A System Simulation Model for Type 2 Diabetes in the Saskatoon Health Region

Jin Zhang
Media Access and Production, University of Saskatchewan
28 Campus Drive, Saskatoon, SK S7N 0X1
(306) 966-2500
Jin.Zhang@usask.ca
Nathaniel Osgood
Winfried Grassmann
Department of Computer Science, University of Saskatchewan
110 Science Place, Saskatoon, SK S7N 5C9
(306) 966-4898
grassman@cs.usask.ca and osgood@cs.usask.ca
Roland Dyck
College of Medicine
107 Wiggins Road, Saskatoon, SK S7N 5E5
(306) 966-7985
Roland.Dyck@usask.ca

Abstract
We describe a System Dynamics model for analysing and scenario analysis in regard to type 2 diabetes. The model has two parts, a normoglycemic part, and a hyperglycemic part. The normoglycemic part is broken down according to weight into normal weight, overweight and obese stocks, all indexed by age, sex and ethnicity, the last consisting of Registered Indian and others. The hyperglycemic part consists of two streams, diagnosed and undiagnosed, both divided according to the severity of hyperglycemia, starting with prediabetes. The model is distinguished from many past models by its more detailed representation of weight, diabetes-related cardiovascular risk, age, and ethnicity effects. The model was carefully calibrated, using a methodology discussed here. Some projections and policy suggestions are presented.

Keywords: Vensim, Diabetes, Calibration, Health

Introduction

Diabetes Mellitus, colloquially referred as “diabetes”, is a chronic syndrome that results from insufficient secretion or inefficient use of insulin. The prevalence of Diabetes Mellitus has increased rapidly due to population aging, urbanization, and increasing prevalence of obesity. An estimated 171 million people suffered from diabetes worldwide in the year 2000; the estimated number grew to more than 180 million in the year 2006. The diabetic population is projected to increase to 366 million by 2030 [1, 2]. In Canada, the prevalence of diabetes is

high and rapidly increasing. National Diabetes Surveillance System (NDSS) data estimated that 1.128 million people – about 4.8% of the total 20+ age population – were diagnosed with diabetes during 1998-99. If one factors in the estimate of the undiagnosed diabetics population, there were a total of 1.7 million Canadians with diabetes during 1998-99 [3, 4]. A recent report from the NDSS shows that the prevalence of diagnosed diabetes increased to approximately 1.783 million in the year 2004-05 [5]. The prevalence including undiagnosed diabetics increased to more than 2 million and is expected to rise to 3 million by the end of the decade [4, 6].

Saskatoon, a city in Saskatchewan, a Canadian prairie province, is not an exception to Canadian's high diabetes prevalence. Diabetes is the chronic disease with the third highest mortality burden in the Saskatoon Health Region (SHR) demographically. Obesity is a well-known risk factor of developing Type 2 diabetes. A high prevalence of obesity in the Saskatoon Health Region makes a considerable contribution to the high diabetes prevalence. The cold winters in Saskatchewan restrict outdoor physical activities and limited recreation facilities restrict indoor physical activities. Many urban residents also find it difficult and costly to access healthier food. Many people in the Saskatoon Health Region are overweight or obese. A total of 30.63% of the residents in the Saskatoon Health Region were overweight and 16.88% were obese in 2005 [8]. Partly as the consequence of the high prevalence of obesity, diabetes is prevalent in the Saskatoon Health Region.

Aboriginal people are a higher risk ethnic group for having diabetes in North America [15,16]. About 8.7% of the populations of the Saskatoon Health Region are of aboriginal origin, and over 50% of those are Status Indians based on Statistics Canada 2001 data. This relatively high percentage of aboriginal population when compared to other areas in Canada also contributes to the elevated diabetes prevalence seen in the SHR.

With the limited budget available for preventing and treating diabetes and its complications, cost-effective intervention policies are needed in order to slow the increase of prevalence of diabetes and lower the burdens imposed by diabetes, and save as many lives as possible.

