many changes of diabetics without complications (D) and diabetics with complications (C). In the DC model the number of new patients (incidence) of diabetes is assumed to be constant. Meanwhile, based on WHO data[19], the number of incidences of diabetes is not constant and tends to increase every year. The increase is due to the lifestyle of the world’s population. One of the factors influencing lifestyle is social interaction. Social interaction is a significant factor affecting lifestyle changes so as to increase the potential of a healthy-susceptible individual into diabetics [20]. In addition, the incidence of diabetes is also often associated with genetic factors from parents who have a history of diabetes[21].

Different from Boutayeb et al.[18], in this paper the number of incidences is not constant. The number of incidences is determined by considering lifestyle factors and genetic factors. Taking into consideration the difference, the DC model was developed into susceptible diabetes complication (SDC) model.

2. The DC Model

The population in this model is diabetic who are classified according to their health conditions into two compartments. There were diabetics without complication (D) and with complications (C).

In this model the types of complication disease was not distinguished. Acute complications or chronic complications were both considered to be the same. Diabetics patients who initially did not exposed to complications can be developed into diabetics with complications. The occurrence of complications rate is denoted by $\lambda$ so that the number of individual on D compartment decreases as much as $\lambda D$ and the number of individual on C compartment increases in the same amount. Patients with diabetes who were exposed to complications can recover from their complications disease, died, or had a disability. Diabetics patients who recover from complications were assumed to still suffer from diabetes, while diabetics who died or disabled were excluded from the population. The recovery rate of complications, the death rate due complications, and disability rate were respectively denoted by $\gamma$, $\delta$, and $\nu$, so the number of individuals in the C compartment decreases as much as $\gamma C$, $\delta C$, and $\nu C$. At the same time, the number of individuals in the D compartment increases as much as $\gamma C$ because of complications recovery.

The DC model also considers natural deaths in both compartments. In both compartments, the natural mortality rates were denoted by $\mu$. Due to a natural death, the number of individual on D and C was decreasing by $\mu D$ and $\mu C$, respectively. The number of diabetes incidences in this model was assumed to be constant at I and categorized as uncomplicated diabetes. The number of individuals in the D compartment increases as much as $I$ because of incidence. So, the DC model according to Boutayeb et al.[14] can be written as

$$\frac{dD}{dt} = I - (\lambda + \mu)D + \gamma C$$
$$\frac{dC}{dt} = \lambda D - (\gamma + \delta + \nu + \mu)C$$

with $I, \lambda, \gamma, \mu, \delta, \nu > 0$. Model (1) is a first-order system of differential equation.
3. Results and Discussion

3.1. The SDC model

Increasing the number of diabetics is influenced by the number of incidences that occur. The incidence of diabetes is largely due to unhealthy lifestyles such as low levels of physical activity, irregular eating patterns, and other unhealthy habits. Patients with diabetes who exposed to complications usually already aware of the weakness of their physical condition. In contrast, diabetics without complications have not been able to maintain their lifestyle, especially in undiagnosed diabetics people.

In daily activities, diabetics without complications tend to have an unhealthy lifestyle. According to Hill et al.[22], the interaction between diabetics who have an unhealthy lifestyle and healthy individuals can lead to lifestyle “transmission”. The impact of lifestyle “transmission” is incidence. Incidence increases the prevalence of diabetics. To determine the number of possible interactions, the number of susceptible individuals should also be known in a new individuals group named susceptible (S). Due to the individuals group of susceptible, the DC (1) model is modified into the susceptible diabetes complication model (SDC). If the rate of interaction causing incidence is denoted by \( \beta \), then the number of incidences that occur due to lifestyle factors is \( \beta S D / N \). Every individual who has recently had diabetes is assumed to have no complications so that the number of individuals in the D compartment increases as much as \( \beta S D / N \). At the same time, the number of individuals in the S compartment also decreases as much as \( \beta S D / N \). Natural deaths also occur in the individual group of S. The natural mortality rate in the S compartment is equals to the natural mortality rate in the individual groups of D and C. Therefore the number of individuals in the S compartment decreased by \( \mu S \) due to natural death.

Besides being caused by an unhealthy lifestyle, the incidence of diabetes is also caused by genetic factors of parents who have a history of diabetes. Children of diabetics have the possibility of having a genetic disorder that causes unproperly pancreas. In this paper, we assume that the healthy individuals have a health children too, while diabetics have children with genetic disorders, but still have the possibility of having a health children. The proportion of genetic disorder’s birth is denoted by \( \rho \) and the birth rate is denoted by \( a \). The number of healthy individual birth is \( a S + a (1 - \rho) (D + C) \) and the number of genetic disorder’s birth is \( a \rho (D + C) \). Birth of healthy individuals increases the number of individuals in the S compartment, while the birth of individuals with genetic disorders increases the number of individuals in the D compartment.

