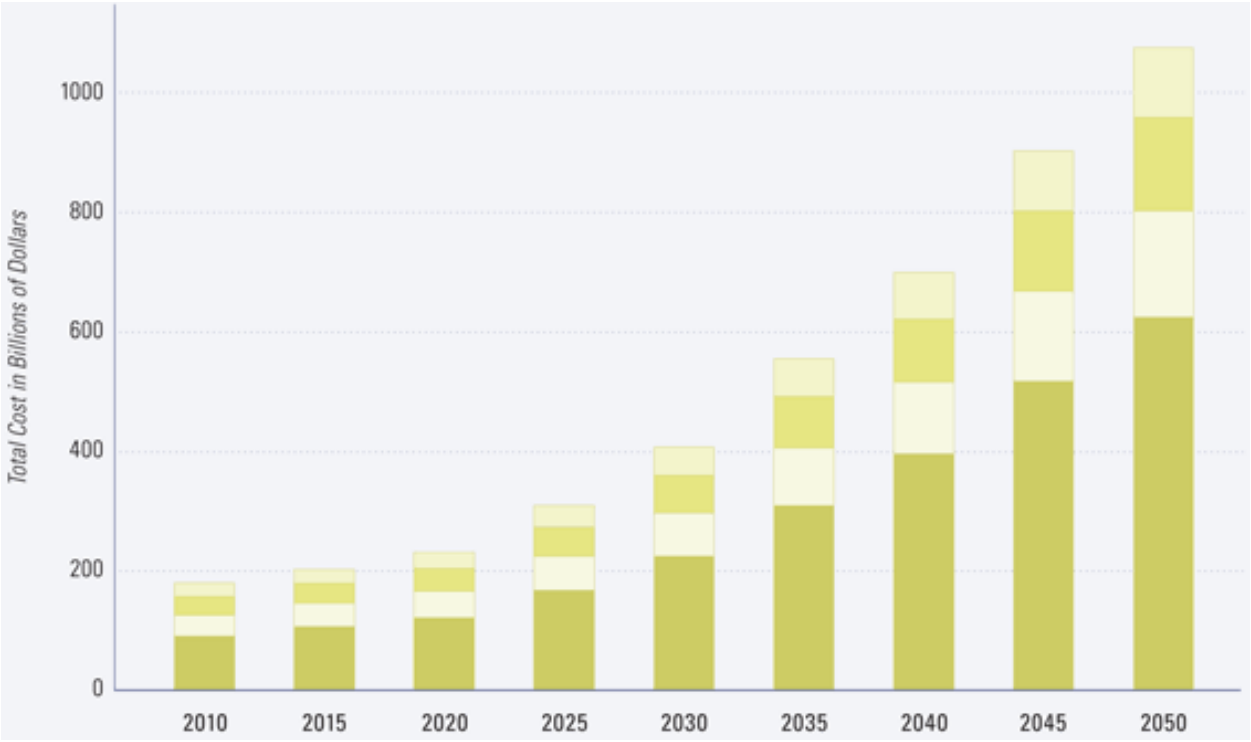


Monte Carlo Simulation of Alzheimer's Disease in the United States: 2010-2060

Michael Blech

Professor Renato Feres

MATH 350

Introduction

Alzheimer's disease is one of the major public health concerns facing the United States over the next 50 years. This progressive disease is currently the sixth-leading cause of death in America, and is expected to cost over $200 billion in 2012 (F&F 2). While these statistics are already troubling, public health experts expect the problem to get worse for two key reasons. First, since advancing age is the main risk factor for Alzheimer's disease, the aging of the baby boom generation will create a disproportionate number of new cases (F&F 10). Second, life expectancy may increase over the coming decades due to improvements in medical technology. If this occurs, Alzheimer's will become more prevalent because the risk of developing the disease increases with age.