For this reason, a group was formed at the University of Saskatchewan, analyzing this problem, which had the full support of the Saskatoon Health Region. One of the members of the group, Dr. Roland Dyck, was a clinician with extensive experience with diabetes, and he has published a number of papers in this area. It was decided to use a Systems Dynamics approach to analyze this problem, and build a detailed model using Vensim. For details of the model, see [7]. When constructing our model, we relied on some earlier studies. In particular, the U.S. Centers for Disease Control and Prevention (CDC) started a modeling project in 2003 to construct a System Dynamic model of diabetes [10]. The project team used the diabetes model to gain a better understanding of the diabetes burden in the U.S. and to evaluate possible interventions. The model simulates the diabetes onset process using the population at risk and the diabetes progression. After calibrating the model with historical data, the project team forecasted growth of diabetes and prediabetes prevalence through 2050 as the baseline for interventions evaluation. Another study we used was done by Rees et al. at Synergia Limited who developed a System Dynamics model to assist development of strategic diabetes policy for Manukau, a large multi-cultural city in New Zealand [11]. For our studies, we relied on both models extensively. However, we added considerable detail to the model, for instance, we indexed the diabetes population according to age, refined the representation of the progression of diabetes, and we adapted the model to the situation in Saskatchewan.

The Vensim Model

Like the CDC model and the New Zealand model, our model contains two core components: the normoglycemic population and the hyperglycemic population. The normoglycemic population section represents people in the SHR whose blood glucose levels lie within a normal range. This population group is divided into three major subgroups by weight categories: normal weight, overweight and obese (see Figure 1). Within a subgroup, the population is classified into smaller scale subgroups by age, ethnicity and gender. There are 17 different age groups, using commonly used 5-year age categories: 0 to 4, 5 to 9, 10 to 14, 15 to 19, 20 to 24, 25 to 29, 30 to 34, 35 to 39, 40 to 44, 45 to 49, 50 to 54, 55 to 59, 60 to 64, 65 to 79, 70 to 74, 75 to 79 and 80 plus. The ethnicity is Registered Indian (RI) and others, the non-RI. Age, ethnicity and gender are represented using subscripts in equations in order to keep the model clean.

As one can see from Figure 1, there is a progression from the normal weight stock to the overweight stock, and finally to the obese stock. All stocks have inflows through births, and outflows through deaths of all causes.

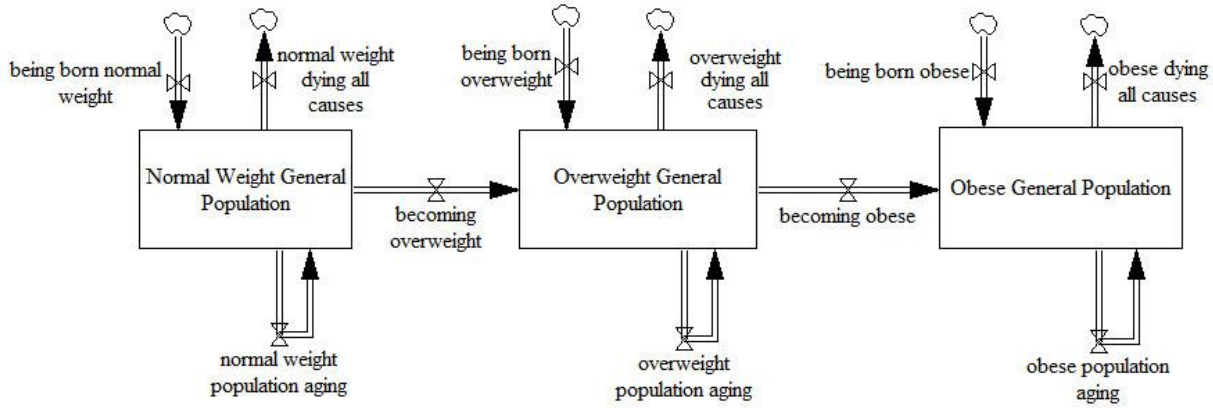

Figure 1: Normoglycemic Population Structure

All normoglycemic subpopulations have a risk of developing hyperglycemia; young age groups have lower risks, elderly age groups have higher risks; the population with lower BMI has lower risks, and the population with higher BMI has higher risks. Reflecting these, a developing prediabetes flow is associated with all general population stocks. Again, these outflows depend on the age, the ethnicity, and the gender. Of course, the outflows are higher for overweight and obese people. These outflows are omitted from Figure 1.

