Disease Prevention Microsimulation Model [DRAFT]

Technical Report, version 3.9.16

2016 **ihsmarkit.com**

Tim Dall
Managing Director
Email: tim.dall@ihs.com
Phone: +1 202 481 9291

Wayne Su
Director
Email: wayne.su@ihs.com
Phone: +1 202 481 7774

Frank Chen
Consultant
Email: frank.chen@ihs.com
Phone: +1 202 481 7795

Table of Contents

TABLE OF EXHIBHITS

Introduction

In the current environment of soaring healthcare costs, decision makers require tools that enable them to make informed choices about managing population health in the most cost effective way possible. The IHS Life Sciences Disease Prevention Microsimulation Model (DPMM) allows users to recreate complex interactions between observed individual characteristics and associated risk factors so that conclusions may be drawn about health and economic outcomes within a specific population. Using findings from published clinical trials and original analysis, the model simulates probabilities of disease onset among customizable populations of interest under user-specified intervention scenarios. In this way, the DPMM facilitates understanding of outcomes in multiple dimensions, at present and in future, and under alternative interventions scenarios. Model outcomes include disease incidence and prevalence, medical costs, other economic outcomes (employment, productivity, disability payments, government revenue and expenditures), quality of life, and mortality.

Conceptual Overview

Microsimulation modeling is based on the observation that each person's health outcomes are dependent on his or her unique characteristics—including demographics (e.g. age, sex, race/ethnicity), biometrics (e.g. body mass index [BMI], blood pressure, cholesterol levels, blood glucose level), and presence or history of disease or other conditions (e.g. diabetes, heart disease, history of stroke). Though no two individuals are exactly alike, data at a trial or population level can inform how one's risk profile might evolve as the person ages. Changes in this risk profile can lead to changes in probability and timing of adverse medical events, which in turn affects medical expenditures, labor force participation, quality of life, and mortality risk.

These causal relationships allow for the use of a Markov Chain model. A Markov model is based on the assumption that health outcomes in the next year (cycle) are dependent on the health profile in the current year.[1] In the same way, an individual's health profile in the next year determines his or her health outcomes two years in advance. This cycle continues throughout the duration of simulation.

Monte Carlo simulation is based on running repeated random sampling in order to obtain more accurate results. Given the heterogeneity of characteristics across individuals and the myriad of differences in possible changes in clinical characteristics and health outcomes, Monte Carlo simulation is an ideal way of dealing with parameter uncertainty. Probabilities of outcomes come from analyses of national surveys, published clinical trials, and peer-reviewed literature. Each individual is assigned a probability for each outcome in each year and this probability is compared to a random number generated from a uniform distribution between 0 and 1. If the probability of the outcome exceeds the random number then the outcome is simulated to occur.

Model Applications

Illustrating the potential clinical and economic benefits of lifestyle intervention in prediabetes

Variance in the policies that governed screening (at the time, guidelines did not uniformly recommend screening for prediabetes) created an opportunity to explore the implications of lifestyle intervention in a pre-diabetes population, in terms of long-term health and economic outcomes. The DPMM was used to model two interventions in three different prediabetes populations over a time horizon of 10 years. Study findings indicated that potential savings were possible for each of the three prediabetes populations, and that lifestyle interventions resulted in benefits that exceeded associated intervention implementation costs.[2]

Assessing the clinical and cost effectiveness of interventions to treat obesity

Obesity remains a major public health concern, and multiple initiatives are being implemented to curb the epidemic. However, the long term effects of these strategies will take several years to manifest, which allows simulation modeling the opportunity to estimate the potential impacts of such interventions. The DPMM is one of the few validated longitudinal simulation models that have examined the effects of obesity on health and economic outcomes. Study results estimated that each excess kilogram of weight contributed to an average of $140 greater annual health costs, and that obesity was associated with increased work absenteeism, mortality, lower probability of employment, income and quality of life.[3]

Projecting future burden of chronic conditions and associated economic implications

The DPMM can be used to estimate the current and future burden of major chronic illnesses at the US national and state level, under different health intervention scenarios.

Model Structure

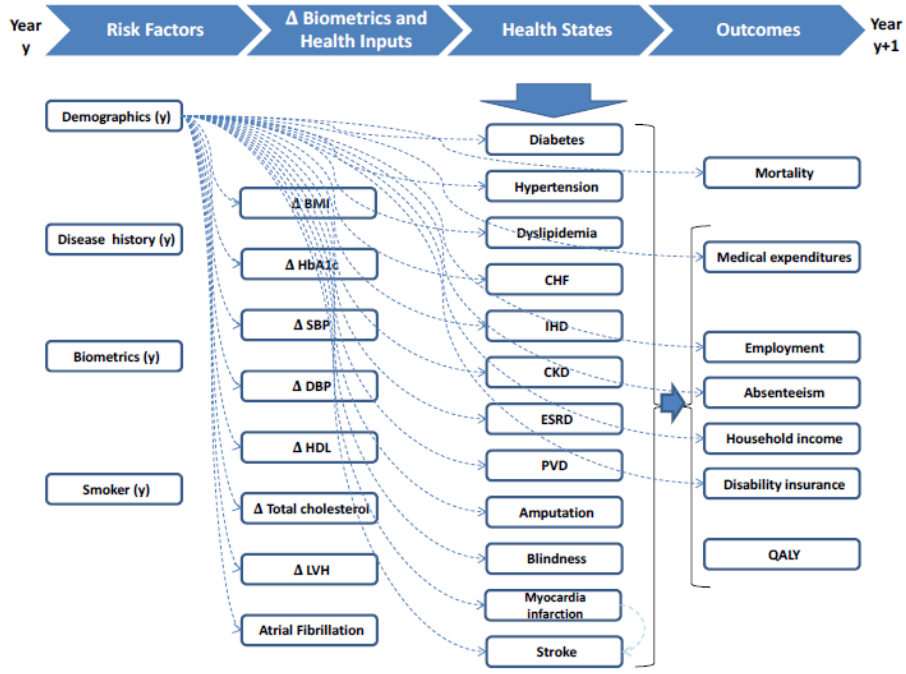
When modeling risk for chronic disease, only a subset of a person's characteristics and health risk factors are directly observable to researchers. These factors typically include: demographics, biometrics, health behavior, presence of chronic diseases, and socioeconomic characteristics. Because the number of unique combinations of disease states is large and the outcomes can vary by a person's risk factors and health-related behavior, a microsimulation approach, where for each person the model can track the presence of each disease and disease risk factor, is useful. The Markov structure used a person's characteristics at time t to predict characteristics at time t+1, with this process repeated annually through the projection horizon unless mortality occurred sooner. The model used Monte Carlo simulation based on running repeated random sampling. Given the heterogeneity of characteristics across individuals and the myriad of differences in possible changes in clinical characteristics and health outcomes, Monte Carlo simulation allowed for analysis of individual prediction uncertainty.

