2017 CMBEC40 Conference
Winnipeg MB
May 23–26, 2017

MODELING THE EMERGENCE AND PROGRESSION OF CARDIOVASCULAR DISEASES IN THE JAMAICAN POPULATION

Orlando Simpson and Sergio Camorlinga
Department of Applied Computer Science
University of Winnipeg, Winnipeg, MB, Canada
simpson-o@webmail.uwinnipeg.ca and s.camorlinga@uwinnipeg.ca

ABSTRACT— In this paper, we developed an Agent-Based Model (ABM) that simulated the emergence and progression of Cardiovascular diseases (CVDs) in Jamaica. The preliminary results presented demonstrate that the model has merits. Therefore, it can assist policy makers to understand the complex interdependency of the risk factors and how that translates to the emergence of CVDs in the population.

Keywords—Agent-based modeling, modeling and simulation, complex adaptive system, cardiovascular diseases, risk factors, epidemiology, physiological, NetLogo

I. INTRODUCTION

According to the WHO [1] CVDs are the leading cause of death with approximately 17.5 million in 2012. However, developing countries accounted for 80% of CVD related deaths. For example Jamaica, a developing country corroborated these results as CVDs accounted for four out of five deaths in 2004 [2-4].

CVDs consist mainly of coronary heart disease (CHD) and stroke [1-5]. CVDs are caused by a cluster of risk factors interacting multiplicatively to promote a vascular risk [5]. Risk factors for developing CVD are classified into behavioral and physiological attributes [1][5]. Behavioral attributes include diet, physical activity and smoking. In contrast, physiological attributes include blood glucose, blood pressure, cholesterol, weight, age and sex [5].

Agent-Based Modeling (ABM) paradigm provides the flexibility to represent from very detailed models to highly abstract models. ABM [6] can be used to represent heterogeneous population and capture the dynamics of each individual's risk factors. ABM is a bottom-up approach. Therefore, it can be used to model the interdependencies of the risk factors (Fig. 1) and observe the possible emergence of CVD at the population level [6]. Simulation models can overcome these complexities because they require fewer restrictive assumptions than purely mathematical models.

The primary prevention for CVD is the avoidance of known risk factors by choosing a healthy lifestyle [5][7]. According to the WHO [1], these include avoid smoking, eating a healthy diet, engaging in sufficient physical activity etc. It is said that a reduction of 1% in body weight, lowers systolic blood pressure an average of 1 mmHg [5][7]. According to Cheng et al. [7], lowering systolic blood pressure conferred 36% reduction in strokes, 27% in non-fatal myocardial infarctions and CVD related deaths. Therefore, modeling and simulation could be used to explore "what-if" scenario planning in a virtual environment to determine what risk factors are the best biomarkers for a specific CVD outcome. E.g. to estimate the population-attributable portion of CVD arising from smoking. As such, the model is beneficial to clinicians, epidemiologists, policy makers and health organizations.

In this paper, we used the multivariate Framingham risk score algorithm [8] to determine the health outcome (e.g. healthy, CHD, stroke and death) from the interrelated risk factors. We used NetLogo for our model implementation. It is a multi-agent programmable modeling environment developed at Northwestern University [9] using steps from Simpson and Camorlinga study[10].

Two major contribution of this paper are: 1) the detail level of abstraction that captures the individual health progression realistically. e.g. determining if an individual will move from a state of physically active to physically inactive as shown in Fig. 1. From these risk factors, it is

more accurate to determine the health outcome using the Framingham CVD risk score. However in the literatures surveyed, most modeled the CVD health outcomes without the explicit use of risk factors and 2) calibration of the Framingham CVD risk score to the Jamaican population where it previously underestimated.

II. RELATED WORK

Computer simulation models have been used in health research and policy since the 1960s [6]. For instance, in preparing for global pandemics such as H5N1 and 9/11 terrorist attacks [6]. The surge in systems science to public health lies in 1) the rapid growth of CPU power and memory and 2) the ability to test different hypothesis in a virtual environment where mistakes are allowed to be made.