We now come to the component dealing with the hyperglycemic population (see Figure 2). Like in the CDC model and in the New Zealand model, we have two streams, an undiagnosed (undx), and a diagnosed (dx) stream. Each stream contains a progression of the severity of the diabetes. It starts with prediabetes, continues with diabetes without complications, and progresses to diabetes with early-stage macrovascular complications. The final stage is late stage macrovascular complications which we assume is always diagnosed. Hence, for this last stage, there is no undiagnosed stock. In summary, there an Undx Prediabetic Population a Dx Prediabetic population, and Undx Diabetic population without complications, a Dx Diabetic population without complications, and so on. The term diagnosis as used here means diagnosis of the correct stage. Consequently, there is a flow from dx Prediabetic Population to the undx

Diabetes Without Complications population, meaning that though prediabetes is diagnosed, diabetes is not. As in the component dealing with normoglycemia, all populations are also classified by age, ethnicity and gender. As before, the progression to the next stage depends on these characteristics. Also, at each stage, there are outflows due to deaths. They are omitted in Figure 2.

It is possible that persons having prediabetes may recover. This leads to three recovery outflows from prediabetes: one to the Normal General Population, one to the Overweight General Population, and one to the Obese General Population.

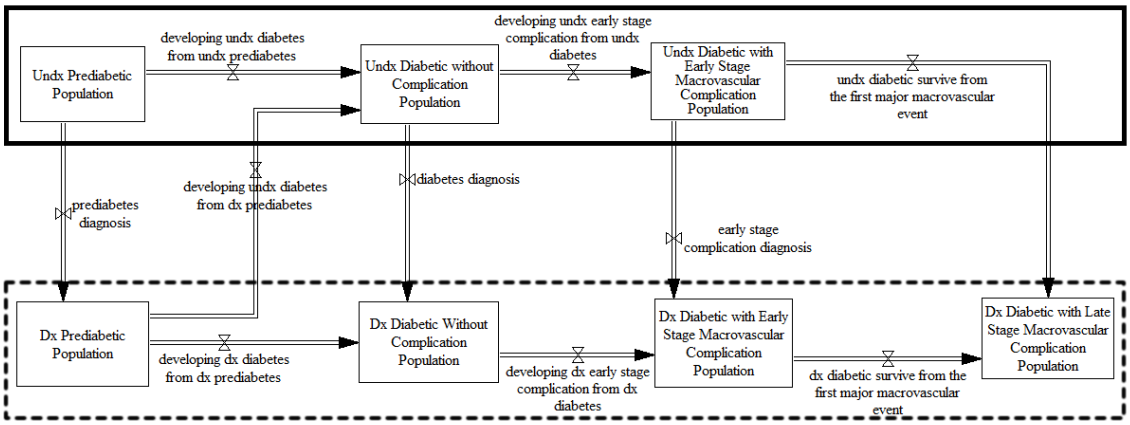

Figure 2. Development of Hyperglycemia

The stages of diabetes are characterized as follows. Patients with early stage macrovascular stocks start to develop macrovascular diseases caused by diabetes, such as coronary diseases and cerebrovascular disease.

The population in the final diabetes progression stage is the population who survived from the first attack of diabetes related macrovascular diseases, such as a heart attack or stroke. After the first attack of macrovascular diseases, any undiagnosed macrovascular complication is assumed manifest itself; hence, there is no undiagnosed population stock represented in the final stage. Diabetic patients stay in this stage until their death.

The Determinations of the Parameter Values and the Initial Values

Because of indexing by age, ethnicity and gender, there is a great number of stocks: there are 10 progression stages, 17 age groups, 2 ethnicities and 2 sexes, which means that there are 680 stocks in total. All these stocks have to be initialized, that is, we had to find 680 initial values, which had to be obtained from existing data bases or scientific papers, or, if they were missing, which had to be inferred by using analogies or expert opinion. Finding initial value for undiagnosed hyperglycemic stocks, in particular, posed a substantial challenge. There are also a great number of flows between the stocks, and to find the rates of these flows, we used numerous equations, and the parameter of these equations had to be obtained. Some of them were available from our data sources, and others were determined by analogy and/or expert opinion.