The transition states in the model and an overview of the relationships between risk factors and the sequence of transitions are summarized in the following exhibits. Each linkage in these exhibits represents a prediction equation or model parameter quantifying the relationship between a risk factor and clinical outcome.

Model overview

A person starts the year with certain risk factors including demographics; biometrics such as body mass index (BMI), systolic and diastolic blood pressure levels, hemoglobin A1c (HbA1c) level, cholesterol levels; current smoking status; and the presence of various chronic diseases or past history of adverse medical events (e.g., stroke, cancer or myocardial infarction). Demographics (age and sex, and sometimes race and Hispanic ethnicity) are inputs to almost every prediction equation in the model. (Exhibit 1) The change in a person's biometrics as he or she ages will vary by current age and sex, with race and ethnicity also playing a role for some biometrics. Demographics are independent predictors in the equations to model disease incidence probability, mortality, annual medial expenditures, and other economic outcomes modeled. For some prediction equations, there are separate equations for men and women. For others, sex enters the equation as a binary indicator. The DPMM contains prediction equations of disease risk during the year based on patient characteristics and health risk factors. The ending values at year y become the starting values for year y+1 and the process repeats through the projection horizon.

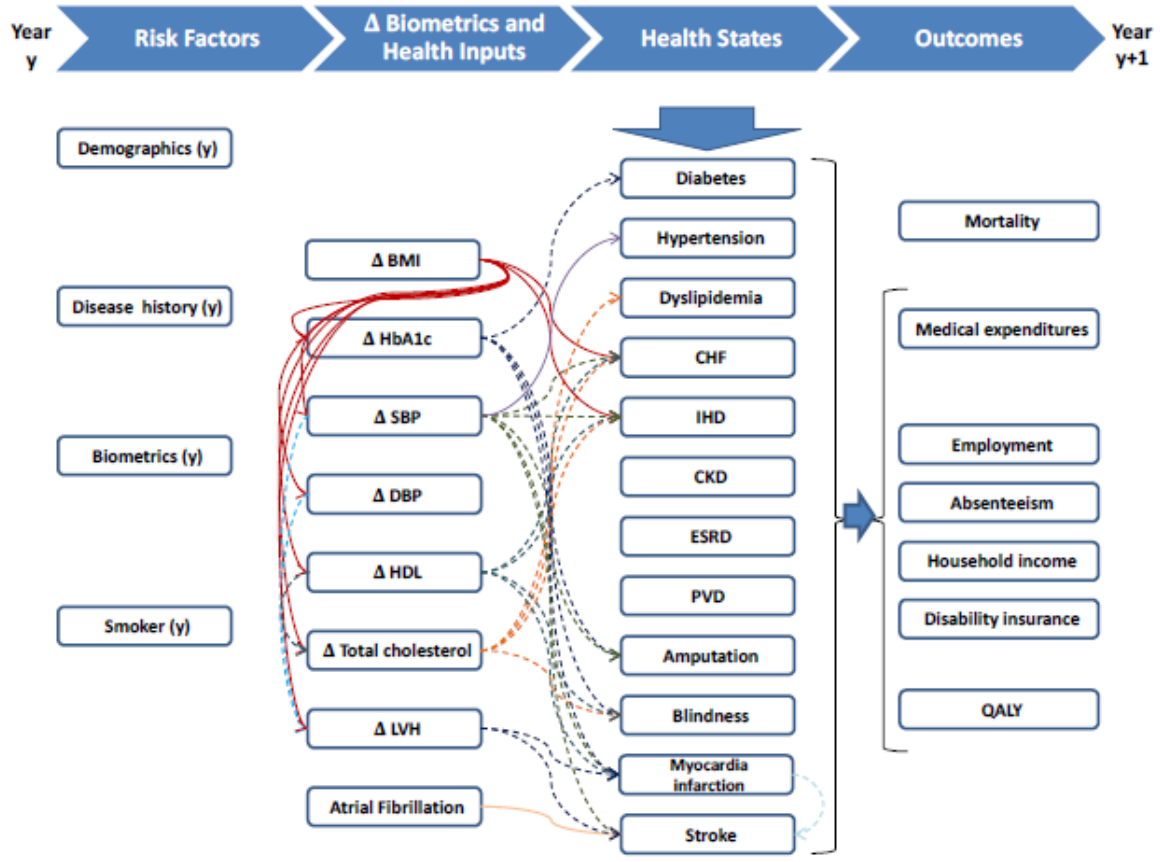
Exhibit 1. Impact of Demographics



Note: Connecting lines show the items in the model that are linked. Abbreviations: BMI=body mass index, CHF=congestive heart failure, CKD=chronic kidney disease, DBP=diastolic blood pressure, ESRD=end stage renal disease, HbA1c=hemoglobin A1c, HDL=high-density lipoprotein, IHD=ischemic heart disease, LVH=left ventricular hypertrophy, PVD=peripheral vascular disease, QALY=quality adjusted life year, SBP=systolic blood pressure.

Change in some biometrics is linked to change in other biometrics, and change in biometrics is combined with current biometric levels as inputs to the prediction equations for disease onset (Exhibit 2). BMI is a key biometric, and change in BMI is linked to change in many other biometrics as well as an independent predictor for congestive heart failure (CHF) and ischemic heart disease (IHD). HbA1c is directly linked to diabetes onset, but HbA1c also is linked to risk for myocardial infarction, amputation, and retinopathy. Atrial fibrillation (AF) is not modeled as an end state, but is included in the model solely as a risk factor for stroke. It is not a condition modeled for cost or QALY purposes, and we do not assume that AF increases the risk for diabetes or that diabetes increases the risk of AF.

