SOCIETY OF ACTUARIES

LIVING to 100
SOCIETY OF ACTUARIES INTERNATIONAL SYMPOSIUM

Mortality and Longevity

Aging and Retirement

Chronic Diseases and Longevity Risk: An Application to Type II Diabetes Insurance Products

January 2020

Chronic Diseases and Longevity Risk: An Application to Type II Diabetes Insurance Products

AUTHOR

Hsin-Chung Wang, Ph.D.
Aletheia University

Jack C. Yue, ASA. Ph.D.
National Chengchi University

Ting-Chung Chang, Ph.D.
Chihlee University of Technology

SPONSOR

Society of Actuaries:
Aging and Retirement Research
Committee on Life Insurance Research
Financial Reporting Section
Mortality and Longevity Research
Product Development Section

CONTENTS

Chronic Diseases and Longevity Risk: An Application to Type II Diabetes Insurance Products

Abstract

Prolonging life is a global trend and population ageing is speeding up in many countries. As a result, more medical expenditures are used for chronic diseases since the elderly often have chronic diseases. For example, about 3/4 and 1/2 of Taiwan's elderly have at least one and two chronic diseases, respectively. Diabetes is a common chronic disease and many serious health conditions are connected with diabetes. Total estimated cost associated with diabetes in the United States was $245 billion in 2012. However, many people are not aware that diabetes is common and at least 23.8% of patients do not know they have diabetes. (Source: U.S. National Diabetes Statistics 2017) The numbers of deaths related to diabetes are increasing and it becomes the 5th and 4th cause of death in 2017 for the Taiwan men and women, respectively (Source: Taiwan's Ministry of Health and Welfare).

In this study, we aim to evaluate the cost of diabetes and to design insurance products to deal with the disease, using the data from Taiwan's National Insurance (NHI). In particular, we are interested in the incidence and mortality rates of diabetes, as well as its medical usage. We use mortality models, such as the Lee-Carter and Age-Period-Cohort models, to explore their trends. Also, we will use the continuous prescription for chronic diseases to determine whether people know they have diabetes, since people with the continuous prescription can have 3-month refillable prescription under the NHI. The empirical study is based on two data sets from the NHI, and both data sets are one-million random samples of Taiwan people: one for the group of ages 0-99 and the other for ages 65-99, accounting for about 4.6% and 45.7% of Taiwan's populations in each age groups. The advantage for using the data set of ages 65-99 is to increase the sample size of the elderly and thus increasing the stability of analysis results.

We found that the incidence and mortality rates of diabetes change with a constant and stable path, and the Lee-Carter model can provide fairly satisfactory estimates. The analysis results also indicate that the people with diabetes without taking diabetes medication have higher mortality rates, than those taking diabetes medication regularly. We also demonstrate how these results can be used to design insurance products associated with diabetes, which can help the insured and their families to face the consequence. In addition, we discuss different criteria for judging whether people have diabetes in commercial insurance and show that they can be related to the possibility of moral hazard for the diabetes products. The judging criteria of diabetes vary a lot between different doctors.

Keywords: Diabetes, Chronic Diseases, Mortality Models, Longevity Risk, National Health Insurance

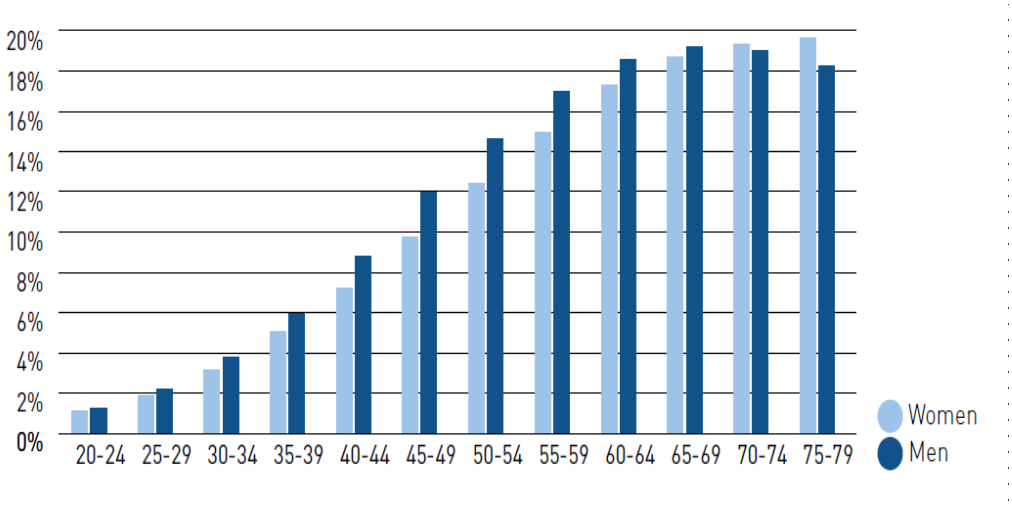
Section 1: Introduction

Prolonging life expectancy is a global trend in the 21st century and population ageing is becoming more apparent in many countries. As a result, people are paying more attention to life planning after retirement since recently medical expenses are capturing more attention, in addition to the economic needs. The elderly have higher medical utilization, especially due to chronic illness and conditions. For example, the annual medical expenses of the elderly in Taiwan are five times of the national average. More medical utilization of Taiwan's elderly is associated with metabolic syndrome diseases, including heart disease, stroke and type 2 diabetes. However, cancer is still the focus in Taiwan. Many people are not aware of the danger of metabolic syndrome diseases to their health and now they are the leading death cause in Taiwan.

Among all metabolic syndrome diseases, diabetes does not receive as much attention as heart disease and stroke, but people gradually realize its impact on health (Lin et al., 2004; Zhang et al., 2010; da Rocha Fernandes et al., 2016). The global prevalence rate of diabetes has grown from 4.7% to 8.5% between 1980 and 2014.[1] In addition, there were 1.5 million people who died of diabetes in 2012, and according to the estimate, the direct expenses of diabetes were over $82.7 billion. In the International Diabetes Federation (IDF) Diabetes Atlas 2017, the prevalence rates of diabetes for women and men 20-79 years are 8.4% and 9.1%, respectively, and they increase with age (Figure 1). The total healthcare expenditure on diabetes will reach USD 958 billion (18-99 years) in 2045.

Diabetes is also one of the major health threats in many countries and its importance increases with time. Many death causes (e.g., kidney and cardiovascular diseases) are highly correlated to it, although it might not be the direct cause. Approximately 4 million people aged between 20 and 79 years are estimated to die from diabetes in 2017 (i.e., equivalent to one death every eight seconds), compared to 1.5 million in 2012. Population ageing will keep the global prevalence rate of diabetes rising, since the incidence rate increases with age. (Wild et al., 2004; Shaw et al., 2010; Danaei et al., 2011; Whiting et al., 2011).

Figure 1
PREVALENCE RATE OF DIABETES BY AGE AND SEX, 2017



Source: IDF DIABETES ATLAS, Eighth Edition, 2017

[1] According to the 2016 report of World Health Organization, adult diabetic all around the world has increased from 108 million people in 1980 to 422 million people in 2014. In addition, and the direct expenses of diabetes were over $ 82.7 billion.

An increasing prevalence rate is not the only problem when we deal with diabetes. Many people do not realize that they have diabetes, based on past studies and our previous work. According to the National Diabetes Statistics 2017, Center for Disease Control and Prevention (CDC), the percentage of adults with diabetes increased with age, reaching a high of 25.2% among those aged 65 years or older. There are 30.3 million people of all ages (9.4% of the U.S. population) that had diabetes and 7.2 million (23.8%) were not aware of or did not report having diabetes in 2015. Holman et al. (2011) observed that 27.1% of diabetic patients were undiagnosed in 2009. Dwyer-Lindgrenet al. (2016) pointed out about 20-40% of the counties in the U.S. had undiagnosed diabetes in 2012. From the 2016 and 2017 IDF reports, around 46.5% and 50.0% of the world's diabetes patients were unaware of their disease, respectively. Apparently, there will be high proportion of undiagnosed diabetes in Taiwan.

We expect that the influence of diabetes to our health and life expectancy will continue to increase, and the insurance industry can play an important role. But we cannot rely on the prevalence rate of diabetes alone to design insurance products. We need other information, such as incidence rates, mortality rates, and medical utilization, of diabetes. Unfortunately, only a few reliable data sets are available. In this study, we use data sets from Taiwan's National Health Insurance Research Databases (NHIRD) to explore the feasibility of designing diabetes related products. Taiwan launched National Health Insurance (NHI) in 1995 and almost all Taiwan citizens are covered. All the hospital visits, including inpatient, outpatient, and surgical records, are stored and the Taiwan government will release various sample data sets every now and then. Thus, we should have reliable estimates of diabetes based on the sample data sets.

The current paper has been arranged as follows. We briefly introduce the data sets used in this study, which are from Taiwan's NHIRD, and define the definition of diabetes and mortality models in Section 2. Empirical data analysis, including the process of big data analysis is given in Section 3. Modelling the incidence and mortality rates is given in Section 4. The final section contains discussions about the study of diabetes, the applications of diabetes and suggestions.