Our data sources can be grouped into four categories: local health authority reports, statistical surveys and the Canadian Census, research papers, and experts' opinion. We first consulted data from local health authority reports. The Saskatoon Health Region Authority releases a report every year summarizing health region operations in the previous year [8, 9, 12, 13]. These annual reports contain, among others, demographic information and its dynamics about the health region. In particular, they provide detailed demographic information especially regarding residents with Registered Indian status, who are at particularly high risk with respect to diabetes [16]. Annual reports describe the health status of residents in the region by using important health indicators; prevalence of diabetes is one of health indicators in the reports. We were able to find the diabetes prevalence in the health region directly from the reports. Besides the annual reports, the Saskatoon Health Region Authority released a detailed health status report every five years [14]. This report listed some key indicators, including population size, structure, birth rate, death rate and more, as attributes of overall health status of SHR population. The values of many parameters in our model were drawn from in this report.

Statistical surveys are powerful instruments for collecting quantitative information in a population. They obtain first-hand raw information directly from the targeted population. For our research we obtained data from several statistical surveys, such as the National Population

Health Survey (NPHS), the Canadian Community Health Survey (CCHS) and Canadian Census, collected by Statistics Canada and other agencies.

We reviewed some pioneering studies of the applications of the System Dynamics methodology, which, aside from helping us greatly to build our model, also provided some useful data. Two of these studies were specifically for diabetes and intervention: the CDC model [10] and the New Zealand model [11]. These two models targeted a different population, they maintained a similar structure, and required similar data, and we were therefore able to use some common data (e.g. concerning disease progression rates) for the initial values of parameters in our model.

We also consulted with some local diabetes experts. Their opinions and valuable suggestions provided a good starting point for parameterization of the model. Their opinion also gave us reasonable value ranges when we could not find accurate values for model parameters.

Though we spend a considerable time searching for data to find parameter values and initial values, and we looked at all available data very carefully, there was no way to find all these values in this fashion. To find the remaining values, we used calibration. Hence, we selected additional historical time series, and for each of these historical series, we added the necessary code to create the corresponding model. We then determined the unknown parameter in such a way that the model series were as close as possible to the historical time series. We selected the following five time series: the total population in the SHR, the normoglycemic population in the SHR, the total mortality rate in the SHR, the prevalence rate of diabetes in the SHR and the incidence rate of diabetes in the SHR. All these series were broken down according to age, ethnicity and sex. We then minimized the squared relative differences between the model series and the historical series, using the following formula:

$$\text{discrepancy} = \left[\frac{h - m}{\left[\frac{h + m}{2}\right]}\right]^2.$$

Here, h denotes the historical series, and m the model series.

In the first runs, we just added all discrepancies and minimized their sum. When doing this, the results were unsatisfactory for the following reason. Some of the historical data had very small sample sizes, whereas in other cases, the samples sizes were huge. For instance, the number of registered Indians below 25 years of age with diabetes is very small, whereas the corresponding population of the normoglycemic population of the SHR below 25 is large. The problem is now that a better fit to some data tends to cause a worse fit for others. Data from small samples are affected much more by random variations than data from large samples. When using equal weights, calibration may try to adjust parameters to what are essentially random variations at the expense of obtaining a good fit for data based on large samples. In our model, this led to substantial distortions. For this reason, it was necessary to give data based on small samples a smaller weight than the data based on large samples. Hence, before adding the discrepancies, we multiplied them by some weight, were the weight reflected the sample size of the historical data.

In this way, we could estimate unknown initial values, certain delays, and a number of parameters affecting the rates of moving between the stocks by minimizing the squared differences. Minimizing squared differences is, of course, well known from statistics, but it can also be justified by other arguments. For instance, it avoids having very large deviations.

Using the Model to Predict the Diabetes Burden, and Find Policies for Improvements

Our model is based on earlier models which we improved in several important aspects. One important change over the New Zealand model was that we added the age structure to the hyperglycemic stocks. Of course, our model is not perfect either. For instance, our simulation does not consider immigration and emigration from and to the SHR. One reason for this is that the migration patterns cannot predicted very well, and data are lacking.