The population on the SDC model is defined as the union of three compartments, there are S, D, and C. The union of those individual compartments represent the number of people living in a region (N). In this model, there was \( vC \) individuals excluded from the population due to disability. The elimination of individuals with disabilities creates a contradiction with the definition of population, so that the SDC model needs to be modified. Individuals with disabilities are part of the living population and their presence does not need to be excluded from the population. Individuals with disabilities are retained in the population by being part of the C individual group.

Thus, the SDC model is a mathematical model in the epidemiology of diabetes that takes into account lifestyle factors and genetic factors as causes of incidence and does not exclude disabilities from the population. The SDC model can be expressed as

\[
\begin{align*}
\frac{dS}{dt} &= \alpha S + \alpha(1 - \rho)(D + C) - \beta S D / N - \mu S \\
\frac{dD}{dt} &= \alpha \rho (D + C) - (\lambda + \mu)D + \gamma C \\
\frac{dC}{dt} &= \lambda D - (\gamma + \delta + \mu)C,
\end{align*}
\]

with \( S(0) > 0 \), \( D(0) > 0 \), and \( C(0) > 0 \). The parameters \( \alpha, \beta, \gamma, \delta, \lambda, \mu, \rho > 0 \) and \( 0 \leq \rho \leq 1 \) respectively denote birth rate, interaction rate, recovery rate of complications, death rate due complications, occurrence rate of complications, natural mortality rate, and the proportion of genetic disorder’s birth. The SDC (2) model is a first-order nonlinear differential equation system.
The solution of SDC model (2) can be used to determine the prevalence of diabetes. Prevalence is the standard measurement of diseased individuals in the population. The prevalence of diabetes can be calculated by the formula $\frac{(D(t) + C(t))}{N(t)}$.

3.2. Application

The SDC model (2) is used to determine changes in the prevalence of diabetes in the United States. The population is assumed to be homogeneous and does not consider the distribution of population areas or the distribution of age categories. Complications are limited to the five most commonly diseases which are found in people with diabetes. There are cardiovascular diseases (CVD), chronic kidney disease, lower extremity conditions (LEC), visual impairment, and ketoacidosis. Data were obtained from the Centers for Disease Control and Prevention (CDC)[23][24] and World Bank[25] 2000-2014. Data 2000-2011 are used to estimate the parameters, while the 2012-2014 data are used to measure the accuracy of the model.

Based on data for the year of 2000-2011, we obtain the value of birth rate is 0.01623, interaction rate is 0.16263, recovery rate of complications is 0.37141, death rate due complications is 0.0068, the occurrence rate of complications is 0.67758, and the natural mortality rate is 0.00764. While the proportion of genetic disorder’s birth is equal to 1/13 = 0.077. We substitute those parameters values to (2), so the SDC model in the United States is

$$\begin{align*}
\frac{dS}{dt} &= 0.008595S + 0.01498(D + C) - 0.1626SD/N \\
\frac{dD}{dt} &= 0.1626 SD/N - 0.68395D + 0.37265C \\
\frac{dC}{dt} &= 0.6776D - 0.3858C
\end{align*}$$

(3)

The solution of model (3) is determined using fourth-order of Runge-Kutta algorithm. The algorithm requires initial value. The initial value refers to the number of individuals S, D, and C in 2011 ($t = 0$). Based on CDC$^3$, 6.67% of 310.5 million people in United States are diabetics which consists of 9.65 million diabetics without complications and 11.05 million diabetics with complications, while the other 289.8 million is a healthy individual but susceptible to diabetes. Based on those data, we get initial value (in the millions of peoples), that is

$$(S(0), D(0), C(0)) = (289.8, 9.65, 11.05)$$

(4)

We determine the solution of (3) with an initial value of (4) in the first 3 years period (2011-2014). The solution value is compared with the data from CDC[23]. From this comparison, we know the relative error values in each compartments. The relative error values in S, D, and C, as well as relative error the prevalence of diabetes for 2011-2014 are shown in Table 1.

| Year | Relative error | Prevalence |
|------|----------------|------------|
| 2011 | S 0.00055, D 0.07625, C -0.09924 | -0.01756 |
| 2012 | S 0.00075, D 0.10426, C -0.15162 | -0.03161 |
| 2013 | S 0.00528, D 0.00642, C -0.18298 | -0.09736 |
| Av.  | S 0.00219, D 0.06231, C -0.14461 | -0.04884 |

The accuracy of the SDC model (3) in predicting the prevalence of diabetes in the United States can be determined by relative error in Table 1. The relative error rates for the individual compartments of S, D, and C are 0.00219, 0.06231, and -0.14461, respectively. The average relative error is quite small enough so that the model can be said to be accurate enough to predict the number of individuals in all three groups. While the average relative error in predicting the prevalence of diabetes is -0.4884, it means that the predicted prevalence of diabetes using the SDC model is only 4.88% higher than the actual data. Thus, the SDC model (3) is quite accurate in predicting the prevalence of diabetes in USA.