Section 2: Data and Methodology

The NHI covered about 99.6% of Taiwan's citizens at the end of 2018, which means that the NHIRD records almost the whole population in Taiwan and it becomes an important research resource. Quite a lot of studies proposed using the NHIRD to estimate, for example, the, the hospital utilization and incidence rates of certain diseases in Taiwan (Chiang et al., 2017; Yue et al., 2018). In this study, we use two sample data sets from NHIRD: Longitudinal Health Insurance Database 2005 (LHID2005) and Elderly (65+) Longitudinal Health Insurance Database 2005, to acquire estimates of incidence, mortality, and medical utilization related to diabetes. These data sets contain one million randomly selected people who were still alive in 2005 and trace their health records between 1996 and 2013. Various types of records are involved in these data sets: "registry for beneficiaries" (personal identification, or ID file), "ambulatory care expenditures by visits" (outpatient visit or CD file), and "inpatient expenditures by admissions" (DD file).

The difference between these two data sets is that the samples in LHID2005 are between ages 0 and 99 while those in the elderly data are ages 65-99, accounting for about 4.6% and 45.7% of Taiwan's populations. According to the past work, most of diabetes patients are type 2 and the age-specific prevalence rates increase with age, especially for people over 65 (Wild et al., 2004; Cockram, 2000). The diabetes related mortality was also higher among the elderly (Manderbacka et al., 2011). Hence, we should focus our goal on the group of ages 45-99 and we obtain our estimates mainly from the elderly data set due to the consideration of sample size. Note that the data quality, as well as the data collection, of NHIRD is not reliable when Taiwan started the NHI in 1995. Thus, usually researchers use the data after 2000 and we also only use the data for the period 2003-2012.

Also, unlike cancer and catastrophic illness (Yue et al., 2018), diabetes does not have a unified definition in Taiwan yet. The government, insurance industry and doctors have different opinions about the symptoms of diabetes and this creates difficulties in designing diabetes insurance products. Table 1 shows the definition of diabetes criteria in past studies. The study settings, such as study objects, sample sizes, and observation periods, are quite different in past work. In the next section, we will compare different criteria of type 2 diabetes and evaluate them using data from the two data sets from NHIRD. The evaluation is based on, for example, whether the estimated incidence and mortality rates are consistent and stable between different ages and years, using the mean absolute percentage error (MAPE) as the evaluation criterions. The MAPE is defined as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|\varepsilon_i|}{Y_i}\times 100\%,$$

where Y_i and ε_i are the observed value and residual of observation i, $i = 1, 2, \ldots, n.$ According to Lewis (1982), a prediction with MAPE of less than 10% is treated as highly accurate, and that with MAPE greater than 50% is considered inaccurate.

Table 1

STANDRDS OF VARIOUS DIABETES STUDIES

Source	Definition
Diabetes dataset (NHRI)	ICD9: 250、6488、7751、7902、6480、A181。
Longitudinal Cohort of Diabetes Patients (NHRI)	Patients diagnosed with ICD codes (250, 6480, A181); prescription for hypoglycemia; 2+ diabetes clinics within one year; 1+ diabetes clinic for hypoglycemic outpatient
Chang et al., 2010	Patient with 1+ diabetes diagnosis codes (250, A181) within 365 days; or 3+ diabetes clinic visits per year
Li et al., 2012	Patient with 1+ hospitalized diabetes diagnosis codes (250, A181); or 3+ diabetes clinic visits per year
Jiang et al., 2012	3+ outpatient diagnostic codes (250, A181) or discharged from diabetic inpatients within 1 year
Lin et al., 2013	3+ outpatient diagnostic codes (250, A181), or 1+ diabetes inpatient as a discharge diagnosis
Lin et al., 2015	Patient with diabetes codes (250x0, 250x2) on discharge, or 3+ outpatient treatments for type 2 diabetes in a year
Lee et al., 2016	Patients admitted to the hospital with diabetes diagnosis code (250, A181) or 3+ diabetes outpatient visits per year

We can also obtain the mortality rates of diabetes patients from the NHI data. Although the NHIRD does not have a complete list of death records, we can still acquire stable and reasonable estimates of mortality rates, with the help of medical records. It was shown that the death criteria can provide fairly accurate estimates of mortality rates, especially for the elderly and the catastrophic illness patients (Chen et al., 2015; Yue et al., 2018; Yue et al., 2019). We adapt the criteria used in the previous studies and judge whether a patient is dead or not based on the following criteria:

Condition 1: Withdrawal in Registry for beneficiary's dataset (ID) and no medical visits for two consecutive years.

Condition 2: Last outpatient records are emergency visits, and no medical visits for two consecutive years.

Condition 3: Last inpatient records are more than 30 days and no medical visits for two consecutive years.

Condition 4: No medical visits for two consecutive years for Catastrophic Illness patients.

Condition 5: Last medical records are death, suicide, or discharged and no medical visits for two consecutive years.

Although the data set used covers almost half of Taiwan's elderly population, the incidence and mortality estimates can have a lot of fluctuations for higher age groups. We will introduce graduation methods to smooth the age-specific rates. Two types of smoothing techniques are used: Partial SMR and stochastic models. The partial SMR (Lee, 2003) is a modification of SMR (Standar d Mortality Ratio) and it is to smooth the mortality rates of small populations using the information (with respect to SMR) from a large (namely, reference) population. The SMR is often used in epidemiology and is defined as

$$\mathrm{SMR} = \frac{\sum_x d_x}{\sum_x e_x} \tag{1}$$

where d_x and e_x are the observed and expected numbers of deaths for age *x*, respectively. Note that SMR > 1 (SMR < 1) indicates whether an area has overall higher (or lower) mortality rates than the reference population and thus the SMR can be treated as a mortality index. Wang et al. (2018) showed that the partial SMR can be used to stabilize the estimates from stochastic models.

For the partial SMR, the graduated rates satisfy

$$v_x = u_x^* \times \exp\left(\frac{d_x \times \hat{h}^2 \times \log(d_x / e_x) + (1 - d_x / \sum d_x) \times \log(\mathrm{SMR})}{d_x \times \hat{h}^2 + (1 - d_x / \sum d_x)} \right) \tag{2}$$

or the weighted average between raw mortality rates and SMR, where $\hat{h}^2$ is the estimate of parameter h^2 for measuring the heterogeneity (in mortality rates) between the small area and big area. To avoid unreasonable results, Lee suggests that

$$\hat{h}^2 = \max\left(\frac{\sum\left((d_x - e_x \times SMR)^2 - \sum d_x\right)}{SMR^2 \times \sum e_x^2} \ , \ 0 \right) \tag{3}$$

The larger $\hat{h}^2$ is the more in heterogeneity. When the number of deaths is smaller, there will be larger weight from the large area, and the graduated value equals $\mathrm{SMR} \times u_x^*$ when the number of deaths is zero.

Stochastic models can be treated as a group of graduation methods. In addition to the frequently used Gompertz model for the elderly, we use the Generalized Age-Period-Cohort (GAPC) model (Villegas et al., 2016) to fit the incidence rates and mortality trends of diabetes patients. The popular Lee-Carter model (Lee and Carter, 1992) is a special case of GAPC model. We will also discuss the spillover effect of diabetes by, for example, considering the morbidity rates of ailments related to diabetes, since it is believed that diabetes is connected to many metabolic syndrome diseases. We should briefly introduce the Gompertz model and Lee-Cater model before showing the results empirical analysis in the next section.

1) Gompertz model:

Originally, the Gompertz model was for modeling the mortality rates of higher ages and it assumes that the force of mortality μ_x at age x satisfies

$$\mu_x = BC^x \tag{4}$$

where $B > 0$ and $C > 1$ are model parameters. Equation (4) can be converted to the form of central death rate m_x.

2) The Lee-Carter (LC) model:

If m_{xt} denote the central death rate or incidence rate for a person aged x at time t. The LC model assumes that

$$\log(m_{xt}) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \varepsilon_{x,t}, \tag{5}$$

with $\sum_x \beta_x^{(2)} = 1$ and $\sum_t \kappa_t^{(2)} = 0$, $\beta_x^{(i)}$ are age related parameters ($i = 1, 2$), and $\kappa_t^{(2)}$ represents the time related parameter. Note that $\beta_x^{(1)}$ is the general mortality level and $\beta_x^{(2)}$ is the mortality improvement rate at age x, and $\kappa_t^{(2)}$ is usually a linear function in time. The term $\varepsilon_{x,t}$ denotes the error term and is assumed to be white noise, with 0 mean and relatively small variance.

3) Renshaw-Haberman (RH) model (Renshaw and Haberman, 2006):

The RH model can be treated as a version of LC model with an extra cohort component,

$$\ln(m_{xt}) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_{t-x}^{(3)}, \tag{6}$$

where $\sum_x \beta_x^{(2)} = 1, \sum_t \kappa_t^{(2)} = 0, \sum_x \beta_x^{(3)} = 1, \sum_{x,t} \gamma_{t-x}^{(3)} = 0,$ and the parameter $\beta_x^{(i)}$ denotes the average age-specific mortality, $\kappa_t^{(2)}$ represents the general mortality level, and $\gamma_{t-x}^{(3)}$ reflects the cohort-related effect.