Since it takes a long time for people to develop diabetes, we need a long time horizon for our simulation to get reasonable results. Hence, we ran the model to simulate the diabetes burden

from year 2001 to year 2101 based on current conditions. To interpret the results of our simulation, it is important to consider the population size and its age structure.

Figure 3 illustrates the trajectory of the total population in the SHR. The projection of the total population slowly increases during the first 20 years of the simulation period. After the total population reaches its peak level around the year 2025, it starts to decrease slowly due to the low birth rates.

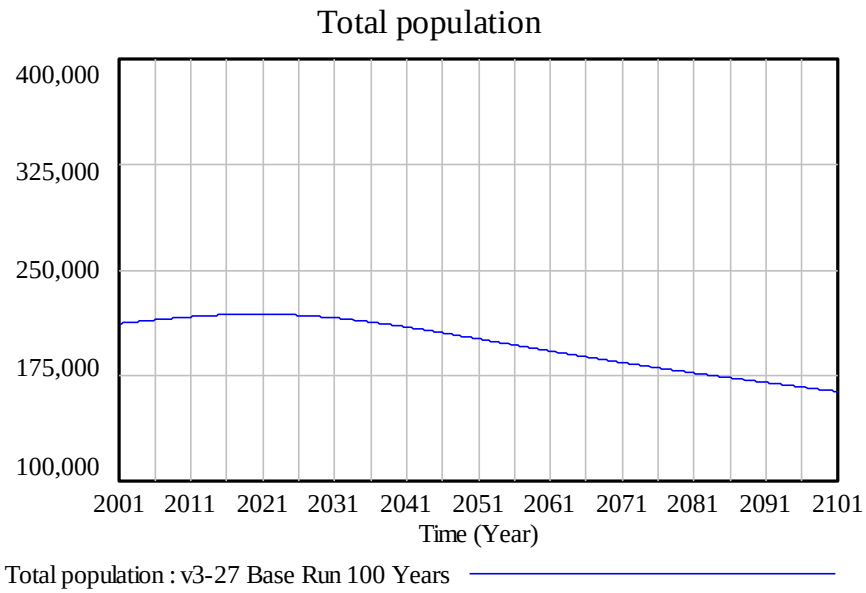

Figure 3. Total Population

The diabetic population in the SHR is the most important indicator of diabetes burden. To interpret the results obtained, note that due to the baby boomers, the population will get older for some time, but after the bulk of the baby boomers has gone through, the average age in the population will decrease. Since older people are more likely to get diabetes, this means that the incidence of diabetes will at first increase, but it will decrease in the long run. Hence, the prevalence of diabetes shows a similar pattern as the population. The total diabetic population is, of course, the product of the two time series, which means that this pattern is even strengthened as show in Figure 4.