Exhibit 2. Impact of Biometrics

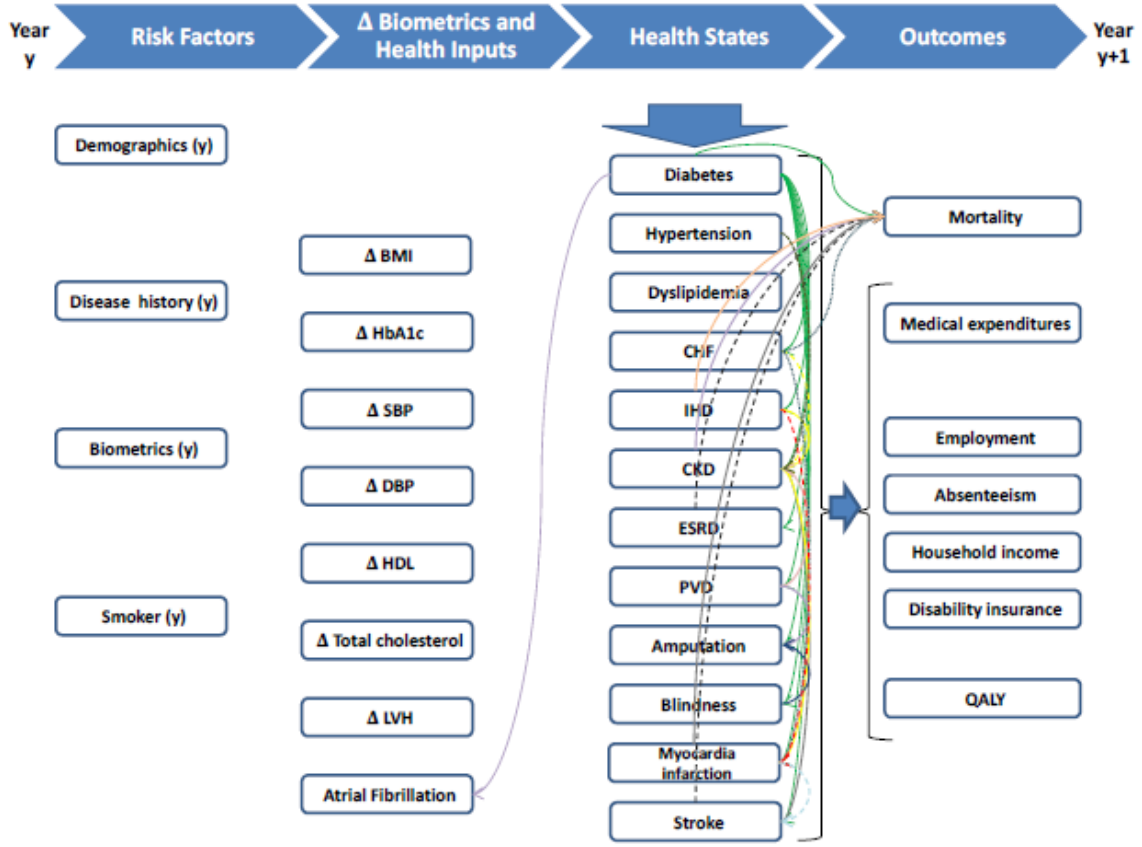


Note: Connecting lines show the items in the model that are linked. Abbreviations: BMI=body mass index, CHF=congestive heart failure, CKD=chronic kidney disease, DBP=diastolic blood pressure, ESRD=end stage renal disease, HbA1c=hemoglobin A1c, HDL=high-density lipoprotein, IHD=ischemic heart disease, LVH=left ventricular hypertrophy, PVD=peripheral vascular disease, QALY=quality adjusted life year, SBP=systolic blood pressure.

The diabetes module of DPMM is used to model the progression of diabetes and its associated impacts, and Exhibit 3 illustrates how the risk for all other health states rises when an individual develops the condition. A

positive diagnosis of diabetes affects the prediction equations for the sequelae in one of three ways. One, for some sequelae the prediction equations for people with diabetes come from sources like the UKPDS study that is specific to a population with diabetes, while for the prediabetes population the prediction equations come from other sources for a non-diabetic population. Two, diabetes sometimes enters the prediction equation as a dichotomous variable indicating presence of disease. Three, for some sequelae the time since diabetes onset is an input to the prediction equation. Many of the disease states are linked to mortality risk. Almost all the disease states are inputs to medical expenditures, economic outcomes, and quality of life.

Exhibit 3. Impact of Morbidity when modeling diabetes



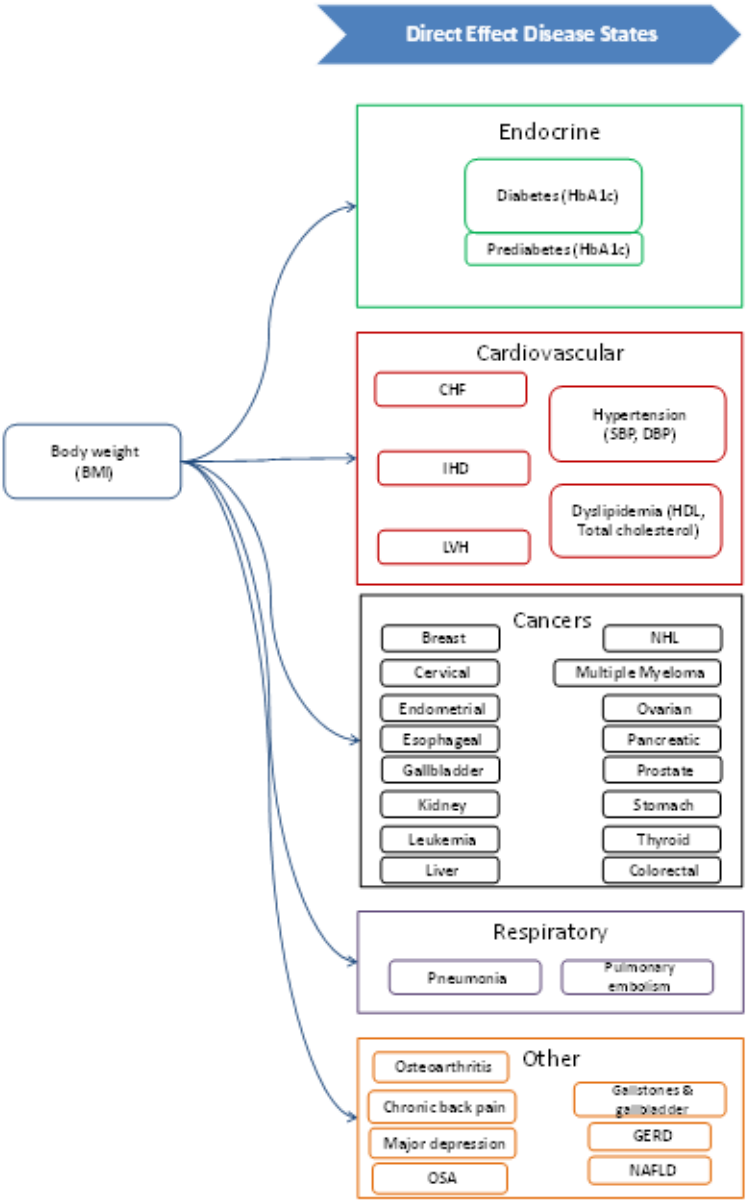
Note: Connecting lines show the items in the model that are linked. Abbreviations: BMI=body mass index, CHF=congestive heart failure, CKD=chronic kidney disease, DBP=diastolic blood pressure, ESRD=end stage renal disease, HbA1c=hemoglobin A1c, HDL=high-density lipoprotein, IHD=ischemic heart disease, LVH=left ventricular hypertrophy, PVD=peripheral vascular disease, QALY=quality adjusted life year, SBP=systolic blood pressure.