4) Cairns-Blake-Dowd (CBD) model (Cairns et al., 2006):

The CBD model was designed to model mortality rates of higher ages and to deal with the longevity risk in pensions and annuities. For the CBD model, it assumes that the mortality rates satisfy

$$\text{logit}(m_{xt}) = \log\frac{m_{xt}}{1-m_{xt}} = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)}, \tag{7}$$

where the parameters are $\beta_x^{(i)}$ and $\kappa_t^{(i)}$ ($i = 1, 2$) denote the average age-specific mortality and the general mortality levels. If we assume $\beta_x^{(1)} = 1$ and $\beta_x^{(2)} = x - \bar{x}$, then the model has a simple parametric form:

$$\text{logit}\,(m_{xt}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) \,. \tag{8}$$

5) The Age-Period-Cohort (APC) model:

The Age-Period-Cohort (APC) model is a popular tool for modelling disease incidence and mortality in epidemiology. Heuristically speaking, if we consider the notion of Analysis of Variance, the LC model considers the effects of Age and Age×Period (Interaction), while the APC model considers three main effects: Age, Period, and Cohort:

$$\ln(m_{xt}) = \alpha_x + \kappa_t + \gamma_{t-x}, \tag{9}$$

where $\sum_{c=t-x} \gamma_c = 0, \sum_c c\gamma_c = 0$ (Cairns et al., 2009).

Section 3: Empirical Data Analysis

As mentioned previously, there are no unified standards for judging if a person is with diabetes in Taiwan. For the catastrophic illness (CI), Taiwan's NHI has a concrete and rigorous standard and review process. This helps the insurance industry prevent insurance dispute and develop the CI related products (Yue et al., 2018). The CI products is one of the most popular health products in Taiwan now. We should use the empirical results in this section to provide suggestions of choosing possible criteria of diabetes. Since the sample size is fairly large, we expect that, if the criteria used are reasonable, the prevalence, incidence and mortality rates should be consistent and stable between ages and years, as well as satisfying the experts' (such as doctors') standards. Of course, consistency is only one of the required factors and we should also take the moral hazard into consideration in order to reduce the insurance risk.

Note that the disease records in the NHRD follow the format ICD-9 (International Classification of Diseases, 9th Revision) and thus we will use the ICD code to determine if people are the diabetes patients. In particular, we are interested in the type 2 mellitus diabetes (T2DM), accounting for 95% of diabetes cases in Taiwan. (Source: Health Promotion Administration, Ministry of Health and Welfare) The ICD code of T2DM is 250, and the cases of type 1 diabetic (ICD code 250×1, 250×3) are excluded in this study. But, unlike the CI patients, we cannot simply rely on the ICD code to decide the diabetes patients, since the ICD code does not reveal the severity of diseases. We need to include other conditions, similar to the criteria of judging diabetes in Table 1.

The use of a prescription drug is often included in deciding diabetes. Unfortunately, there are concerns in the data quality of prescription drug records and some people may even use diabetes prescription drugs for weight loss, according to our consultation with doctors. Another reason for not using prescription drug in deciding diabetes is that patients may seek alternative treatments. Garrow (2006) pointed out that there are 46.7% of diabetes patients using complementary and alternative medicine. Also, we cannot develop a complete list of medicines used for diabetes patients. Thus, we consider another type of record for chronic diseases such as diabetes and it is called refillable (continuous) prescriptions for patients with chronic illness (RP). The idea behind RP is to reduce the number of hospital visits for the chronic patients. The patients with RP are qualified to have three months of prescription drugs, which means that they only need to visit doctors every three months. The RP has been enforced since 2003 and it significantly reduces the number of hospital visits. Right now, there are about 100 types of illness which can have RP based on the doctors' recommendations.

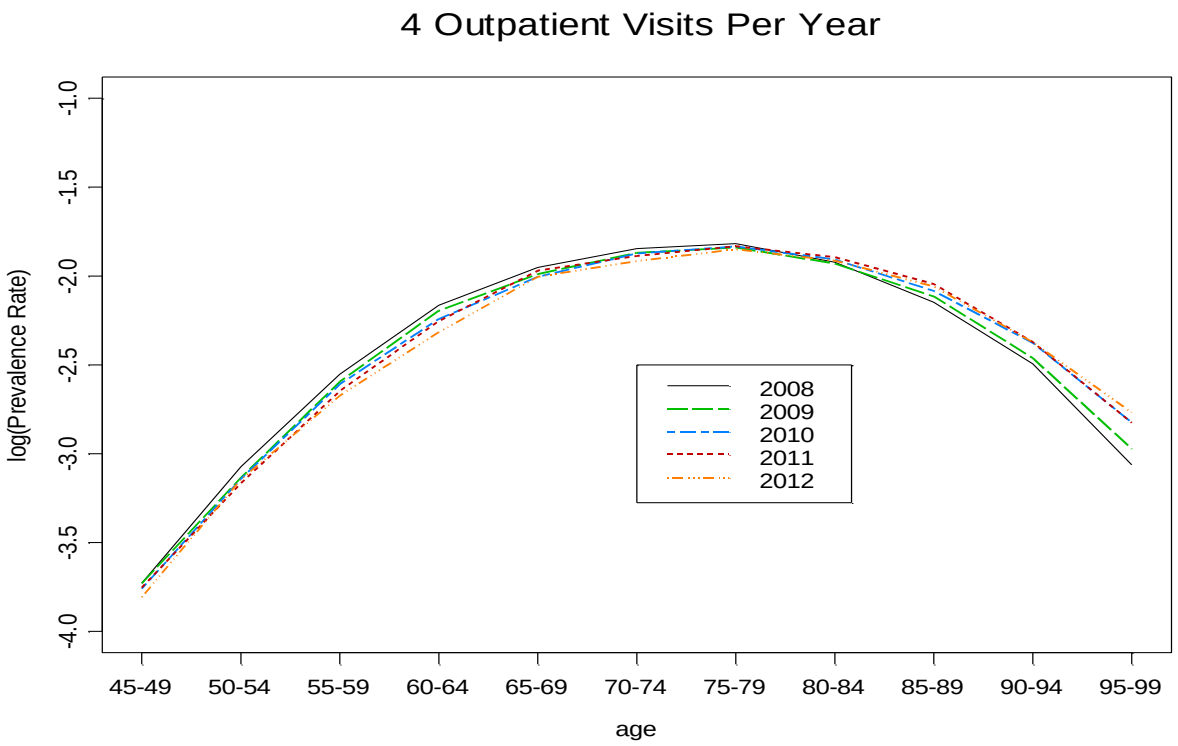
Because diabetes usually is not immediately fatal, patients often stop visiting doctors or forget to take medication when the symptoms of diabetes (such as hyperglycemia) relieve. This would create difficulty in calculating the incidence rates of diabetes, i.e., failure in identifying first-time patients. Thus, we adapt rules, similar to the idea of a

washout period, used in Taiwan's health insurance products. In Taiwan, in order to reduce the possibility of moral hazard and over-estimation, usually a two-year observation (or probationary) period is used. For example, if consumers want to purchase cancer insurance, then they need to provide their medical records in the last two years, showing that they don't have cancer. The idea of a two-year observation period will be imposed to determine the incidence rate of diabetes.

With the two-year observation period, we can calculate the incidence rates based on the numbers of outpatient visits and RP's. Intuitively, more outpatient visits would reduce the possibility of false positive. Lin et al. (2005) showed that the accuracy of overall diabetes diagnosis in NHI claims data was 74.6% and it increased to 96.1% for the cases with 4 or more outpatient visits. Using the criterion of 4 outpatient visits per year, we found that the prevalence rates of T2DM are very stable in 2008-12 and look like a reverse U-shaped curve, reaching the peak around age 80 (Figure 2). We also compute the prevalence rates of T2DM using the criteria 2 and 3 outpatient visits per year. As expected, they are higher than those of 4 outpatient visits per year, but the results vary a lot for different years. Thus, we should focus on the results of 4 outpatient visits per year.