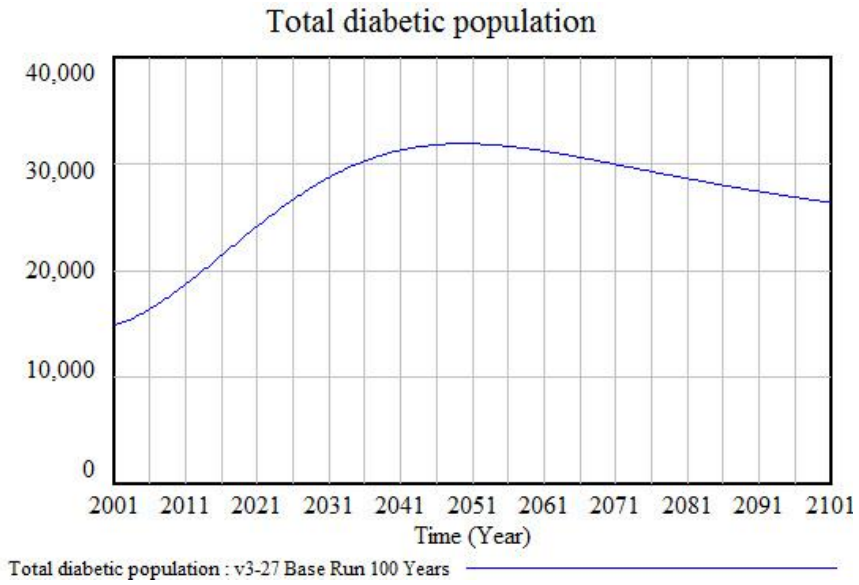

Figure 4. Diabetic Population

To highlight the importance of diabetes, let us consider relative figures. The population in the SHR in 2001 was roughly 215,000, and it is forecast to increase to 220,000 by 2031. The number of diabetes patients in 2001 was roughly 14,200, which is 6.9 percent of the population. This increases to 28,800, which is 13.1 percent of the population by 2031. Hence, the diabetes burden increases significantly. One reason for this increase is the prevalence of overweight and obese people, which increases from 14.7% in 2001 to 19.4% in 2031. The increase of obesity and the resulting increase in the diabetes burden make the introduction of policies ameliorating the situation imperative. For this reason, we first conducted a sensitivity analysis to find which parameters are most sensitive to changes, and we then tried a number of policies. We changed a number of parameters by plus or minus 20% and observed the effects. We found that the overweight incidence rate and the obesity incidence rate both have a strong impact on the diabetes population, especially if they are combined. Also, the number of years to develop diabetes from prediabetes has a strong effect. This affects two rates, the rate from undx prediabetes to diabetes, and the rate from dx prediabetes to diabetes.

Intervention policies based on lowering the overweight incidence rate and the obesity incidence rate - such as fitness classes, support for recreational facilities, and programs to make nutritious food more affordable and accessible when compared with less nutritious food, can

help to lower the diabetes prevalence, as shown in the Figure 5. In this figure, the upper line represents the projected diabetes prevalence in the base case, while the lower line represent the projected diabetes prevalence resulting from lowering the overweight incidence rate and obesity incidence rate fivefold. As one can see, these policies are very effective in slowing down the rate at which the diabetic population and diabetes prevalence are rising, and with it the diabetes related burdens.

These kinds of interventions will not lower the diabetic population from its original path immediately since they are not intervening directly on the diabetic population directly. The changes in the overweight and obese population will affect the diabetic population with levels of delays and manifest their impacts in the long run.

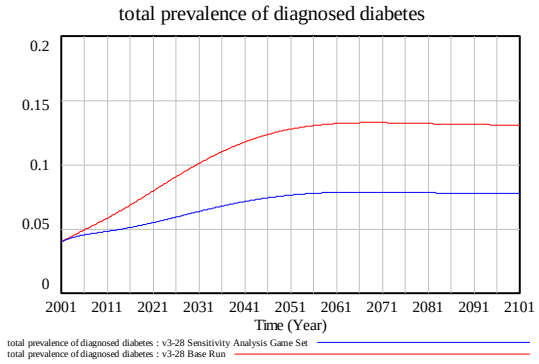

Figure 5. Changes in Diabetes Prevalence from Lowering Overweight Incidence Rate and Obese Incidence Rate Fivefold

If we implement an intervention policy to delay diabetes onset in the prediabetic population, it could lower diabetes incidence and the diabetes prevalence greatly. Figure 6 illustrates the improvement in the diabetes prevalence and the number of new incident cases by doubling the average years to develop diabetes from undiagnosed prediabetes and the average years to develop diabetes from diagnosed prediabetes. The diabetes prevalence and incident of new cases (illustrated by bottom lines) are much lower than the baseline. Hence, using measures to

increase the time to pass from prediabetes to diabetes is effective in reducing the diabetes burden.

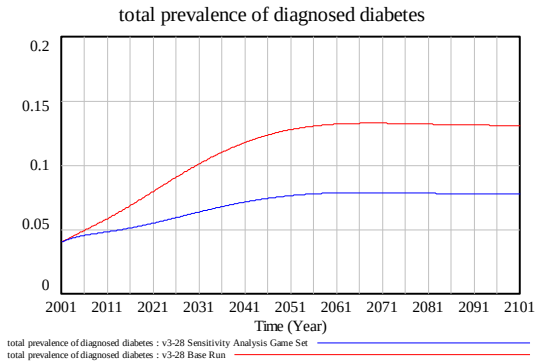

Figure 6. Changes in Diabetes Prevalence from Doubling Average Years to Develop Diabetes from Diagnosed Prediabetes and Undiagnosed Prediabetes

Conclusions

In this paper, we discussed a model used to gain insight into the burden caused by diabetes. We will continue to refine this model, and to use it to further our study in this important area. Unfortunately, the policies we suggest only work over relatively long time scale, much longer than the typical planning horizon of policy makers, and this may make the introduction of these policies difficult.