For obesity module of DPMM, Exhibit 4 illustrates how BMI is directly linked to an individual's risk for the following:

- Endocrine problems—specifically, prediabetes and diabetes
- Cardiovascular problems—hypertension, dyslipidemia, congestive heart failure (CHF), ischemic heart disease (IHD), and peripheral vascular disease (PVD), and ventricular hypertrophy (LVH)[1]
- Cancer—breast, colorectal, endometrial, kidney, leukemia, liver, non-Hodgkin's lymphoma, pancreas, stomach, thyroid, and other cancers with lower incidence
- Respiratory conditions—asthma, pneumonia, and pulmonary embolism

[1] LVH is modeled solely as a risk for myocardial infarction and stroke.

Exhibit 4. Impact of Body Weight

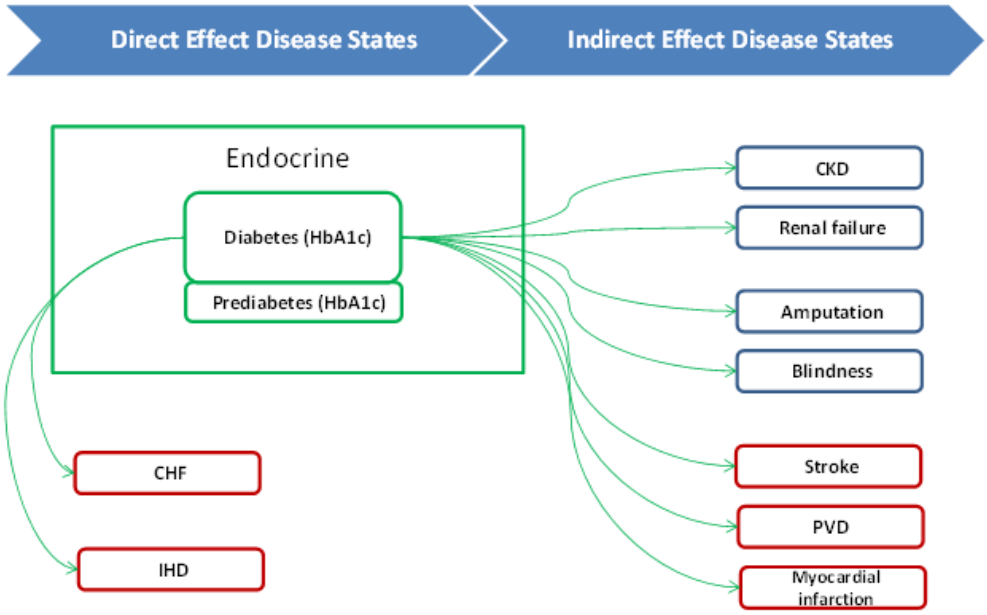


Note: Connecting lines show how BMI links to other modeled conditions. Abbreviations: BMI=body mass index, CHF=congestive heart failure, DBP=diastolic blood pressure, GERD= gastroesophageal reflux disease, HbA1c=hemoglobin A1c, HDL=high-density lipoprotein, IHD=ischemic heart disease, LVH=left ventricular hypertrophy, NAFLD=non-alcoholic fatty liver disease, OSA=obstructive sleep apnea, SBP=systolic blood pressure.

The link between body weight (measured by BMI) and HbA1c (used to define prediabetes and diabetes status), and between BMI and cardiovascular disease is based on prediction equations from the published literature

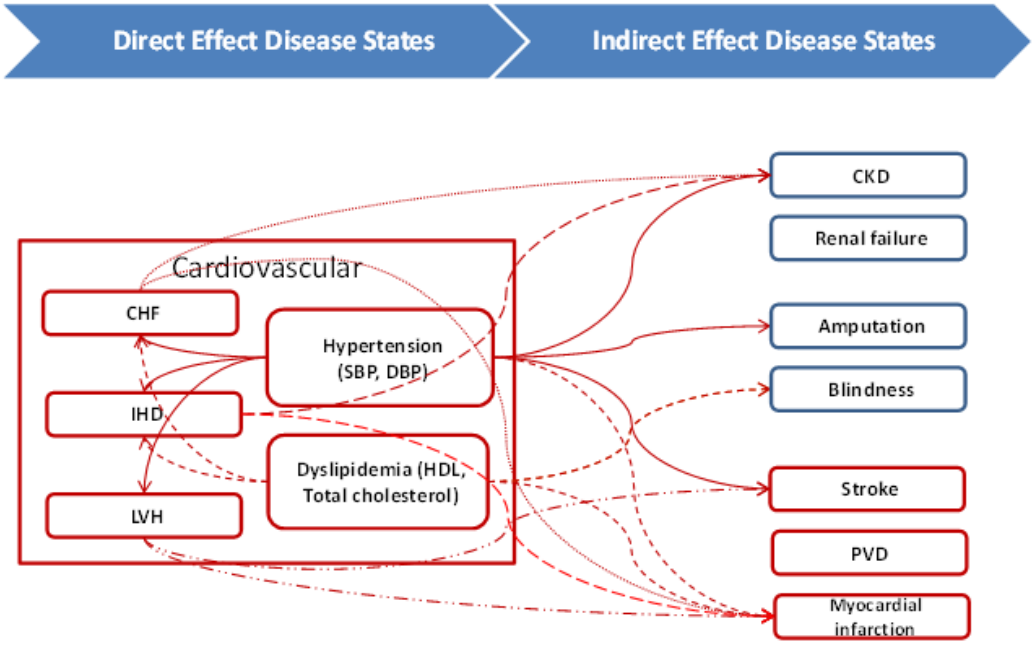
(described later). The links between BMI and cancers and the link with respiratory and other medical conditions also came from published literature. Body weight is indirectly linked to other conditions—such as diabetes sequelae— through the effect of excess body weight on insulin resistance, and ultimately elevated blood glucose levels as shown in Exhibit 5. Similarly, BMI is indirectly linked to various cardiovascular and other conditions through its direct links to select cardiovascular diseases. For example, risk for stroke is associated with blood pressure and cholesterol levels, which in turn are correlated with body weight. For myocardial infarction, body weight is linked indirectly through multiple routes—diabetes, hypertension, dyslipidemia, CHF, IHD, and LVH.[4;5]

Exhibit 5. Endocrine effects



Note: Connecting lines show the items in the model that are linked. Abbreviations: BMI=body mass index, CHF=congestive heart failure, CKD=chronic kidney disease, HbA1c=hemoglobin A1.