The Framingham Heart study led to the development of a CVD prediction algorithm [8]. Tulloch-Reid et al. [4] stated that most of these risk factors were designed for Caucasian in high income countries. However, it is one of the only CVD risk score to be assessed and validated outside an high income Caucasian population. Moreover, Chia et al. [11], stated the Framingham CVD risk score predicts fairly accurate for men while slightly overestimating for women in Malaysia. From our literature survey, all the of CVD related models were created for developed countries and none for developing countries.

III. METHODOLOGY

Fig. 1 shows the effects of the risk factors and state transitions for our CVD model. The oval shapes represent the state and the transitions are represented by the broken arrows. The baseline parameters were mainly extracted from Ferguson et al. [2], Taubman et al. [12] and Li et al. [13] studies. Others were estimated and calibrated accordingly. In a nutshell the model include: 1) creating individuals with the specified health characteristics from a text file, 2) starting the simulation with 1 time step representing 1 year, 3) iterating through each individual to calculate the transitional probabilities for each risk factor, 4) calculating the risk of CVD and death, 5) updating each individual age and health characteristics, and 6) saving each individual health characteristics.

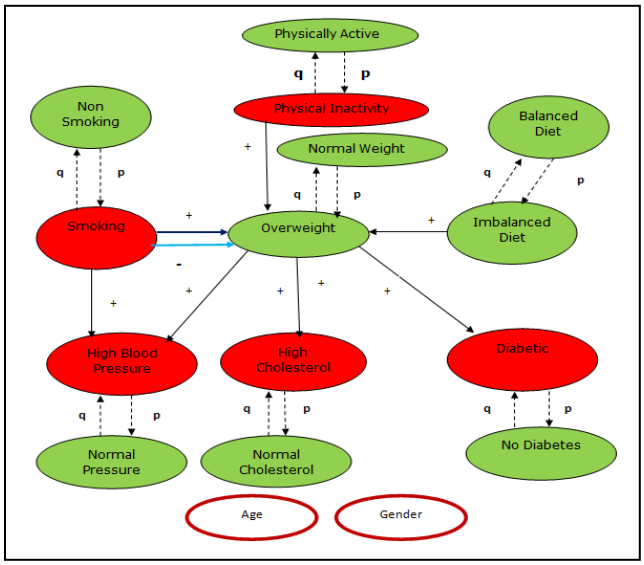

Figure 1 Complex interdependency of risk factors and state transitions with probability of "$p_{1..n}$" and "$q_{1..n}$".

The individuals in NetLogo is known as "agents". For example, each agent is simulated through a specified number of years or until age 74 or death before moving unto another agent in the model. Each agent health characteristics is defined by age, gender, body weight, cholesterol, blood pressure, blood glucose smoking, physical activity, diet and CVD history. Fifty (50) agents were generated based on distribution of the real Jamaican population from the Statistical Institute of Jamaica [14] and Ferguson et al. [2] study.

This model uses one time step to represent one year. Within this time step, the transition between risk factor states can occur. Therefore, age increase by 1 at each time step. Age and sex are independent variables that are not affected by other risk factors. On the other hand, all other risk factors (dependent variables) are affected by one or more risk factors. For example, high blood pressure is affected by age and body weight. Therefore, the relative risk associated with being overweight is multiplied by the transitional probability of normal to high blood pressure.

For every risk factors state, there is a known probability of transition to every other state. For example, an individual was noted as having a transitional probability of "p_1= 0.049" [13] moving from physically activity to another

state of physical inactivity. Therefore, a random number is generated and compared to the transition probability threshold. If the random number is less than the stated transition probability threshold, then that individual will move from physically activity to physical inactivity. Simulation rules are general programming statements e.g. testing if the individual will go in a physical active state: If random number "<" [occurrence probability threshold of physical activity]. While "q_1" is a transitional probability that moves in the opposite direction. E.g. physically inactive to physically active.

Each individual's health profile progress is based on the interaction of all the risk factors. At the end of each year, the CVD risk is generated using the Framingham risk score. Then the risk of CVD death is calculated based on the mortality rate of dying from CVD [14]. In contrast, a non-CVD death is calculated by subtracting the CVD mortality rates from the overall mortality rates of Jamaica [14]. When an agent ends up in the "death" state, it is removed from the population in the model.