Given the burden of Alzheimer's disease on patients, caregivers, and the healthcare system, it is imperative to create models that accurately predict the long-term prevalence of the disease. In 2007, Brookmeyer et al. used Monte Carlo simulation to simulate the global prevalence of Alzheimer's disease through 2050. Based on the simulation, they predicted that the number of Alzheimer's cases would increase from 26.6 million in 2006 to 106.2 million in 2050 (Brookmeyer 190). They also evaluated the potential impact of new interventions that delay the onset and progression of Alzheimer's disease. The simulation described in this paper is based partially on the work of Brookmeyer et al., but with two key differences. First, this simulation focuses specifically on the United States population, and second, the simulation models both prevalence and mortality. Both of these statistics are important because prevalence is an indicator of the burden of the disease at a given time, while mortality rates reflect the impact of the disease over a longer time period.

Alzheimer's disease lends itself well to Monte Carlo simulation for several key reasons. First, the primary risk factor for Alzheimer's disease is advancing age, and it is easy to find

accurate information about the age breakdown of the population. Second, the disease is currently progressive and fatal, so predictions about long-term mortality can be made with a reasonable degree of certainty. Third, the disease is very rare in individuals below the age of 50, so the impact of the disease over the next 50 years can be analyzed without considering birth rates. For these reasons, Monte Carlo simulation is a useful tool for analyzing the future prevalence and mortality of Alzheimer's disease.

Data and Model

The Markov chain in this model has 146 states. There are 140 transient states, each of which is characterized by three pieces of information: (1) Age group, (2) gender, and (3) stage of Alzheimer's disease. The age groups are divided into 5-year intervals, and there are 22 age groups in total. The stage of Alzheimer's disease is also divided into five-year intervals, with the numbers 1-4 indicating the current stage of the disease. The number "0" is used to indicate that an individual does not have the disease. Each of these 140 states are transient because it is impossible to remain in the same age group when transitioning between states. There are also six absorbing states that are designed to track deaths in the population, both by gender (M or F) and by cause of death (Alzheimer's disease, other causes, or a combination of both). The model simulates changes in the population over five-year intervals (i.e. 2010, 2015, …, 2060). When the model transitions from one state to another, healthy individuals can remain healthy, develop Alzheimer's disease, or die of other causes. Individuals who initially have Alzheimer's can advance to a new stage of the disease, die from the disease, or die from a combination of the disease and other factors.

The initial state vector was based on Census data that breaks down the U.S. population by age and gender (Census 1). The current prevalence of Alzheimer's disease within each age group was estimated using statistics from the Alzheimer's Association. These statistics indicate that 5.4

million Americans currently have Alzheimer's disease, with 4% of cases occurring in people under age 65, 6% in ages 65-74, 44% in ages 75-84, and 46% in people above age 85 (F&F 14). Given the lack of available information on when these cases were diagnosed, the model assumes that most of the cases (80%) are currently in the first stage (1-5 years old), while only 1% of cases have progressed to the fourth stage (16-20 years). This assumption is consistent with the fact that only 3% of people with Alzheimer's disease live for more than 14 years after the diagnosis (Mölsä 1).

The transition probabilities were estimated based on past research and actuarial data. For individuals of a given age and gender, the baseline probability of death was taken from actuarial data compiled by the Social Security Administration (Actuarial 1). At a given age, the lower death probability for females reflects the fact that females generally live longer than males. To estimate the baseline probability of developing Alzheimer's disease, two prior studies on the topic were analyzed. Brookmeyer et al. estimated incidence rates as an exponential function of age (Brookmeyer 187). Through a linear regression, the researchers derived the following equation for the incidence rate of Alzheimer's disease (where t is an individual's age): $Incidence\ rate = 0.117e^{.127(t-60)}$. Similarly, Kawas et al. estimated the age-specific incidence rates of Alzheimer's by analyzing an elderly population in Baltimore. Since this paper did not provide a distribution for incidence rates, an exponential distribution was created that roughly reflects the incidence rates they found: $Incidence\ rate = 0.16e^{.13(t-60)}$. The Kawas distribution seems to provide more accurate results than the Brookmeyer distribution, so the Kawas distribution was used as an estimate of age-specific incidence probabilities.