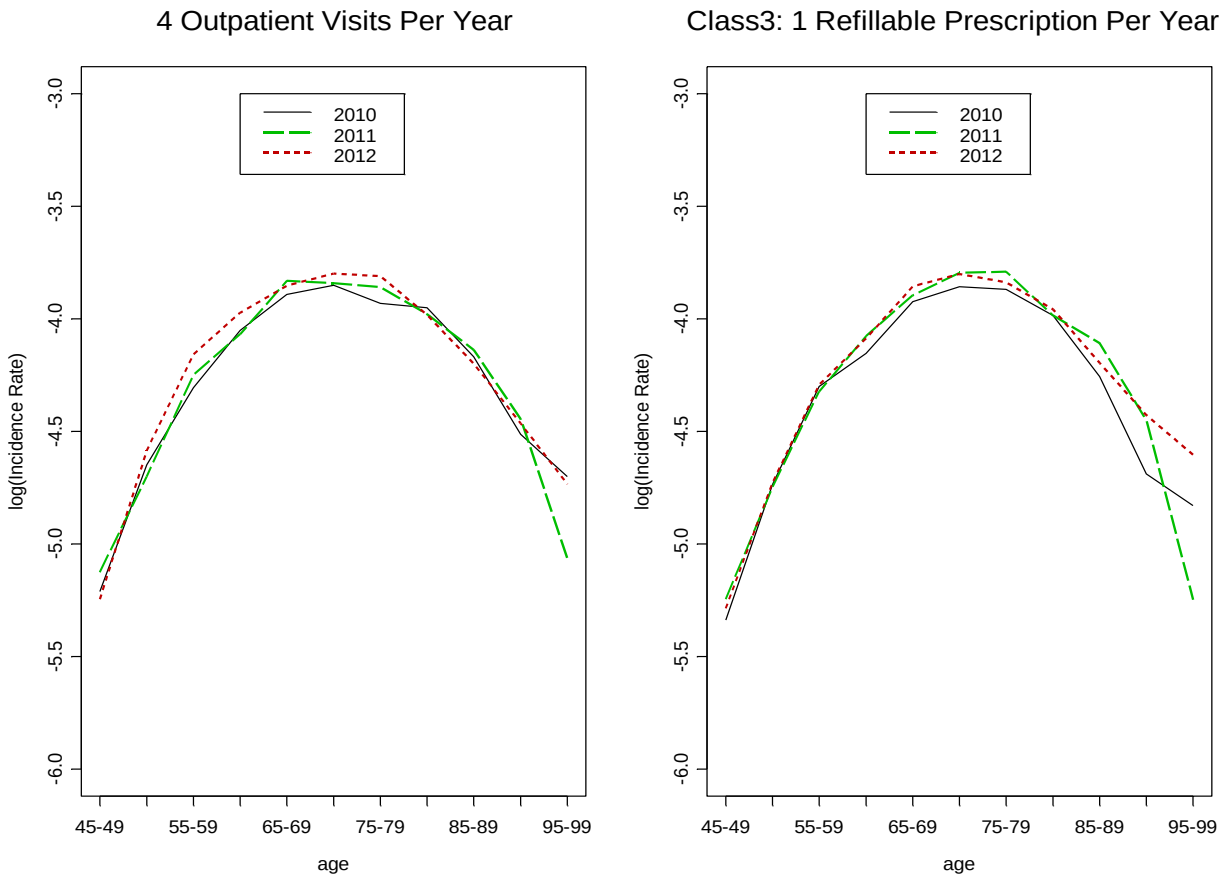
Figure 2

PREVALENCE RATES OF 4 OUTPATIENT VISITS PER YEAR (2008-2012)



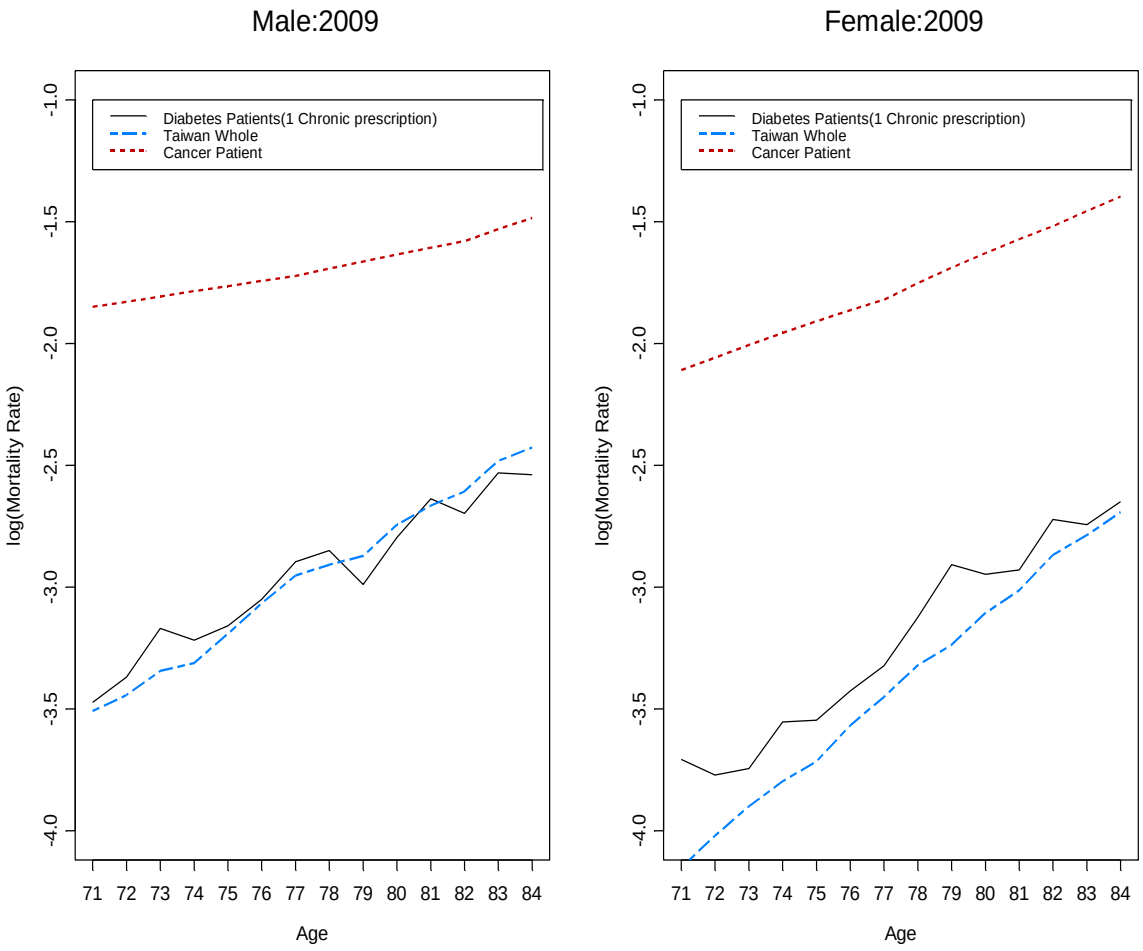
The incidence rates using the criterion of 4 outpatient visits per year are also very stable and like a reverse U-shaped curve as well (left panel of Figure 3). The incidence rates reach the peak 2% around age 75. We also consider the incidence rates of T2DM using the criterion of one RP per year and the results are very interesting (right panel of Figure 3). The incidence rates based on 4 outpatient visits per year and one RP per year are almost identical, and it seems that the patients with the diabetes refillable prescription are likely to have diabetes. Since the RP is easy to confirm, we should use one RP per year to determine the diabetes patients for the rest of this study.

Figure 3
INCIDENCE RATES OF T2DM (2010-12 FEMALE)
(LEFT: 4 OUTPATIENT VISITS PER YEAR; 1 RP PER YEAR)



We also calculate the mortality rates of diabetes patients, according the five criteria of death judgement in the last section. We should use the results of 2009 as a demonstration and discuss the mortality's trend in the next section. Figure 4 shows the age-specific mortality rates of diabetes patients ages 71-84, comparing to those of Taiwan's population and Taiwan's cancer patients. As expected, the male mortality rates are higher and, since the diabestes is a chronic disease, the mortality rates of diabetes patients are much smaller than those of cancer patients. However, surprisingly, the mortality rates of diabetes patients are simialr to those of Taiwan average people, only the female mortality rates are slightly larger. This result is somewhat different to those of previous studies, where older diabetes patients have higher mortality rates (Manderbacka et al., 2011; Kirkman et al., 2012; Castro-Rodríguez et al., 2016). Of course, the mortality rates of diabetes patients are related to the definition of diabetes. Our definition is connected to the willingness of constantly receiving medication and these patients may have lower mortality rates comparing with diabetes patients cannot or choosing not to take regular medication.

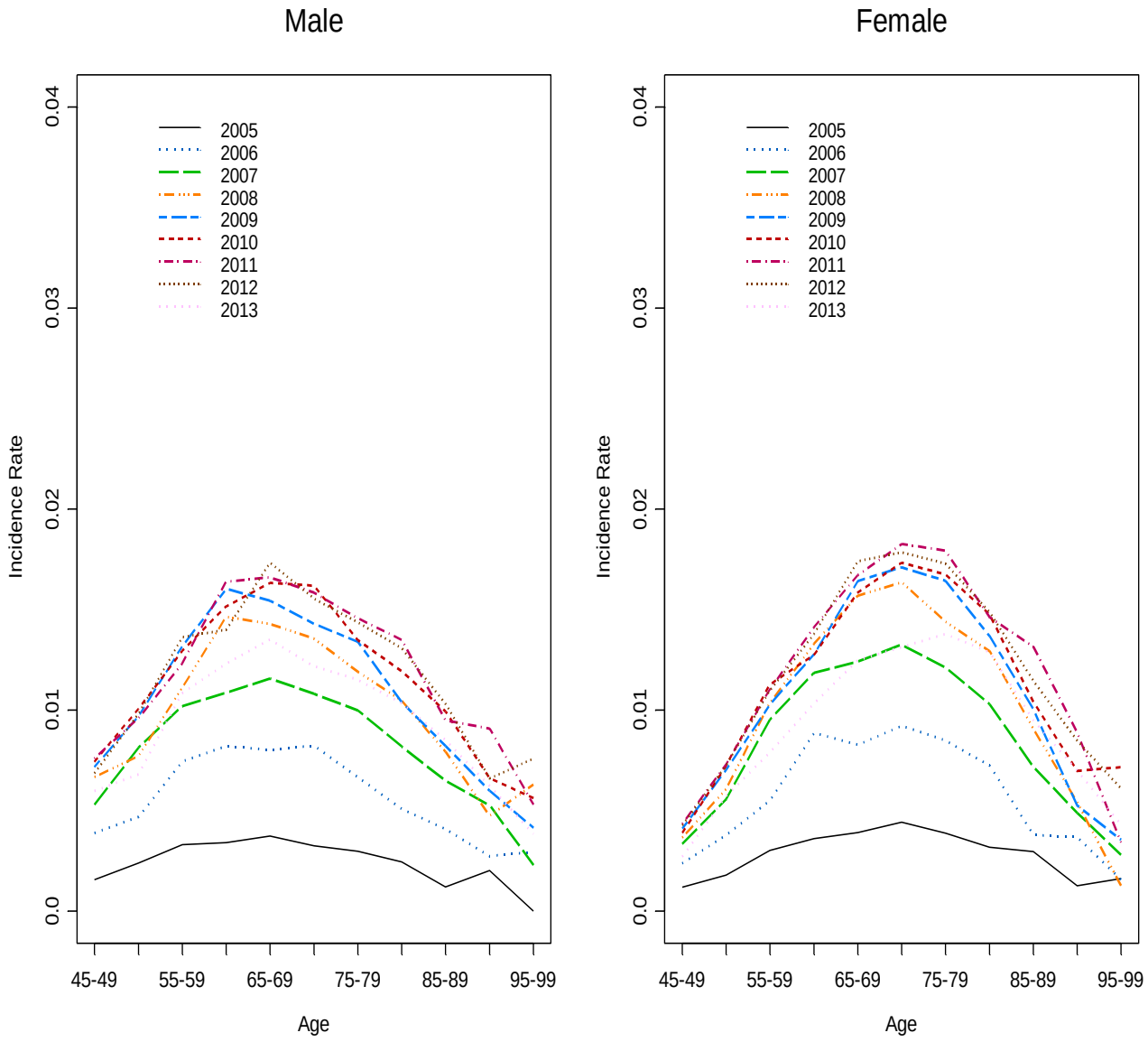
Figure 4
MORTALITY RATES OF DIABETES PATIENTS (2009)