The CVD history of an individual is updated and maintained for the entire length of the simulation. For the subsequent years, each individual's CVD risk profile is based on the state (e.g. agent 1 and year 1 = not smoking, normal blood pressure etc.) of the risk factors of the previous year. Once an individual has a history of CHD or stroke, he/she will have a higher risk of dying due to CVD.

IV. VALIDATION AND RESULTS

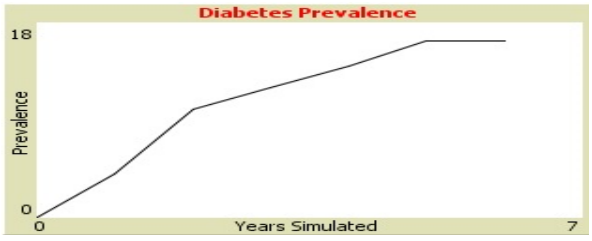

Figure 2 shows the prevalence of Diabetes from our model of the Jamaican population.

The development process included: 1) literature survey, 2) model formulation and 3) validation against actual data from Tulloch-Reid et al. [4] Jamaican study and from the Statistical Institute of Jamaica [14]. The simulation started at year 2002 and ended 2008 with an artificial population of 50. This was repeated for 10 simulations to decrease the uncertainty in the results based on the randomness in the model (fig.'s 2-4). This model was not able to be cross validated since it is first CVD model for Jamaica. One way sensitivity was conducted on each parameter. As such, the simulated CVD risk was compared to studies from [4] and [14]. The model showed a good fit for the simulated data vs. actual data based on the five measures in table 1. For example, the model had a 1.1 % relative error for people that are diabetic (fig. 2). Therefore, the model showed that it had a 98.9% accuracy in replicating the proportion of people who had diabetes. Moreover, the model categorized 95.2% of the population having a CVD risk score of less than 10%. In contrast, the Tulloch-Reid et al. study[4], stated that 90% of the Jamaican population had less than 10% risk of getting CVDs (fig. 3). This meant the model had a relative error of 5.8% which is 94.2% accuracy as shown in Figure 3. Based on the same study [4] in 2008, the prevalence of CVDs in Jamaica was estimated to be 3.39% or 81,617 persons. In contrast, our model produced a result of 3.66% (fig. 4 and table 1). This equates to 8% relative error and an accuracy of 92%.

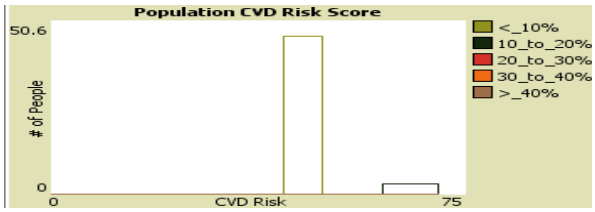

Figure 3 shows the CVD risk in the Jamaican population from our model.

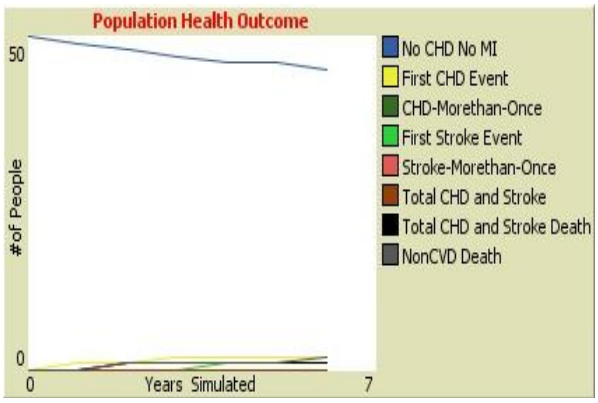

Figure 4 shows the health outcomes including CVD prevalence from our model of the Jamaican population.

Table 1. Shows the simulated results vs. the actual results. "A" - actual and "S"- simulated. The relative error column compares the mean simulated results against the actual results. The actual results are estimated from Tulloch-Reid et al. study [4] and Statistical Institute of Jamaica [14] using formula (mean simulated - actual)/actual. A Relative Error is an experimental error in approximation. A "-" means the simulated results underestimated the actual while "+" means overestimated.