For individuals who already have Alzheimer's, the probability of death was estimated based on publicly available data about the progression of the disease. Two separate studies by Mölsä, Martilla, and Rinne have revealed that life expectancy after a diagnosis of Alzheimer's

disease is about seven years, while only 3% of people with Alzheimer's disease live for more than 14 years after the diagnosis (Mölsä 103; Mölsä 159). Both the probability of developing Alzheimer's and the progression of the disease were assumed to be irrespective of age or gender. This assumption is consistent with the statement in the Brookmeyer paper that epidemiological data are not "sufficient to more precisely characterize rates of disease progression" (Brookmeyer 188).

These baseline probabilities for developing Alzheimer's disease, dying from Alzheimer's disease, and dying from other causes were then combined to calculate the probabilities for the transition matrix. The updated probabilities account for the mutual exclusivity of certain events within the five-year period. For example, the baseline probability of developing Alzheimer's was updated to reflect the possibility that an individual who would have developed the disease might have died of other causes beforehand. This process essentially turns a conditional probability (the probability of developing the disease conditional on not dying from other causes) into a lower non-conditional probability (the overall probability of developing the disease). As an example, the baseline probability of developing Alzheimer's disease for males in the 85-89 age group is 30.99%, but the probability used in the model is only 18.71%. Similarly, for a person who already has Alzheimer's disease, the baseline probabilities of death from Alzheimer's disease and death from other causes were combined to determine the non-conditional probability of each event occurring.

The basic model does not account for the potential impact of increased life expectancy on the prevalence and mortality rates of Alzheimer's disease. Life expectancy is an important factor to consider because the risk of developing Alzheimer's disease increases almost exponentially at older ages. To reflect this, two separate modifications were made to the original model. In the first modification, the baseline probabilities of death from causes other than Alzheimer's were

reduced by 5% for individuals of a given age and gender. In the second modification, these baseline probabilities were reduced by 10% to reflect even greater breakthroughs in medical technology. As long as the model is accurate, these increases in life expectancy should increase both the expected prevalence and mortality rates of Alzheimer's disease.

Results and Analysis

The projected prevalence of Alzheimer's disease for both males and females is shown in Exhibit 1. The model implies that in 2060, there will be approximately 4.25 million males with the disease and approximately 7.4 million females with disease. In total, there are expected to be 11.65 million cases in 2060, a 115.7% increase over the current prevalence of 5.4 million cases. The shape of the graph reflects the significant increase in Alzheimer's prevalence as the baby boomer generation ages, followed by a slight decline when subsequent generations start to age. The next two exhibits reflect the impact of increased life expectancy on the prevalence of Alzheimer's disease. When the probabilities of death for individuals of a given age and gender are decreased by 5%, the model predicts that 12.3 million Americans will have Alzheimer's in 2060 (4.5 million males, 7.8 million females), a 5.6% increase over the number of cases predicted in the base case. Similarly, when the probabilities of death are reduced by 10% (Exhibit 3), the model predicts 17.3 million cases of Alzheimer's in 2060 (6.5 million males, 10.8 million females), a 48.5% increase over the base case prediction. This sharp increase reflects the significant marginal effect that increased life expectancy would have on the prevalence of Alzheimer's disease.

Exhibit 4 shows the projected mortality rates of Alzheimer's disease in the base case. The model implies that in the 50-year period from 2010-2060, about 10 million males and 17.2 million females will die from Alzheimer's disease. If we include Alzheimer's patients who die

from other causes, the projected number of deaths is 17 million males and 22 million females. Exhibit 5 and Exhibit 6 illustrate how increased life expectancy would affect the mortality rates of Alzheimer's disease. When life expectancy is increased by a relatively small amount (Exhibit 5), the projected number of deaths from Alzheimer's disease only increases from 27.2 million people (10 million males, 17.2 million females) to 27.6 million people (10.1 million males, 17.5 million females). However, when life expectancy increases more significantly (Exhibit 6), the projected number of deaths increases from 27.2 million people to about 36.3 million people (13.5 million males, 22.8 million females). It is also interesting to note that as life expectancy increases, Alzheimer's patients become more likely to die from the disease itself than to die from other causes, since the model assumes no improvements in the treatment of Alzheimer's disease.