Section 4: Model Evaluation and Applications

In this section, we will use stochastic models to forecast the incidence rates and mortality rates of diabetes and evaluate the impact of diabetes due to the population ageing. The diabetes patients are those who have one RP per year. Also, we consider the Generalized Age-Period-Cohort (GAPC) models in this study, with the help of graduation methods if the population sizes are small. We should choose the models which can well describe the patterns of diabetes patients, using them to design diabetes products and to provide suggestions for policy planning.

Figure 5
INCIDENCE RATES OF DIABETES (2005-2013)



Before evaluating the models, we first examine the data quality. As mentioned earlier, the consistency is one of our concerns in choosing the criteria of diabetes, and it can also be used in selecting the data period. For example, Figure 5 shows the diabetes incidence rates of Taiwan male and female in 2005-2013. The results of first three years (2005-2007) are obviously different than other years. Taiwan's NHI enforced the RP in 2003, in order to ease the burden of patients with chronic diseases, and the patients only need to have outpatient visit every 3 months. It takes time to let

all people get used to the new rule and thus we tend to think that the results in the early years, such as 2005 and 2006, should not be included in the model evaluation.

Nonetheless, to avoid misjudgment, we still conduct the model evaluation with those data in the early years. In specific, three data periods are considered: 2005-2013, 2007-2012, and 2008-2013, and we should verify if the results of model fitting are influenced by the data in the early years. Also, two age settings are considered: single-age and 5-year age groups. For the single-age case, the age range is 45, 46, …, 89, while the age range is 45-49, 50-54, …, 95-99 for the 5-year age case. Due to the population size, we do not consider ages 90 and beyond for the single-age case. Note that, in addition to the GAPC models, we can combine graduation methods such as the PSMR method with mortality models (Wang et al., 2018) and we should use the combination PSMR and LC model as a demonstration. The GAPC models are conducted using the R package StMoMo.

Table 2
FITTING MAPE OF INCIDENCE RATES (5-AGE GROUPS: 45-99)

	2005-2013		2007-2012		2008-2013		Average
	Male	Female	Male	Female	Male	Female	
LC	49.87	9.80	6.12	7.97	6.04	10.00	14.97
APC	6.40	6.71	4.49	7.03	4.07	4.89	5.60
PSMR	144.64	9.46	6.98	9.72	5.91	9.47	31.03
PSMR+LC	147.49	11.69	7.56	9.92	6.29	10.13	32.18
CBD	338.46	83.92	43.80	88.21	41.59	84.65	113.44
RH	---	78.38	68.55	68.48	37.09	74.94	65.49

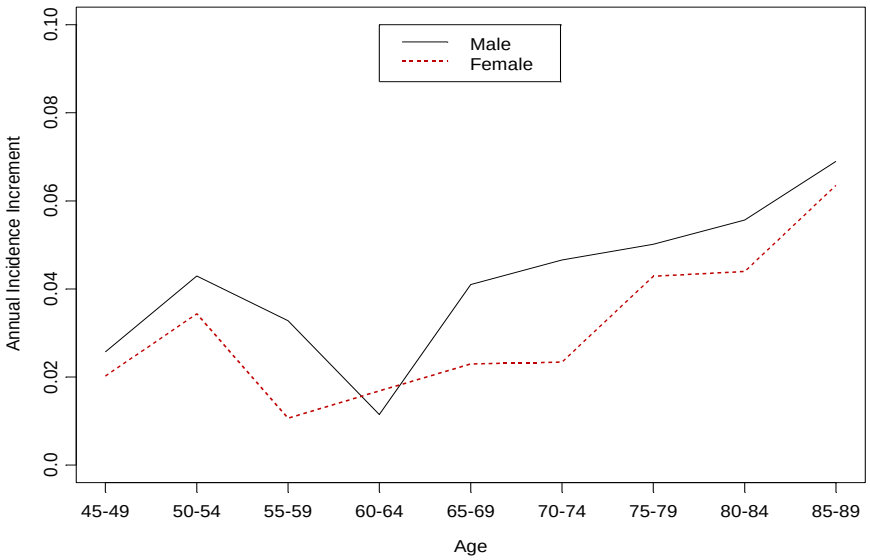
Note: In 2005, the incidence number of 95-99 was 0 and the RH model does not converge.

We first discuss the 5-year age case and Table 2 lists the fitting results with respect to MAPE for all three time periods. The APC model has the best fitting performance in all three periods and it is not influenced by the data in the early years. If we omit the data in 2005 and 2006, the LC, APC, PSMR, and PSMR+LC models all have satisfactory fitting results. Maybe we should only consider ages 45-89 and thus the fitting results of single-age case are slightly different. Except the CBD model, all models have satisfactory fitting results for the single-age case (Table 3). Among all models, RH, LC and APC have the best performance.

Table 3
FITTING MAPE OF INCIDENCE RATES (SINGLE-AGE: 45-89)

	2005-2013		2007-2012		2008-2013		Average
	Male	Female	Male	Female	Male	Female	
LC	10.37	10.16	7.80	7.22	8.22	7.65	8.57
APC	10.70	9.95	7.38	7.73	7.99	7.71	8.58
PSMR	10.87	10.34	8.51	8.39	8.77	8.24	9.19
PSMR+LC	13.21	12.46	8.95	8.66	8.95	8.62	10.14
CBD	30.24	40.40	26.34	38.43	25.87	37.23	33.08
RH	9.57	8.71	6.09	6.18	6.11	6.18	7.14

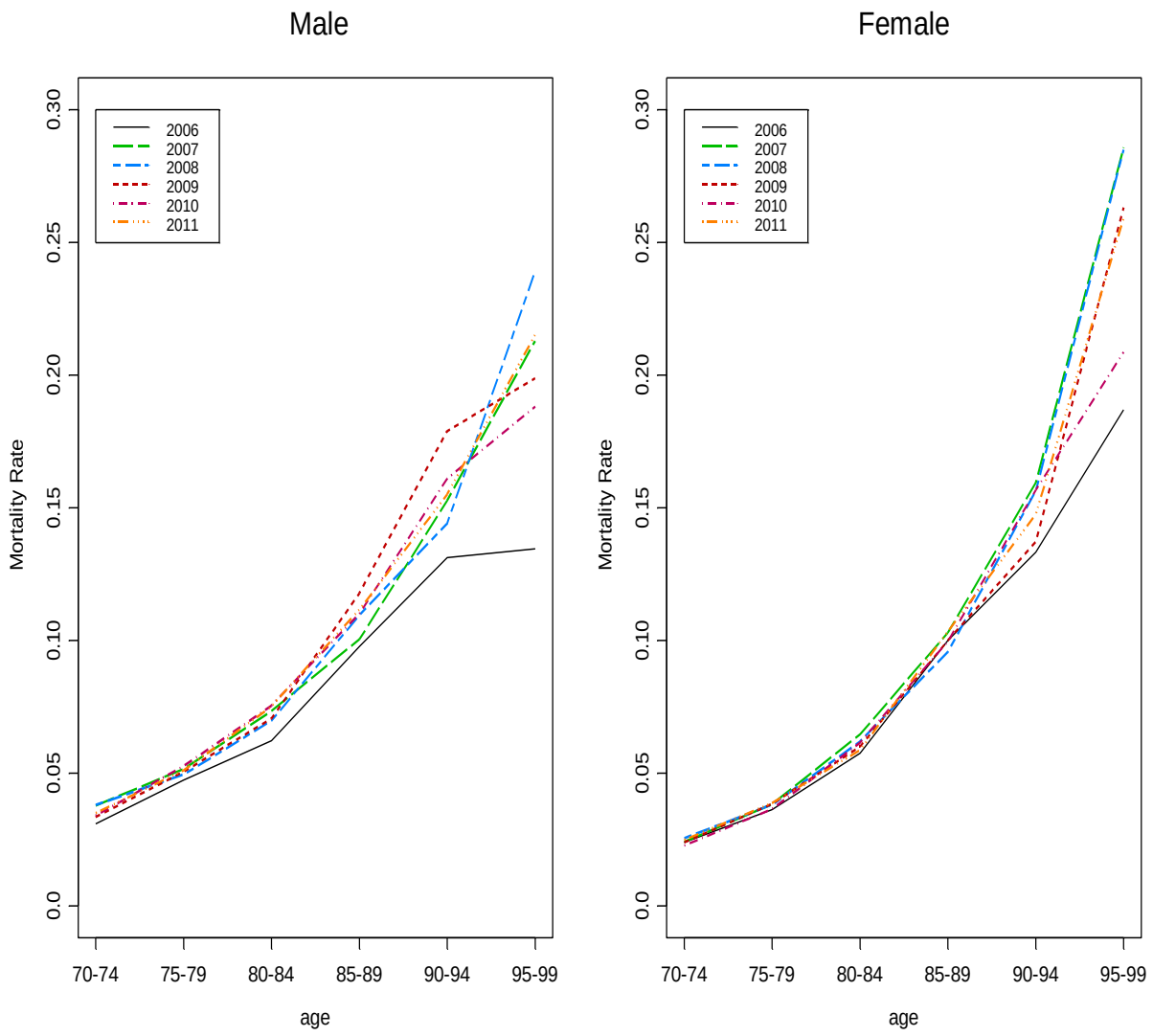
Figure 6
ANNUAL INCREMENT OF DIABETES INCIDENCE RATES (LC MODEL)