The prevalence of diabetes based on SDC model (3) solutions within the 2011-2030 timeframe is shown in Figure 1. Figure 1 shows that the prevalence of diabetes in 2011 is 6.67% and is predicted to
reach 14% by 2030, more than doubling (110%) in the first 19 years. Prevalence of diabetes tends to increase every year with an average increase of 4% per year.

![Prevalence of diabetes 2011-2030](image)

**Figure 1.** Prevalence of diabetes 2011-2030

The effect of an increase in prevalence is an increase in the number of deaths due to diabetes. In 2011 there were 73831 individuals who died of diabetes. The number of deaths due to continue to increase every year, an average increase of 5% higher than the number of deaths in the previous year. Deaths due to diabetes by 2030 are predicted to reach 206260 people, more than double than deaths in 2011. If accumulated, then there are 2.74 million individuals who die of diabetes within the 2011-2030 period.

**References**

[1] L. F. Berkman, I. Kawachi. 2001 Social epidemiology, Oxford University Press, NY
[2] D. W. Tudor, 1985 *Math. Bioscience*, 73, 131-147
[3] W. P. London, J. A. Yorke, 1973 *J. Epidemiology*, 98, 453-468
[4] D. Schenzle, 1984 *IMA J. Math. Appl. Med. Biology*, 1, 169–191
[5] H. W. Hethcote, 1983 *Am. J. Epidemiology.*, 117, 2–13
[6] X. Lin, H. W. Hethcote, P. van den Driessche, 1993 *Math. Biosci.*, 118, 181–195
[7] J.A.Jacquez,C.P.Simon,J.S.Koopman,L.Sattenspiel,T.Perry, 1988 *Math.Biosci.*,92,119-199
[8] Syafruddin, S. Noorani, M.SMd., 2013. Lyapunov function of SIR and SEIR model for transmission of dengue fever disease. *International Journal of Simulation and Process Modelling*, 8(2/3), 177-184.
[9] Rangkuti, Y.M. Side, S. Noorani, M.S.M., 2014. Numerical analytic solution of SIR model of dengue fever diseasein South Sulawesi using homotopy perturbation method and variational iteration method. *Journal of Mathematical and Fundamental Sciences*.46A(1), 91-105
[10] Side, S. Sanusi, W. Aidid, M.K. Sidjara, S., 2016. Global stability of SIR and SEIR model for tuberculosis disease transmission with lyapunov function method. *Asian Journal of Applied Sciences*. 9(3), 87-96.
[11] Side, S., 2015. A susceptible-infected-recovered model and simulation for transmission of tuberculosis. *Advanced Science Letters*. 21(2), 137-139.
[12] A. Azizah, P. Widyangsih, D.Saputro, 2017 *Journal of Physics Conf. Series*, 855, 012008
[13] E. Bonyah, K. O. Okosun, 2016 *Asian Pacific Journal of Tropical Disease*, 6, 673-679
[14] Sutanto, A. Azizah, P. Widyaningsih, D. Saputro, 2017 *AIP Conference Proceedings*, 1847

[15] International Diabetes Federation (IDF), 2015 IDF Atlas Seventh Edition

[16] World Health Organization (WHO), 2017 The Top 10 Causes of Death

[17] A.D. Deshpande, M. Harris-Hayes, M. Schootman, 2008 *Physical Therapy* 88, 1254–1264

[18] A. Boutayeb, E. H. Twizell, K. Achouayb, A. Chetouani, 2004 *BioMedical Engineering Online* 3, 20

[19] World Health Organization (WHO), 2016 *Report of a WHO/IDF Consultation*,

[20] T. L. Gary-Webb, S. F. Suglia, P. Tehranifar, 2013 *Current Diabetes Reports* 13, 850–859

[21] S. S. Rich, *Journal of The American Society of Nephrology*, 2006 17, 353–360

[22] J. Hill, M. Nielsen, M. H. Fox, *The Permanente Journal*, 2013 2, 67–72,

[23] Centers for Disease Control and Prevention (CDC), 2015 *National Data of Diabetes*,

[24] Centers for Disease Control and Prevention (CDC), *Deaths: Final Data 2000-2014, National Vital Statistics Reports (NVSR),* (2014).

[25] World Bank, Birth and Death Rate, Crude, (2015).