Measures	**(A)-Value(%)**	**(S)-Mean (%)**	**Relative Error**
Diabetes Prevalence [4]	16.18	16.35	0.011
High Blood Pressure Prevalence [4]	48.4	40.58	-0.162
Overweight Prevalence [4]	36.25	34.72	-0.042
Population CVD risk of <10% [4]	90	95.2	0.058
Population CVD Prevalence [4] [14]	3.39	3.66	0.080

V. CONCLUSION

In this paper, we have proposed an Agent-based modeling and simulation approach to: (1) assess the emergence of CVD over a period of time and (2) better understand how CVD emerges from the interrelated risk factors of an individual. The preliminary simulated results show that the model had a good fit in most categories based on five measures. Population-wide efforts for reducing CVD risk factors and CVD deaths reduction have proven effective in some populations [5]. These include food labeling for nutritional content, educational programs to promote decreased saturated fats and sodium, promotion of smoking cessation, advocating physical activity for weight reduction etc. Therefore, the model could be used to determine which risk factors have the greatest overall effective (e.g. Physical inactivity and imbalance diet increase obesity, which increase blood pressure, cholesterol, blood glucose etc.) on the emergence on CVDs within Jamaica. As such, it could assist developing countries to allocate limited resources accordingly to prevent and manage CVDs. This is important as developing countries have limited resources. Consequently, this could reduce a country's health care cost, increase productivity and decrease its economic burden. Future work include more calibration, adding more validation measures such as CVD mortality rate and assessing the impact of lifestyle changes on the prevalence of CVDs in Jamaica. This include avoid smoking, eating a healthy diet and engaging in physical activity.

ACKNOWLEDGEMENTS

Research supported by NSERC DDG Grant program and the President's world leaders scholarship by the University of Winnipeg.

REFERENCES

[1] World Health Organization, (2016). Available URL: http://www.who.int/mediacentre/factsheets/fs317/en/. Accessed September 30, 2016.

[2] T. Ferguson et al., "Progression from prehypertension to hypertension in a Jamaican cohort: incident hypertension and its predictors".West Indian Med J, vol. 59, no. 5, pp. 486–493, 2010.

[3] T. S. Ferguson et al. , "Prevalence of prehypertension and its relationship to risk factors for cardiovascular disease in Jamaica", vol. 8, p. 20, 2008.

[4] M.K. Tulloch-Reid et al., "Excess CVD risk burden in Jamaican women does not influence predicted 10-year CVD risk profiles of Jamaica adults," 2013.

[5] R.E. Harris, Epidemiology of Chronic Disease: Global perspectives, Jones and Bartlett, 2013, pp 25-112.

[6] U.Wilensky , An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo, MIT Press, 2015.

[7] A. Cheng, J. Braunstein, D. C., C. Dennison, and C.Nass. "Reducing global risk for cardiovascular disease: Using lifestyle changes" Clin. Cardiol., vol. 25, no. 5, pp. 205–212, 2002.

[8] Framingham Heart Study (2015). Available URL: www.framinghamheartstudy.org/about-fhs/. Accessed Jan. 1, 2015.

[9] U.Wilensky.NetLogo.http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1999.

[10] O. Simpson and S. Camorlinga, "A methodology to create Complex Adaptive System models that support Cardiovascular Diseases simulation," in Systems Conference, British Columbia, 10.1109/SYSCON.2015.7116756, pp. 224 - 229, 2015.

[11] Y. C. Chia et al., "Validation of the Framingham general cardiovascular risk score in a multiethnic Asian population: a retrospective cohort study," BMJ Open, vol. 5, no. 5, pp. e007324–e007324, 2015.

[12] S. L. Taubman, J. M. Robins, M. A. Mittleman, and M. A. Hernán, "Intervening on risk factors for CHD," Int. J. Epidemiol., vol. 38, no. 6, pp. 1599–1611, 2009.

[13] Y. Li, N. Kong, M. Lawley, and J. Pagan, "Assessing lifestyle interventions to improve CVD health using ABM", IEEE winter simulation conference, 2014.

[14] Statistical Institute of Jamaica (2012). Available URL:http://statinja.gov.jm/demo_socialstats/newbirthdeathmarriage.aspx. Accessed August 10, 2014.