Future Research

This Monte Carlo simulation provides a framework for predicting the future prevalence and mortality rates of Alzheimer's disease in the United States. The model includes several simplifying assumptions about Alzheimer's disease that could be corrected for in future models. First, it assumes that the age-specific probabilities of developing Alzheimer's disease will remain constant over the next 50 years. In reality, these probabilities will likely decrease as researchers come to better understand the causes of Alzheimer's disease and the ways in which individuals can reduce their risk. In future research, the current framework could be used to conduct a sensitivity analysis evaluating how lower incidence rates would impact prevalence and mortality. Second, the simulation of increased life expectancy in this model assumes that all reductions in death probabilities occur immediately, when in reality these improvements will occur gradually over time. To more precisely simulate increased life expectancy, researchers could create a Monte Carlo simulation that would iterate through the transition matrix after each stage and reduce the probability of death for individuals in the older age groups. In this alternative

simulation, we would expect the prevalence of Alzheimer's to look more like an exponential distribution, since each successive increase in life expectancy would increase the number of people at high risk for developing Alzheimer's disease.

Conclusion

The main results of this Monte Carlo simulation are consistent with the hypotheses discussed in the introduction. The model projects a consistent increase in Alzheimer's prevalence over the next 35-40 years, which coincides with the aging of the baby boomer generation. In addition, the modifications to life expectancy result in an increased number of Alzheimer's cases and an increased mortality rate for the disease. Both of these findings are consistent with the fact that advancing age is a major risk factor for developing Alzheimer's disease. While the precision of the model is reduced by some of the simplifying assumptions, the underlying framework could be used as a building block for future simulations of Alzheimer's disease.

Simulation of Alzheimer's prevalence over 50-year period:

```
N=10;
IniPop=xlsread('Population 2.xlsx');
TransMatrix=xlsread('Transition matrix 2.xlsx');
PopMatrix=zeros(N,length(IniPop));
PopVec=IniPop*TransMatrix;
PopMatrix(1,:)=PopVec;
AD_male(1,:)=sum(PopVec(12:70))-sum(PopVec(16:5:66));
AD_female(1,:)=sum(PopVec(85:143))-sum(PopVec(89:5:139));

for i=2:N
PopMatrix(i,:)=PopVec*TransMatrix;
PopVec=PopMatrix(i,:);
AD_male(i,:)=sum(PopVec(12:70))-sum(PopVec(16:5:66));
AD_female(i,:)=sum(PopVec(85:143))-sum(PopVec(89:5:139));
end;

subplot(1,2,1);
a=[sum(IniPop(12:70))-sum(IniPop(16:5:66)) AD_male'];
plot(1:N,AD_male');
title('Number of AD (male)');
ylabel('');
xlabel('');

subplot(1,2,2);

plot(1:N,AD_female');
title('Number of AD (female)');
ylabel('');
xlabel('');
```