REFERENCES

[1] S. Wild, G. Roglic, A. Green, R. Sicree and H. King. “Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030.” Diabetes Care, vol. 27, pp. 1047-1053, Oct. 2004.

[2] World Health Organization. "Fact Sheet No312, Diabetes." Internet: http://www.who.int/mediacentre/factsheets/fs312/en/, Sep. 2006. [March 2007].

[3] Health Canada. (Jan. 2003). *Diabetes in Canada*. (2nd Edition). [On-line] Available: http://www.phac-aspc.gc.ca/publicat/dic-dac2/pdf/dic-dac2_en.pdf. [Mar. 2007].

[4] Canadian Diabetes Association. "Diabetes Facts." Internet: http://www.diabetes.ca/Files/Diabetes_Fact_Sheet_Mar04.pdf, Mar. 2004. [Mar. 2007].

[5] National Diabetes Surveillance System. "Diabetes in Canada: Highlights from the National Diabetes Surveillance System, 2004-2005." Internet: http://www.phac-aspc.gc.ca/publicat/2008/dicndss-dacsnsd-04-05/pwdd-iadd-eng.php, Apr. 23, 2008. [June. 2011].

[6] Canadian Diabetes Association. "The Prevalence and costs of diabetes." Internet: http://www.diabetes.ca/Files/prevalence-and-costs.pdf. [June 2011].

[7] J. Zhang. A System Simulation Model for Type 2 Diabetes in the Saskatoon Health Region, M. Sc. Thesis, Department of Computer Science, University of Saskatchewan, July 2011.

[8] Saskatoon Regional Health Authority. "2006-2007 Annual Report to the Minister of Health Living Services." Internet: http://www.saskatoonhealthregion.ca/about_us/documents/shr_annual_report_2006_07.pdf, Jul. 2007. [June 2011].

[9] Saskatoon Regional Health Authority. "2004-2005 Annual Report." Internet: http://www.saskatoonhealthregion.ca/about_us/documents/shr_annual_report_2004_05.pdf. Jul, 2005. [June 2011].

[10] J. Homer, A. Jones, D. Seville, J. Essien, B. Milstein and D. Murphy. "The CDC's Diabetes System Modeling Project Developing a New Tool for Chronic Disease

Prevention and Control," in *Proc. 22nd International Conference of the System Dynamics Society*. July 2004.

[11] D. Rees and B. Orr-Walker. *System Dynamics Modelling as a Tool in Healthcare Planning.* [On-line]. Available: ftp://vps224.2day.com/assets/File/Publications/System%20Dynamics%20Modelling%20as%20a%20Tool%20in%20Healthcare%20Planning.pdf. [June 2011].

[12] Saskatoon Health Region. "2003-2004 Annual Report." Internet: http://www.saskatoonhealthregion.ca/about_us/documents/shr_annual_report_2003_04.pdf. Mar. 2004. [June 2011].

[13] Saskatoon Regional Health Authority. "2005-2006 Annual Report to the Minister of Health and the Minister of Healthy Living Services." Internet: http://www.saskatoonhealthregion.ca/about_us/documents/shr_annual_report_2005_06.pdf. Jul. 2006. [June 2011].

[14] Saskatoon Health Region. "2004 Health Status Report." Internet: http://www.saskatoonhealthregion.ca/your_health/ps_public_health_profinfo.htm. 2004.[Sep. 2008].

[15] P.W. Frank, H.C. Looker, S. Kobes, L. Touger, P.A. Tataranni, R.L. Hanson and W.C. Knowler. "Gestational Glucose Tolerance and Risk of Type 2 Diabetes in Young Pima Indian Offspring." *Diabetes*, vol. 55, issue 2, pp. 460-465. Feb. 2006.

[16] R. Dyck, N. Osgood, T.H. Lin, A. Gao, M.R. Stang. "Epidemiology of diabetes mellitus among First Nations and non-First Nations adults." *Canadian Medical Association Journal*, vol. 182, no. 3, pp. 249-256. Jan. 18, 2010.