Exhibit 6. Cardiovascular effects



Note: Connecting lines show the items in the model that are linked. Abbreviations: BMI=body mass index, CHF=congestive heart failure, CKD=chronic kidney disease, DBP=diastolic blood pressure, HbA1c=hemoglobin A1c, HDL=high-density lipoprotein, IHD=ischemic heart disease, LVH=left ventricular hypertrophy, PVD=peripheral vascular disease, QALY=quality adjusted life year, SBP=systolic blood pressure.

The published equation for stroke risk used in the model includes patient occurrence of atrial fibrillation (AF), PVD, and LVH. These three conditions are solely included in the model to predict stroke occurrence. However, there is insufficient information in the Medical Expenditure Panel Survey (MEPS) to include these conditions in the medical expenditure prediction equations (discussed later) and insufficient information in the literature to model the mortality and quality of life implications of presence of these conditions independent of their effect on stroke.

As shown in Exhibit 7 the obesity module of the DPMM complex relationships between health risk factors, biometrics, health states, and outcomes of interest. This allowed us to model multiple orders of effects from change (Δ) in BMI, reflecting the complexity and interrelated nature of human health, such as the following:

1st order effect: ΔBMI→ΔSBP

2nd order effect: ΔBMI→ΔSBP→ΔCHF risk

3rd order effect: ΔBMI→ΔSBP→ΔCHF risk→Δmyocardial infarction risk

Linkages with many risk factors are not shown in the diagrams above but are included in the OPEM. For example, demographics are in the prediction equations for almost all the health states modeled. Almost every

prediction equation uses age as an input. Many of the equations use sex—either through different prediction equations for men and women or including sex as a dichotomous variable in the equation. Some equations include race and ethnicity as a risk factor. Smoking status is a risk factor in many of the equations (though in the simulations we assumed that smoking status stays constant over time). Risk factors are included in prediction equations based on published literature and data availability. Depending on the nature of disease, not all of the risk factors were significant. For instance, the risk of chronic kidney disease is the same across both genders, and is determined by age, sex, and whether or not the person has hypertension, diabetes, history of MI, stroke, IHD, CHF, and PVD.

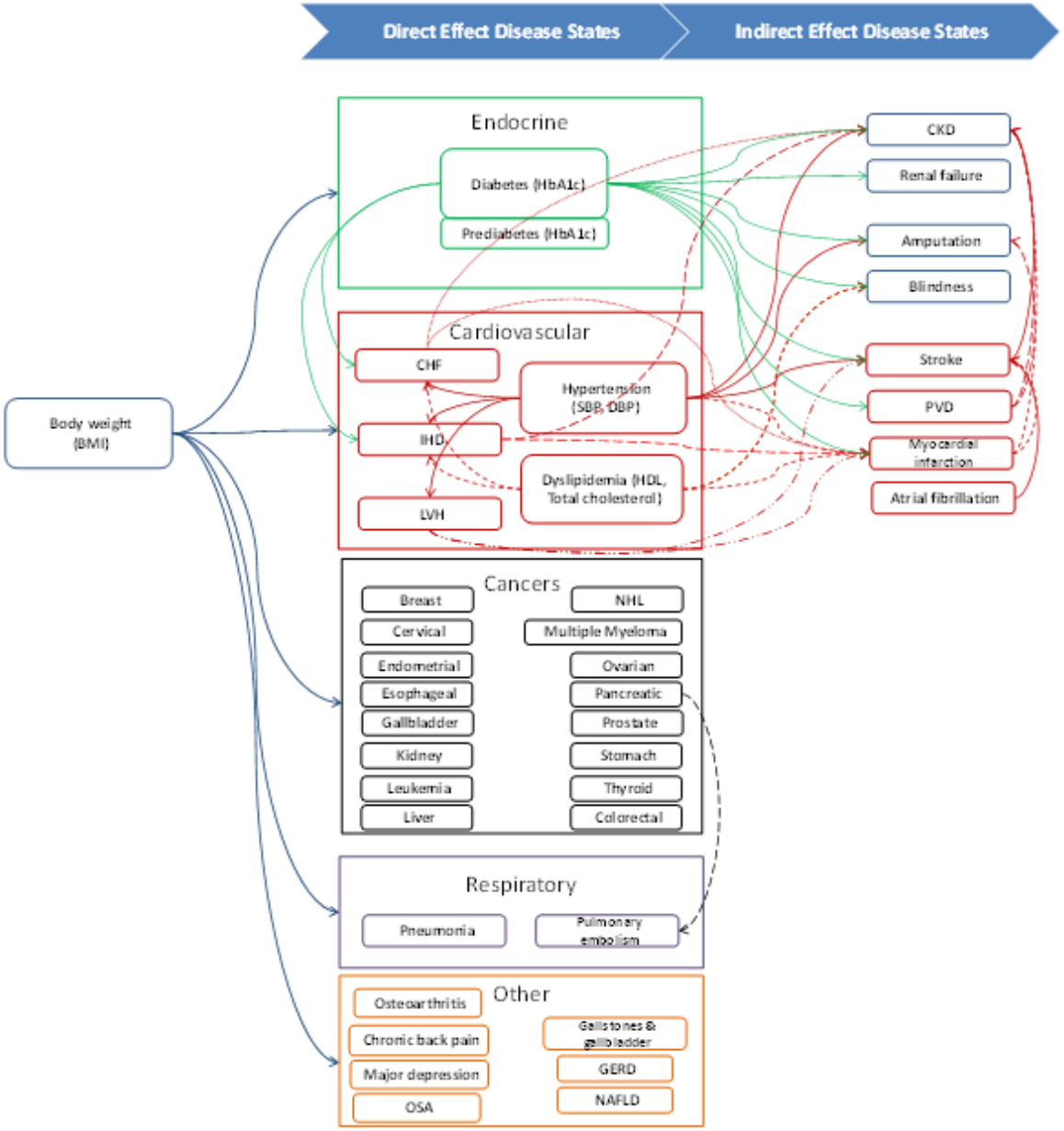
Exhibit 7 depicts the myriad of linkages in the obesity module between obesity and health conditions such as endocrine diseases (diabetes and prediabetes), cardiovascular diseases, cancers, and respiratory diseases, among others.

Due to data limitation, not all clinically established linkages were included in the model. For example, renal failure incidence currently is linked only to age and sex, but with a rate ratio adjustment of 9.0 (from the literature) to reflect that people with diabetes have 9 times the risk for renal failure incidence of a person without diabetes. Hypertension is also a contributor to renal failure risk, but that linkage was not included in the model due to the paucity of published information on that linkage. While the NHANES has a wealth of information on its population, there are gaps in the data that might preclude the use of some prediction equations from the literature. Information like family history of disease or genetic markers would be ideal for the modeling of some cancers and other conditions. Due to a lack of information on inputs used in published risk equations, alternate means of modeling were necessary in some cases. Whether or not the individual with hypertension is taking anti-hypertensive medications is imputed using patient characteristics, but this variable also can be modified as a user-defined scenario.