Since the LC model is frequently used in prediction, we will use it to demonstrate the trend of diabetes incidence rates. For the 5-year age group case and years 2008-2013, we can use the estimates of LC model parameters to acquire the annual increment of incidence rates. In particular, we assume $\kappa_t^{(2)} = a + bt$ in equation (5) and thus the annual increment of incidence rate at age *x* is $\beta_x^{(2)} \times b$ (Yue et al., 2018). Figure 6 shows the annual increments of diabetes incidence rates at all age groups. The annual increments are smaller for younger groups and generally increase with age, reaching to around 6% at ages 85-89. The scale of annual increments is worth to pay attention. But, since the estimates are based on a few years of data, we need to collect more data before making any conclusions.

Modelling the mortality rates of diabetes patients can follow the same process, similar to that of diabetes incidence rates. Due to the nature of our data, we can only estimate the mortality rates in 2006-2011. Our data were drawn in 2005, meaning that the individuals in the sample were still alive in 2005. On the other hand, the death criteria used require two-year washout and thus we cannot estimate the mortality rates in 2012 and 2013. Figure 7 shows the 5-year age mortality rates of elderly diabetes patients in 2006-2011. The mortality rates increase with age and females have larger increment. However, the mortality rates do not have an obvious trend, and only those in 2006 are the lowest. Still, we should verify if the GAPC models can capture the trend of mortality rates.

Figure 7
MORTALITY RATES OF DIABETES PATIENTS (2006-2011)



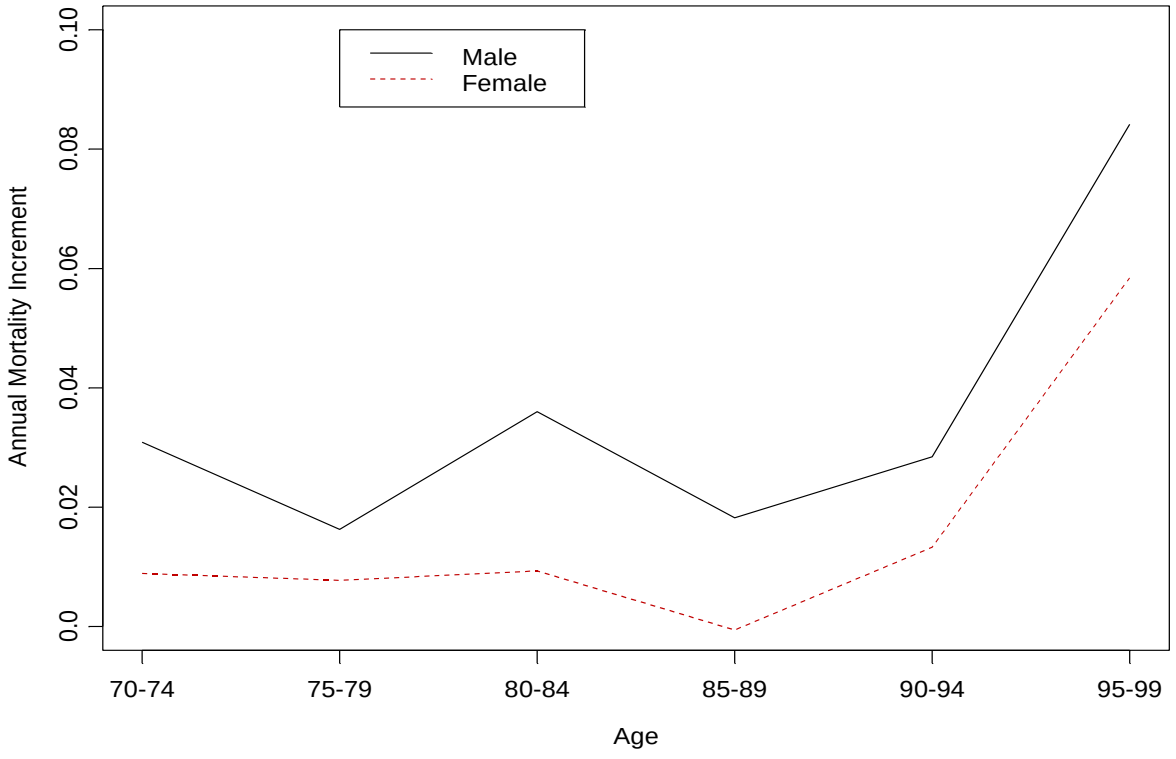
Similar to the case of incidence rates, we also compare the fitting performance of diabetes patients for all models. However, due to the consideration of sample size, the age ranges for the 5-year age group and single-age cases are 70-74, 75-79, …, 95-99 and 71, 72, …, 89, respectively. Table 4 shows the fitting errors with respect to MAPE for all models. Except the RH model, all models have fairly accurate fits, with average MAPE about 5%. Again, for the 5-year age group case in 2006-2011, we can use the estimates of LC model parameters to acquire the annual increment of mortality rates for the elderly diabetes patients (Figure 8). The average annual increments are 3.6% and 1.6% for male and female, respectively. If we only consider age 70-89, then the annual increments would be more stable but slightly reducing to 2.5% and 0.6% for male and female, respectively.

Table 4
FITTING MAPE OF MORTALITY RATES

	5-year age group (70-99)		Single-age (71-89)		Average
	Male	Female	Male	Female	
LC	3.76	2.74	5.48	4.75	4.18
APC	3.27	2.64	4.86	5.13	3.97
PSMR	4.60	4.41	6.37	5.75	5.28
PSMR+LC	4.94	4.41	6.42	5.75	5.38
CBD	5.56	4.41	6.59	5.95	5.63
RH	24.74	11.51	3.84	4.14	11.06

The results of model evaluation for the diabetes incidence and mortality rates suggest that the LC and APC models are preferred. Both models indicate that the incidence and mortality rates will increase with time, but the increments of mortality rates are much smaller. As a result, we expect that the number of diabetes patients will grow with time, especially for the elderly. The diabetes patients usually use more medical resources than those without diabetes, and thus more diabetes patients indicate more spending in Taiwan's NHI. Taiwan government needs to figure out some solutions in order to retain the sustainability of NHI and other social insurance.