Simulation of Alzheimer's mortality over 50-year period:

```
N=10;
IniPop=xlsread('Population 4.xlsx');
TransMatrix=xlsread('Transition matrix 4.xlsx');
PopMatrix=zeros(N,length(IniPop));
PopVec=IniPop*TransMatrix;
PopMatrix(1,:)=PopVec;
DeathAD_male(1,:)=PopVec(71);
DeathAD_female(1,:)=PopVec(144);
DeathOther_male(1,:)=PopVec(72);
DeathOther_female(1,:)=PopVec(145);
DeathCombined_male(1,:)=PopVec(73);
DeathCombined_female(1,:)=PopVec(146);
TotalDeathAD_male(1,:)=PopVec(71)+PopVec(73)
TotalDeathAD_female(1,:)=PopVec(144)+PopVec(146)

for i=2:N
```

```
PopMatrix(i,:)=PopVec*TransMatrix;
PopVec=PopMatrix(i,:);
DeathAD_male(i,:)=PopVec(71);
DeathAD_female(i,:)=PopVec(144);
DeathOther_male(i,:)=PopVec(72);
DeathOther_female(i,:)=PopVec(145);
DeathCombined_male(i,:)=PopVec(73);
DeathCombined_female(i,:)=PopVec(146);
TotalDeathAD_male(i,:)=PopVec(71)+PopVec(73)
TotalDeathAD_female(i,:)=PopVec(144)+PopVec(146)
end;

subplot(2,3,1);

plot(1:N,DeathAD_male');
title('Death from AD (male)');
ylabel('');
xlabel('');
subplot(2,3,2);
plot(1:N,DeathAD_female');
title('Death from AD (female)');
ylabel('');
xlabel('');

subplot(2,3,3);

plot(1:N,DeathCombined_male');
title('Death from Combined (male)');
ylabel('');
xlabel('');
subplot(2,3,4);
plot(1:N,DeathCombined_female');
title('Death from Combined (female)');
ylabel('');
xlabel('');

subplot(2,3,5);

plot(1:N,TotalDeathAD_male');
title('Total Death AD(male)');
ylabel('');
xlabel('');
subplot(2,3,6);
plot(1:N,TotalDeathAD_female');
title('Total Death AD(female)');
ylabel('');
xlabel('');
```

Exhibit 1:

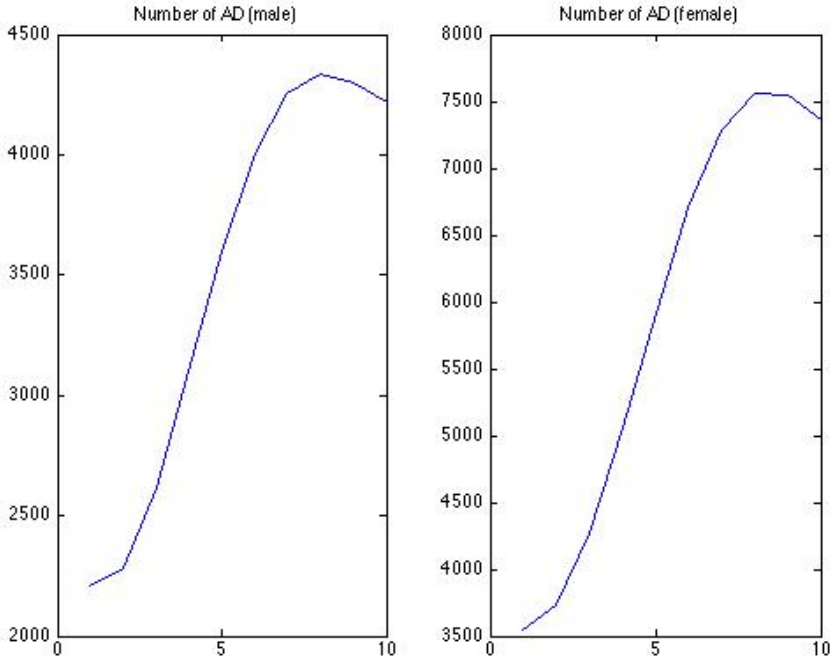


Exhibit 2:

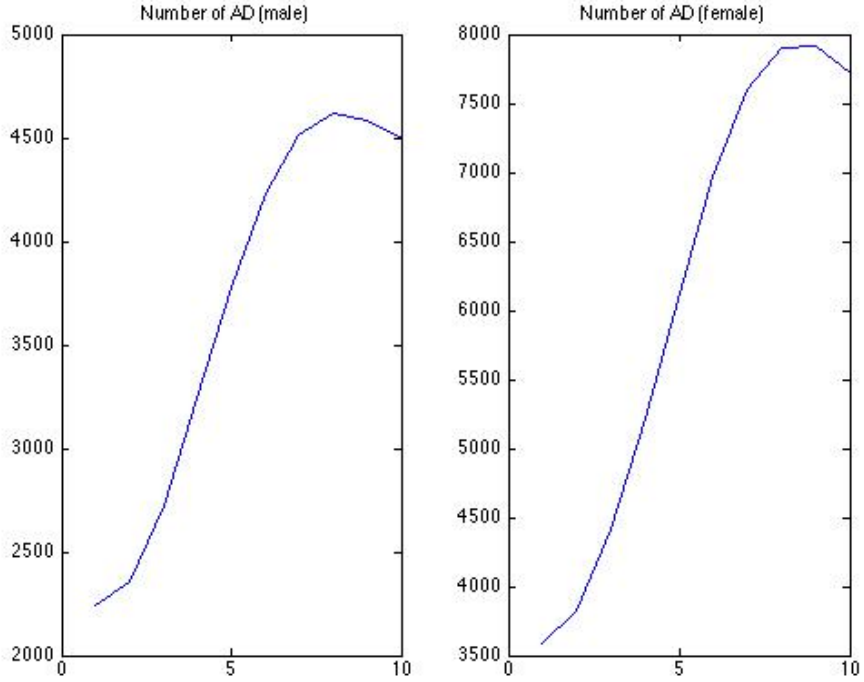


Exhibit 3:

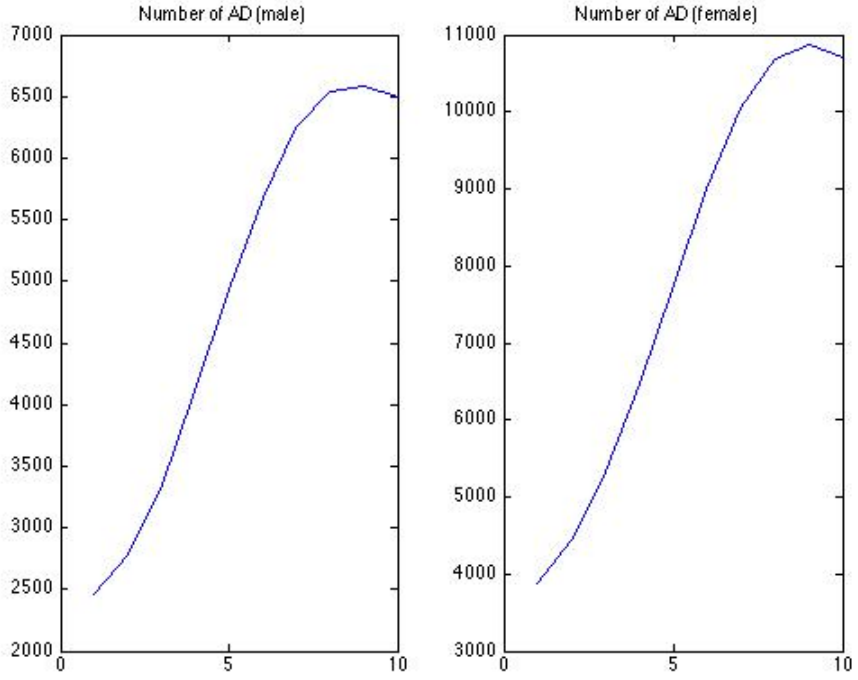


Exhibit 4:

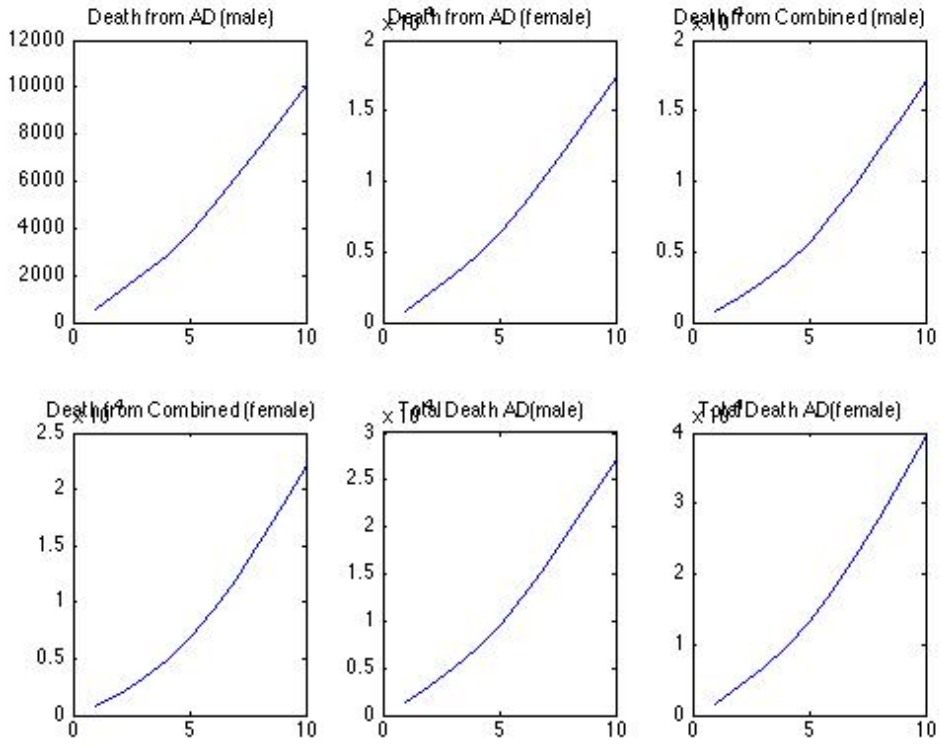


Exhibit 5:

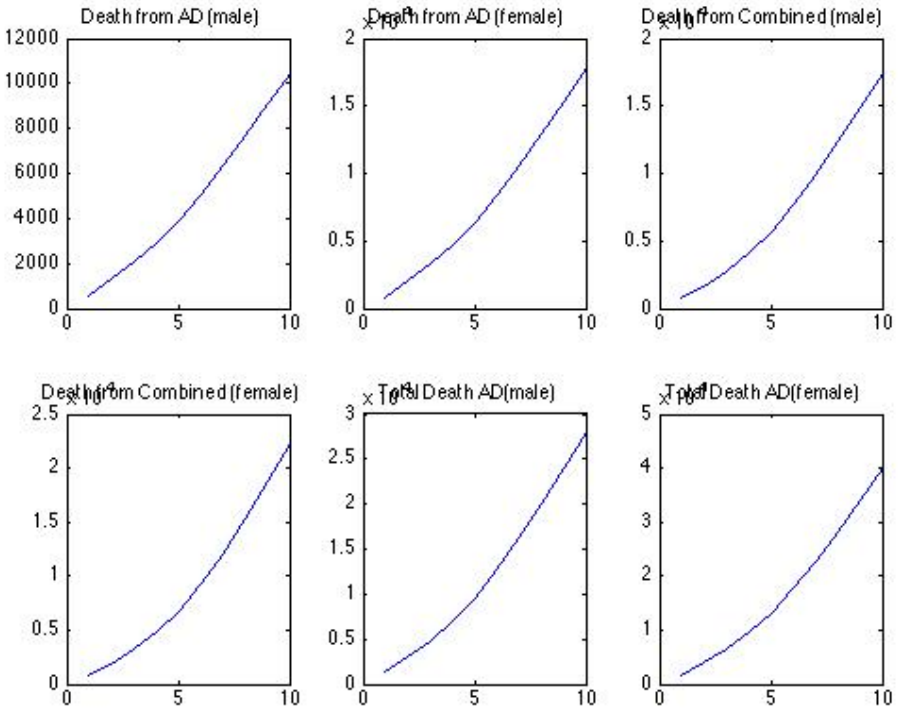
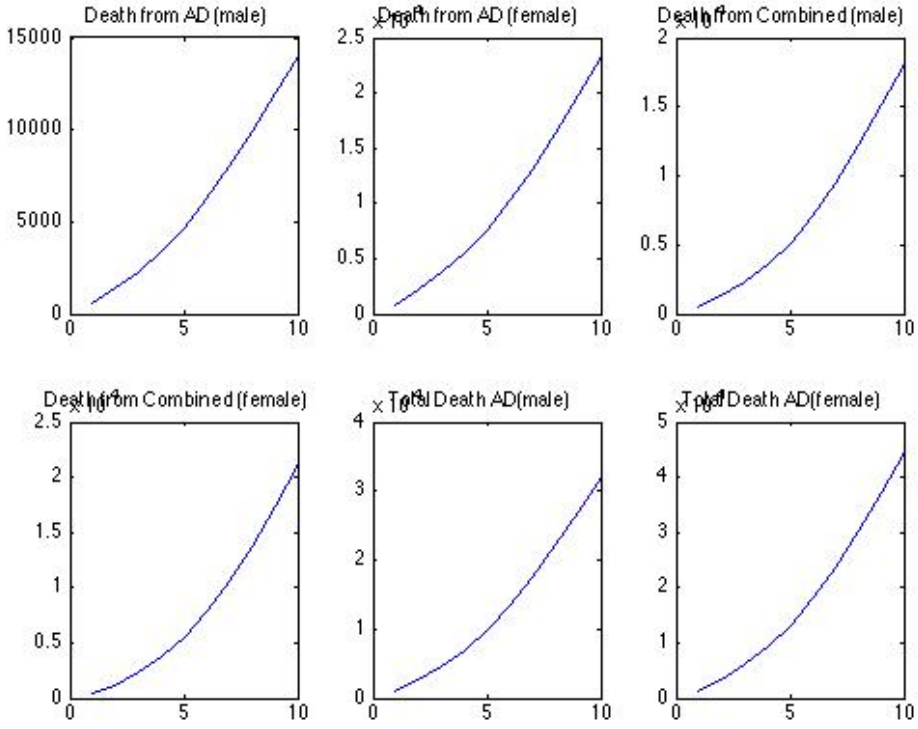


Exhibit 6:



References

"Actuarial Life Table." United States Social Security Administration. http://www.ssa.gov/oact/STATS/table4c6.html

"2012 Alzheimer's Disease Facts and Figures." *Alzheimer's & Dementia* 8.2 (2012): 131-168. Web. 9 Dec 2012.

Brookmeyer, Ron et al. "Forecasting the global burden of Alzheimer's disease." *Alzheimer's & Dementia* 3.3 (2007): 186-191. Web. 10 Dec 2012.

Hebert, L. E. "Alzheimer Disease in the US Population: Prevalence Estimates Using the 2000 Census." *Archives of Neurology* 60.8 (2003): 1119-122. Web. 10 Dec 2012.

Kawas, C. et al. "Age-specific Incidence Rates of Alzheimer's Disease." *Neurology* 54.11 (2000): 2072-2077. Web. 10 Dec 2012.

Mölsä, P.K., R.J. Marttila, and U.K. Rinne. "Survival and Cause of Death in Alzheimer's Disease and Multi-infarct Dementia." *Acta Neurologica Scandinavica* 74.2 (1986): 103-07. Web. 10 Dec. 2012.

Mölsä, P. K., R. J. Marttila, and U. K. Rinne. "Long-term Survival and Predictors of Mortality in Alzheimer's Disease and Multi-infarct Dementia." *Acta Neurologica Scandinavica* 91.3 (1995): 159-64. Web. 9 Dec 2012.

United States Census Bureau. "Age and Sex Composition in the United States: 2011."