Patient mortality increases with the existence of major health conditions such cancer, CHF, CKD, diabetes, renal failure, IHD, myocardial infarction, pneumonia, and stroke.

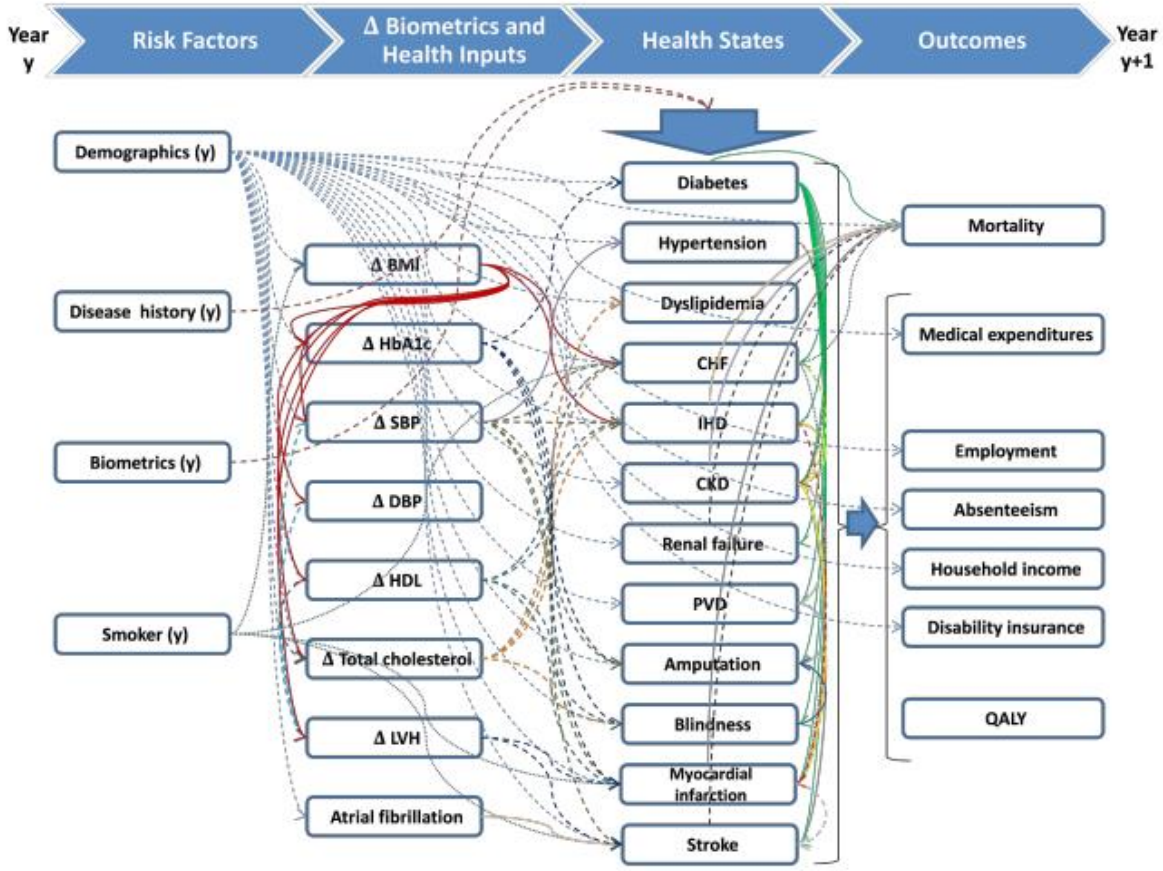
Many of these disease states affect a person's ability and decision to remain in the workforce (employment), their type of work and hours worked (affecting income), and time off from work when sick (absenteeism). Some conditions might qualify a person for disability (Supplemental Security Income [SSI]). The presence of adverse health conditions and chronic diseases also has implications for quality of life. The prediction equations for economic activity, disability, mortality, and quality of life are discussed later.

Exhibit 7. Model overview of DPMM obesity module



Note: Connecting lines show the items in the model that are linked. Abbreviations: BMI=body mass index, CHF=congestive heart failure, CKD=chronic kidney disease, DBP=diastolic blood pressure, GERD= gastroesophageal reflux disease, HbA1c=hemoglobin A1c, HDL=high-density lipoprotein, IHD=ischemic heart disease, LVH=left ventricular hypertrophy, NAFLD=non-alcoholic fatty liver disease, OSA=obstructive sleep apnea, PVD=peripheral vascular disease, SBP=systolic blood pressure.

Exhibit 8. Model Overview Diagram



Note: Connecting lines show the items in the model that are linked. Abbreviations: BMI=body mass index, CHF=congestive heart failure, CKD=chronic kidney disease, DBP=diastolic blood pressure, HbA1c=hemoglobin A1c, HDL=high-density lipoprotein, IHD=ischemic heart disease, LVH=left ventricular hypertrophy, PVD=peripheral vascular disease, QALY=quality adjusted life year, SBP=systolic blood pressure.