Figure 8
ANNUAL INCREMENT OF MORTALITY RATES (LC MODEL)

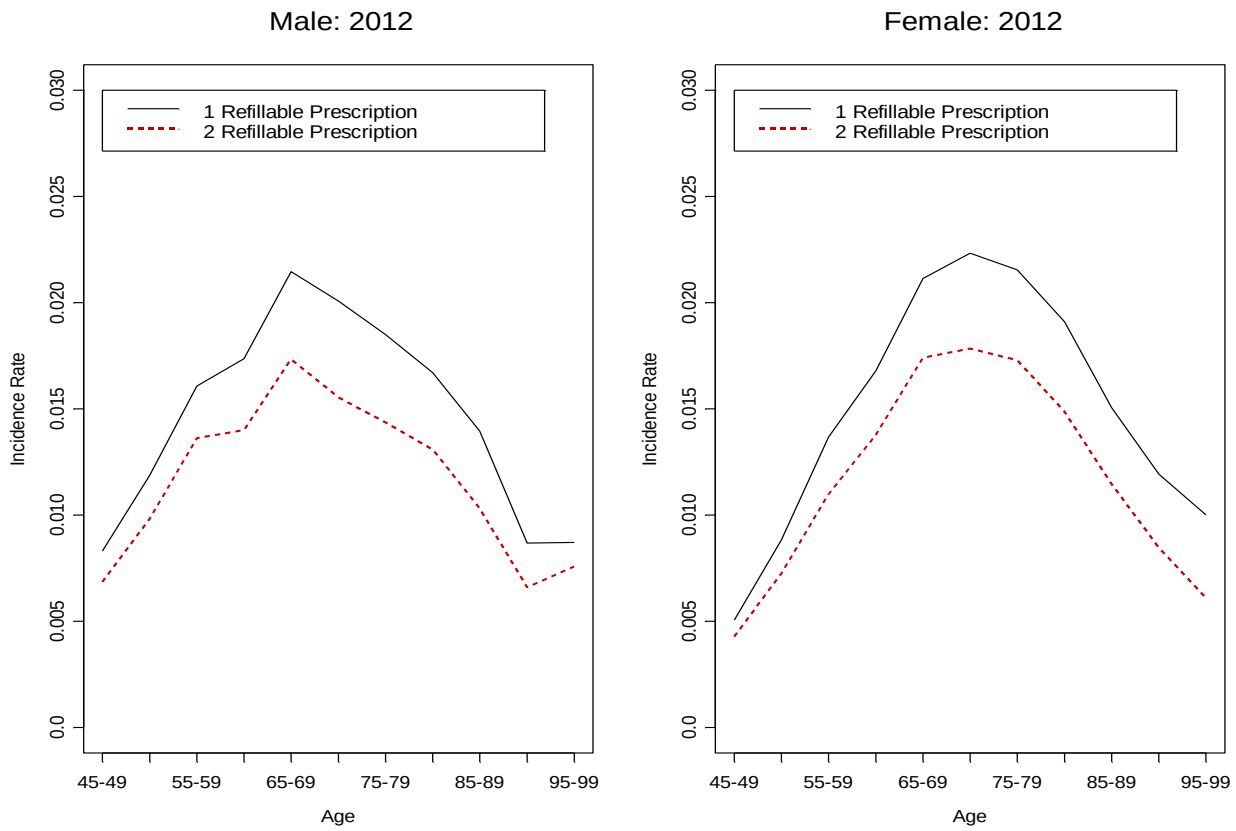


Section 5: Conclusion and Discussion

Population ageing is a common demographic trend in the 21st century. More and larger proportion of elderly people are expected due to the prolonging life expectancy. Chronic diseases, such as metabolic syndrome, replacing infectious and acute diseases, become the main health concern. Diabetes is one of the major metabolic syndrome diseases, but many people don't know they have this diabetes. Unlike stroke and cardiovascular disease, diabetes has not received much attention in the past, probably because it does not directly lead to death. It receives more attention in recent years since studies showed that diabetes is associated with many diseases and more people that have diabetes. We expect that diabetes will have a larger influence on the health and medical expenditure of Taiwan people, and thus we use Taiwan's National Health Insurance Research Database to explore the trend of type 2 diabetes.

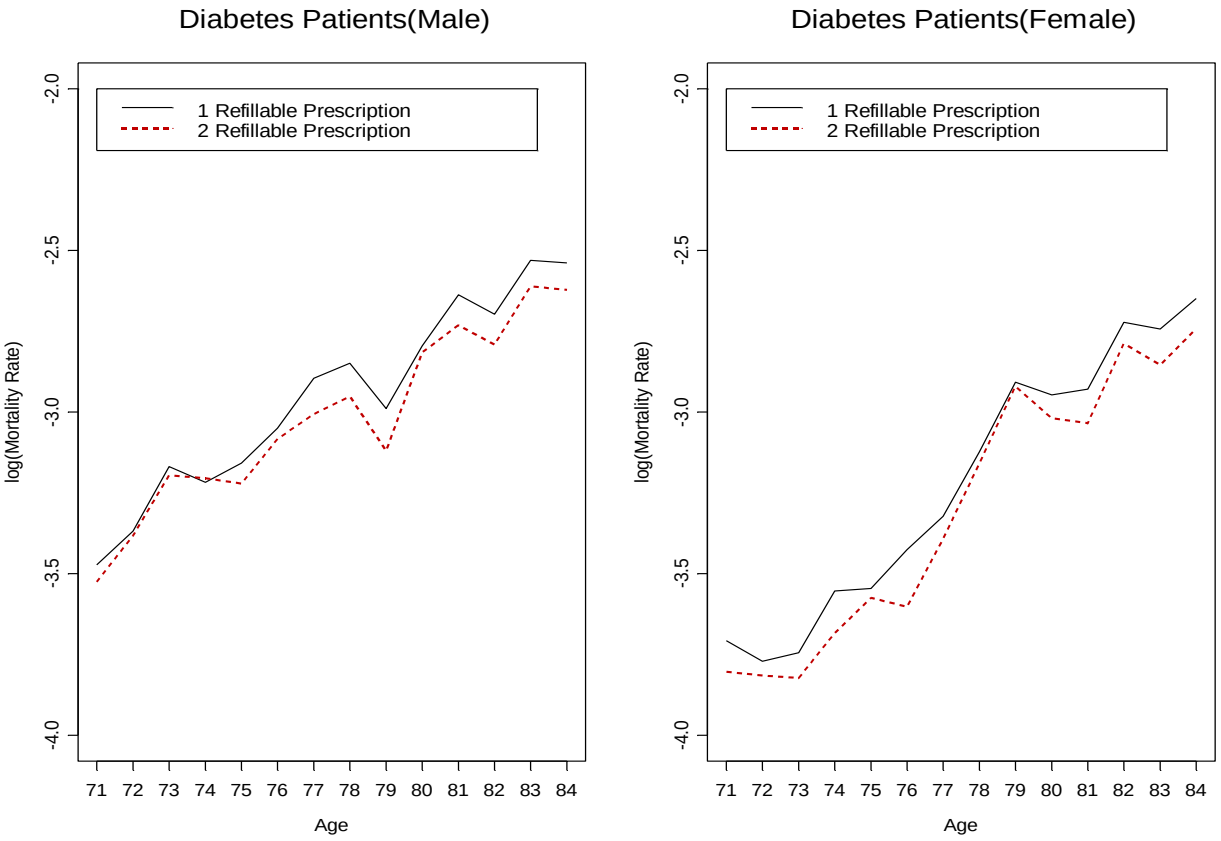
In this study, we evaluated judging criteria of diabetes and calculate its incidence rates and mortality rates. Using the refillable (continuous) prescriptions for patients with chronic illness (RP), we obtained stable incidence rates and mortality rates, which can be used to design diabetes insurance products. Among all GAPC models for fitting the incidence rates and mortality rates, the APC model had the smallest MAPE errors and the LC model was also a feasible choice. If we used the LC model to measure the time trend, we found that the incidence rates of type 2 diabetes gradually increased with time while the elderly mortality rates of diabetes patients change with a stable path.

Figure 9
DIABETES INCIDENCE RATES: 2RP VS. 1RP (2012)



Also, as mentioned in the previous section, the incidence rates (and possibly the mortality rates) of diabetes depend on the judging criteria. The trend of these rates can also be very different accordingly. For the readers' reference, we compared the results of using 1RP per year and 2RP year as the criteria. Figure 9 are the incidence rates of two different criteria in 2012. Interestingly, no matter for the male or the female, the diabetes incidence rates of 2RP are about 20% less than those of 1RP. The results of diabetes mortality rates showed similar pattern. Since we cannot compute the mortality rates in 2012, we compared the mortality rates in 2009 instead. The mortality rates of diabetes patients using 2RP were about 7% less than those using 1RP (Figure 10). It seems that using RP can produce fairly stable estimates for the incidence rates and mortality rates. As for the gaps between 1RP and 2RP, we can use methods such as the spill-over effect, similar to the car insurance's no claim discount (NCD) method, to design medical policies with discount for the insured who continue receiving treatments.

Figure 10
DIABETES MORTALITY RATES: 2RP VS. 1RP (2009)



The ageing population and unhealthy lifestyle can lead to changes in the main causes of death in many countries like Taiwan (Wild et al., 2004). It seems that diabetes is a good indicator for the degree of unhealthy level. According to National Diabetes Statistics Report (2017), the complications of U.S. adults (aged 18 and beyond) diagnosed with diabetes in 2011-2014 include overweight and obesity, physical inactivity, high blood pressure, high cholesterol, and high blood glucose, which are related to metabolic syndrome diseases. Thus, we think that the increasing incidence rates of diabetes in Taiwan and U.S. also indicates increasing medical demands and expenditures, not restricted to the number of deaths. In order to maintain the sustainability of NHI, we suggest that Taiwan's government provides more incentive for diabetes patients to pay extra attention to their health, such as free health exam every two or three years.

For commercial insurance, diabetes can be deemed a sign of potential health problems, and thus we treat it as a risk factor (i.e., those with diabetes as part of the substandard group) for insurance products. Alternatively, we can also adopt the idea of insured product options in designing diabetes products. For example, consumers can purchase the diabetes option. When the insured is diagnosed with diabetes, instead of receiving a benefit payment, he/she can purchase new policies using the standard price rate. This is feasible for life insurance products, since we can acquire mortality rates of diabetes patients. Of course, for health insurance products, we need further study and more information regarding the relation of diabetes and other health conditions.

References

Cairns, A.J.G., Blake, D., and Dowd, K. (2006), "A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration." Journal of Risk and Insurance, 73 (4): 687–718.