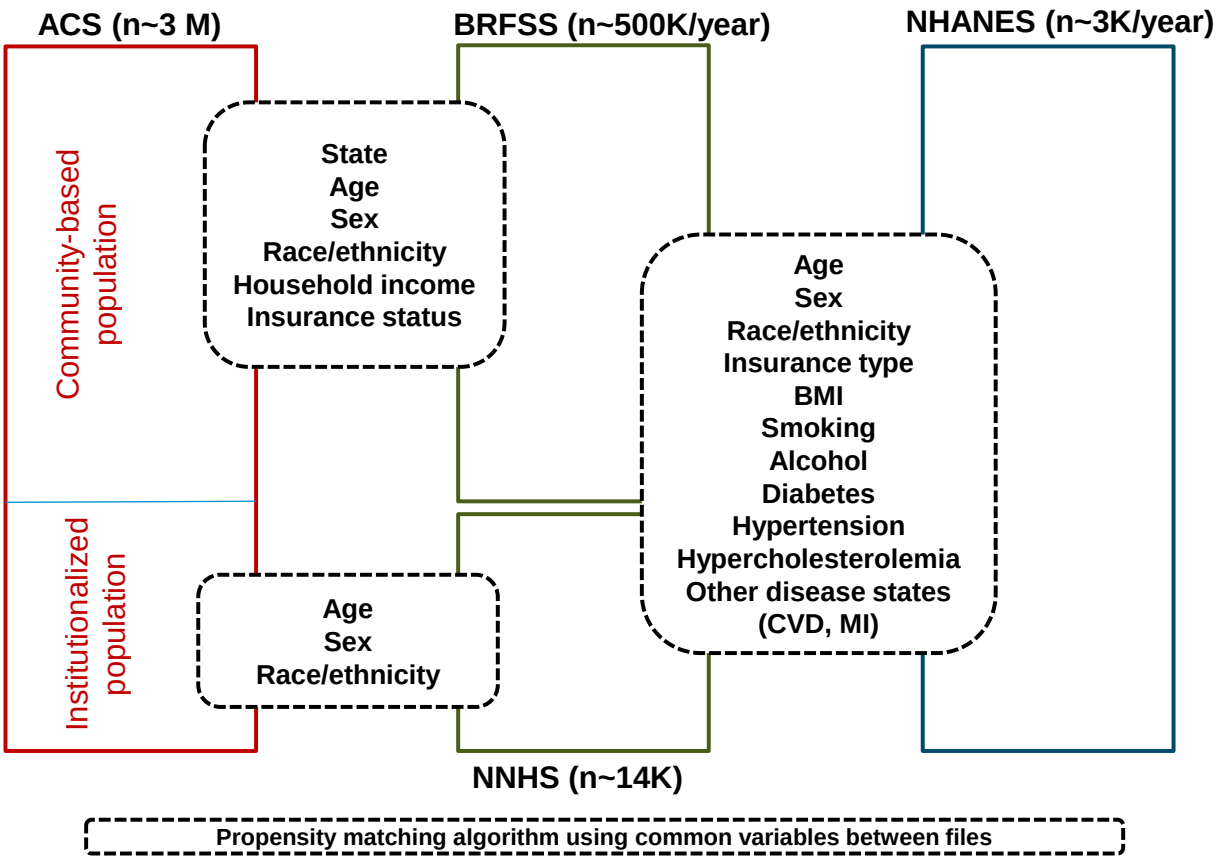
Simulation population

The adult population data sets were generated from multiple public data sources. To achieve the most accurate and complete clinical information for each individual, state level records from the American Community Survey (ACS, 2014) and Behavioral Risk Factor Surveillance System (BRFSS, 2013-2014) were merged to National Health and Nutrition Examination Survey (NHANES, 2005-2014) data through propensity match algorithm based on their age, gender, race, BMI, and insurance, diabetes, smoking, hypertension, and hyperlipidemia status. The combined data files provide metrics on SBP, total cholesterol, HDL-C, and HbA1c as well as other chronic illness

conditions for each US state. In addition, to better estimate the future clinical and economic burden, we produce the state level population projections from 2015 to 2030 based on published and IHS internal state and national projections in which the projected sample weights were assigned yearly to each of the demographic subsets. Each demographic subset is defined as a unique combination of 10-year age group, gender, and race.

Repeated sampling from the above mentioned state population file, using ACS sample weights to determine selection probability, produced representative samples of 100,000 adults for each state. In each modeled year, the sample sizes from the microsimulation model were compared with population projections for every demographic subset. If the actual number of individuals is less than projected population size, then persons with matching demographics are randomly selected to replenish the batch. If the actual model sample size is higher than projected, then the subset size is adjusted by randomly removing a small number of individuals, equal to the difference of the model sample size and the projected sample size. Additionally, as the result of population aging in the model, individuals who are 20 at initial year need to be supplemented each year since no one younger than 20 are included in the modeled adult population. We fulfilled this step by bootstrapping this specific age group of samples each time to maximize the heterogeneity in characteristics. The simulation populations are derived from algorithms that combine the various data sets in Exhibit 9 and can be specified according to the scenarios being modeled and the population being targeted.

Exhibit 9. Algorithm to generate the starting population



Overview of Model Parameters and Data Sources

Relationships between modeled risk factors and state transition probabilities came from published clinical trials, meta-analyses, observational studies, government statistics, and analyses of NHANES data. Where possible, we used published findings from recent meta-analyses. Priority was given to studies based on randomized clinical trials versus retrospective studies, to studies based on longitudinal versus cross-sectional data, to studies based on a diabetic population versus a general population (for those model components that apply to the population with simulated onset of diabetes), and to U.S.-based studies versus studies based on the population in other countries (primarily Europe). The literature search included broad topics such as disease prevalence, as well as narrow topics such as the association between specific risk factors and disease incidence. Key sources for published data include the United Kingdom Prospective Diabetes Study (UKPDS), the Framingham Heart Study, and the Centers for Disease Control and Prevention.

Exhibit 10 summarizes the health outcomes modeled, risk factors used to determine probability of adverse medical events for the populations with and without, and data sources. Individual model components are described later in more detail. Some biometric outcomes are calculated only for the diabetic or non-diabetic populations. This is due to the fact that in some cases, disease prediction equations published for the diabetic population use slightly different risk inputs than do equations published for the non-diabetic population.

Exhibit 10. Data Sources by Model Component

Outcome	Diabetic/Non-Diabetic Population	Risk Factors Considered	Source
Annual biometric changes (in absence of intervention)			
Body mass index	Both	Age, sex, weight category (normal, overweight, obese)	Analysis of 2003-2010 NHANES Sheehan et al., 2003[6]
Cholesterol Ratio	Non-diabetic	Total cholesterol, HDL cholesterol	Calculated by dividing total cholesterol by HDL-C
	Diabetic	Cholesterol ratio at diabetes diagnosis, previous year's cholesterol ratio	UKPDS Outcomes Model, Clarke et al., 2004[7]
Diastolic Blood Pressure	Non-diabetic	Change in BMI Aging	Neter et al., 2003[8] Analysis of 2003-2010 NHANES
Fasting Plasma Glucose	Both	HbA1c	Danaei et al., 2011[9;10]
HbA1C Level	Non-diabetic	BMI, age, cholesterol	Gadde et al., 2011[11] Heianza et al., 2012a[12] Heianza et al., 2012b[13]
	Diabetic	Years since diabetes diagnosis, HbA1C at diabetes diagnosis, previous year's HbA1C	UKPDS Outcomes Model, Clarke et al., 2004[7]
HDL cholesterol	Non-diabetic	Change in BMI	Framingham Heart Study, Wilson et al.,