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., and Balevich, I. (2009), "A quantitative comparison of stochastic mortality models using data from England and Wales and the United States." North American Actuarial Journal 13(1): 1-35.

Chang, C.H., Shau, W.Y., Jiang ,Y.D., Li, H.Y., Chang, T.J., Sheu, W.H., Kwok, C.F., Ho, L.T., and Chuang, L.M. (2010), "Type 2 Diabetes Prevalence and Incidence among Adults in Taiwan during 1999-2004: A National Health Insurance Data Set Study", Diabet Med. 27(6): 636-43.

Chen, C. , Yue, C.J., and Tsai, W. (2015), "The Effect of the 921 Chi-Chi Earthquake on the Mortality Risk of the Middle-Aged and Elderly", Journal of Population Studies 50, 61-99.

Chiang, J., Lin, C., Wang, C. Koo, M., and Kao, Y. (2017), "Cancer studies based on secondary data analysis of the Taiwan's National Health Insurance Research Database", Medicine 96(17): e6704.

Castro-Rodríguez, M., Carnicero, J.A., Garcia-Garcia, F.J., et al., (2016), "Frailty as a major factor in the increased risk of death and disability in older people with diabetes", Journal of the American Medical Directors Association 17: 949-55.

Cockram, C.S. (2000), "Diabetes Mellitus: Perspective from the Asia-Pacific Region", Diabetes Research and Clinical Practice 50 (Suppl. 2), S3–S7.

Danaei, G., Finucane, M.M., Lu, Y., Singh, G.M., Cowan, M. J., Paciorek, C.J., Lin, J.K., Farzadfar, F., Khang, Y.H., Stevens, G.A., Rao, M., Ali, M.K., Riley, L.M., Robinson, C.A., and Ezzati, M. (2011), "National, Regional, and Global Trends in Fasting Plasma Glucose and Diabetes Prevalence since 1980: Systematic Analysis of Health Examination Surveys and Epidemiological Studies with 370 Country-years and 2.7 Million Participants", Lancet 378: 31-40.

da Rocha Fernandes, J., Ogurtsova, K., Linnenkamp, U., Guariguata, L., Seuring, T., Zhang, P., Cavan, D., and Makaroff, L.E. (2016), "IDF Diabetes Atlas Estimates of 2014 Global Health Expenditures on Diabetes", Diabetes Res Clin Pract. 117: 48-54.

Dwyer-Lindgren, L., Mackenbach, J.P., van Lenthe, F.J., Flaxman, A.D., and Mokdad, A.H. (2016), "Diagnosed and Undiagnosed Diabetes Prevalence by County in the U.S., 1999-2012", Diabetes Care 39(9): 1556-62.

Garrow, D., and Egede, L.E. (2006), "National Patterns and Correlates of Complementary and Alternative Medicine Use in Adults with Diabetes", Journal of Alternative and Complementary Medicine 12(9): 895-902.

Holman, N., Forouhi, N.G., Goyder, E., and Wild, S.H., (2011), "The Association of Public Health Observatories (APHO) Diabetes Prevalence Model: Estimates of Total Diabetes Prevalence for England, 2010-2030", Diabetic Medicine 28(5): 575-82.

International Diabetes Federation. (2017), IDF Diabetes Atlas Eighth Edition 2017. Available From: https://www.idf.org/e-library/epidemiology-research/diabetes-atlas/134-idf-diabetes-atlas-8th-edition.html. [Accessed 4 September 2019]

Jiang, Y.D., Chang, C.H., Tai, T.Y., Chen, J.F., and Chuang, L.M., (2012), "Incidence and Prevalence Rates of Diabetes Mellitus in Taiwan: Analysis of the 2000-2009 Nationwide Health Insurance Database", Journal of the Formosan Medical Association 111(11): 599-604.

Kirkman, M.S., Briscoe V.J., Clark N., et al., (2012), "Diabetes in older adults", Diabetes Care 35:2650-64.

Lee, W.C. (2003), "A Partial SMR Approach to Smoothing Age-Specific Rates", Annals of Epidemiology 13(2): 89–99.

Lee, R.D., and Carter, L.R. (1992), "Modeling and Forecasting US Mortality", Journal of the American Statistical Association 87 (419): 659–671.

Lewis, E.B. (1982), "Control of Body Segment Differentiation in Drosophila by the Bithorax Gene Complex", Embryonic Development, Part A: Genetics Aspects, Edited by Burger, M. M. and R. Weber. Alan R. Liss, New York, 269–288.

Li, H.Y., Jiang, Y.D. , Chang, C.H., Chung, C.H., Lin, B.J., and Chuang, L.M., (2012), "Mortality Trends in Patients with Diabetes in Taiwan: A Nationwide Survey in 2000-2009", Journal of the Formosan Medical Association 111(11): 645-650.

Lin, C.C., Lai M.S., Syu C.Y., Chang S.C., and Tseng F.Y., (2005), "Accuracy of Diabetes Diagnosis in Health Insurance Claims Data in Taiwan", Journal of the Formosan Medical Association 104: 157–163.

Lin, C.C., Li ,C.I., Hsiao, C.Y., Liu, C.S., Yang, S.Y., Lee, C.C., and Li, T.C., (2013), "Time Trend Analysis of the Prevalence and Incidence of Diagnosed Type 2 Diabetes Among Adults in Taiwan from 2000 to 2007: A Population-Based Study", BMC Public Health 13: 318. doi: 10.1186/1471-2458-13-318.

Lin, T., Chou, P., Tsai, S.T., Lee, Y.C., Tai, T.Y. (2004), "Predicting Factors Associated with Costs of Diabetic Patients in Taiwan", Diabetes Research and Clinical Practice 63: 119–125.

Lee, H.T., Lai, C.C. Chen W.C., Wu S.H., Lin, H.I (2016), "Increased Lung Cancer Risk among Diabetic Patients--A Nationwide Population-Based Study", Fu-Jen Journal of Medicine 14(4): 175–184.

Lin, W.H., Hsu, C.H., Chen, H.F., Liu, C.C., and Li, C.Y., (2015), "Mortality of Patients with Type 2 Diabetes in Taiwan: A 10-year Nationwide Follow-Up Study", Diabetes Research and Clinical Practice 107(1): 78-86.

Manderbacka, K., Peltonen, R., Koskinen, S. and Martikainen, P., (2011), "The Burden of Diabetes Mortality in Finland 1988-2007 - A Brief Report", BMC Public Health 11: 747.

NCD Risk Factor Collaboration (NCD-RisC) – Africa Working Group., (2017), "Trends in Obesity and Diabetes Across Africa from 1980 to 2014: An Analysis of Pooled Population-Based Studies", The International Journal of Epidemiology 46(5): 1421-1432.

Renshaw, A. E., and S. Haberman., (2006), "A Cohort-Based Extension to the Lee-Carter Model for Mortality Reduction Factors", Insurance: Mathematics and Economics 38(3): 556–570.

Shaw, J.E., Sicree, R.A., Zimmet, and P.Z., (2010), "Global Estimates of the Prevalence of Diabetes for 2010 and 2030", Diabetes Research and Clinical Practice 87(1): 4-14.

United States Renal Data System., (2016), "USRDS Annual Data Report: Epidemiology of Kidney Disease in the United States Volume 2: ESRD in the United States", https://www.usrds.org/2016/download/v2_ESRD_16.pdf.

Villegas, A.M., Millossovich, P. and Kaishev, V. K. (2016), "StMoMo: An R Package for Stochastic Mortality Modelling",R package version 0.3.5. http://CRAN.R-project.org/package=StMoMo.

Wang, H.C., Yue, C.J. and Chong, C.T., (2018), "Mortality Models and Longevity Risk for Small Populations", Insurance Mathematics and Economics 78: 351-359.

Whiting, D.R., Guariguata, L., Weil, C., and Shaw, J., (2011), "IDF Diabetes Atlas: Global Estimates of the Prevalence of Diabetes for 2011 and 2030", Diabetes Research and Clinical Practice 94(3): 311-21.

Wild, S., Roglic, G., Green, A., Sicree, R., and King, H., (2004), "Global Prevalence of Diabetes: Estimates for the Year 2000 and Projections for 2030", Diabetes Care 27(5): 1047-53.

Yue, J.C., Wang, H. C., Yin-Yee Leong, and Wei-Ping Su (2018), "Using Taiwan National Health Insurance Database to Model Cancer Incidence and Mortality Rates", Insurance Mathematics and Economics 78: 316-324.

Yue, J.C., Chien, Y., and Leong, Y. (2020), "Using National Health Insurance Database for Sampling Survey", Survey Research-Method and Application 44: 97-130.

Zhang P., Zhang X., Brown J., Vistisen D., Sicree R. , Shaw J. , and Nichols G., (2010), "Global Healthcare Expenditure on Diabetes for 2010 and 2030", Diabetes Research and Clinical Practice 87: 293-301.

About The Society of Actuaries

With roots dating back to 1889, the Society of Actuaries (SOA) is the world's largest actuarial professional organizations with more than 31,000 members. Through research and education, the SOA's mission is to advance actuarial knowledge and to enhance the ability of actuaries to provide expert advice and relevant solutions for financial, business and societal challenges. The SOA's vision is for actuaries to be the leading professionals in the measurement and management of risk.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and non-actuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org