Outcome	Diabetic/Non-Diabetic Population	Risk Factors Considered	Source
		Aging	1994[14] Analysis of 2003-2010 NHANES
Systolic Blood Pressure	Non-diabetic	Change in BMI Aging	Neter et al., 2003[8] Analysis of 2003-2010 NHANES
	Diabetic	Years since diabetes diagnosis, SBP at diabetes diagnosis, previous year's SBP	UKPDS Outcomes Model, Clarke et al., 2004[7]
Total cholesterol	Non-diabetic	Change in BMI Aging	Framingham Heart Study, Wilson et al., 1994[14] Analysis of 2003-2010 NHANES
Disease onset			
Atrial Fibrillation	Non-diabetic	Incidence rates by age and sex	Kannel et al., 1998[15]
Chronic back pain	Both	Incidence rates by age, relative risks by BMI	Guh et al., 2009[16]
Chronic Kidney Disease	Both	Age, sex, hypertension, diabetes, MI, stroke, IHD, CHF, PVD	The Atherosclerosis Risk in Communities Study and the Cardiovascular Health Study, Kshirsagar et al.,2009[17] Hsu et al.,2000[18]
Congestive Heart Failure	Both	Age, sex, on antihypertensive medication, SBP, total cholesterol, HDL cholesterol, smoking status	Framingham Heart Study, D'Agostino et al., 2008 [19]
Diabetes PreDiabetes	Diabetic	HbA1c, FPG	World Health Organization[20]
	Non-Diabetic	HbA1c, FPG	
End State Renal Disease	Both	Age and sex specific incidence rates Diabetes relative risk adjustment	Hippisley-Cox and Coupland, 2010 [21] Brancati et al., 1997 [22]
Gallstones & gallbladder	Both	Incidence rates by age and sex, Relative risks by BMI	Field et al., 2001[23] Maram et al., 1990[24]
Gastroesophageal reflux disease	Both	Incidence rates by age and sex, Relative risks by BMI	Nilsson et al., 2003[25] Ruigómez et al., 2004[26]
Left Ventricular Hypertrophy	Non-diabetic	SBP, DBP, age, BMI, race	de Simone et al., 1994 [27]
Peripheral vascular disease	Non-diabetic	Incidence rates by age and sex	Hooi et al., 2001 [28]
History of Atrial Fibrillation	Non-diabetic	Age and sex specific prevalence rates for population without diabetes	Nichols et al., 2009 [29]
	Diabetic	Age and sex specific prevalence rates for population with diabetes	
History of Peripheral Vascular Disease	Non-diabetic	Age and sex specific prevalence rates	Selvin and Erlinger, 2004 [30]
	Diabetic	Age and sex specific prevalence rates scaled by odds ratios for diabetics	Zhang et al., 2009 [31]
Ischemic Heart	Non-diabetic	Age, sex, BMI, smoking status, SBP,	Framingham Heart Study, Wilson et al.,

Outcome	Diabetic/Non-Diabetic Population	Risk Factors Considered	Source
Disease		cholesterol ratio	2008 [32]
	Diabetic	BMI, age, smoking status, SBP, cholesterol ratio, diabetes status, sex	
Non-Alcoholic Fatty Liver Disease	Both	Overall incidence rate, relative risks by BMI	Clark et al., 2002 [33] Ratziu et al., 2010 [34]
Osteoarthritis	Both	Overall incidence rate, relative risks by BMI	Guh et al., 2009 [35] Oliveria et al., 1995 [36]
Pneumonia	Both	Incidence rates by age, relative risks by BMI	Kornum et al., 2010 [37] American Lung Association [2] CDC [3]
Pulmonary embolism	Both	Incidence rates by age and sex, relative risks by BMI	Stein et al., 2005 [38] Silverstein et al., 1998 [39]
Breast Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, alcohol	CDC Seer Database [4] Green et al., 2012 [40]
Cervical Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, smoking	CDC Seer Database [d] Reeves et al., 2007 [41]
Colorectal Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, alcohol	CDC Seer Database [d] Moghaddam et al., 2007 [42]
Endometrial Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI	CDC Seer Database [d] Nagle et al., 2013 [43]
Esophageal Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, smoking, alcohol	CDC Seer Database [d] Chow et al., 1997 [44]
Gallbladder Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI	CDC Seer Database [d] Larsson et al., 2007 [45]
Kidney Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, smoking	CDC Seer Database [d] Adams et al., 2008
Leukemia	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, smoking	CDC Seer Database [d] Larsson et al., 2008 [46]
Liver Cancer	Both	Age, sex, race/ethnicity dependent	CDC Seer Database [d]

[2] Pneumonia Fact Sheet: http://www.lung.org/lung-disease/influenza/in-depth-resources/pneumonia-fact-sheet.html
[3] National Vital Statistics Reports, Deaths: Final Data for 2010 http://www.cdc.gov/nchs/data/nvsr/nvsr61/nvsr61_04.pdf
[4] Surveillance, Epidemiology, and End Results Program http://seer.cancer.gov/statistics/summaries.html

Outcome	Diabetic/Non-Diabetic Population	Risk Factors Considered	Source
		incidence rates; relative risk associated with BMI, smoking, alcohol	Larsson et al., 2007 [47]
Lung cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, smoking	CDC Seer Database [d] Peto et al, 2000
Multiple myeloma	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI	CDC Seer Database [d] Birmann et al., 2013 [48]
Non-Hodgkin's lymphoma	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI	CDC Seer Database [d] Lim et al., 2007 [49]
Ovarian Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI	CDC Seer Database [d] Reeves et al., 2007 [41]
Pancreatic Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, smoking	CDC Seer Database [d] Genkinger et al., 2011 [50]
Prostate Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI (reverse relationship)	CDC Seer Database [d] Wright et al., 2007 [51]
Stomach Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI, smoking	CDC Seer Database [d] Chen et al., 2013
Thyroid Cancer	Both	Age, sex, race/ethnicity dependent incidence rates; relative risk associated with BMI	CDC Seer Database [d] Leitzmann et al., 2010 [52]
New chronic conditions			
Alzheimer's disease (AD)		See AD section	See AD section
Asthma		See Asthma section	See Asthma section
Bipolar Disorder (BD)		See BD section	See BD section
Cancers		See Cancers section	See cancers section
Chronic Obstructive Disorder (COPD)		See COPD section	See COPD section
Congestive Heart Failure (CHF)		See CHF section	See CHF section
Updated cancers		See Cancers section	See Cancers section
Depressive Disorder		See Depressive Disorder section	See Depressive Disorder section
Myocardial		See MI section	See MI section

Outcome	Diabetic/Non-Diabetic Population	Risk Factors Considered	Source
Infarction (MI)			
Osteoporosis		See Osteoporosis section	See Osteoporosis sectin
Schizophrenia		See Schizophrenia section	See Schizophrenia section
Stroke		See Stroke section	See Stroke section
Adverse event			
Amputation	Diabetic	HbA1C, SBP, history of PVD, blindness, time since diabetes diagnosis	UKPDS Outcomes Model, Clarke et al., 2004 [33]
Blindness	Diabetic	Age at diabetes diagnosis, time since diabetes diagnosis, sex, HbA1C, SBP, cholesterol ratio	UKPDS Outcomes Model, Clarke et al., 2004 [33]
Myocardial infarction	Non-diabetic	Age, sex, SBP, smoking status, cholesterol ratio, history of LVH	Anderson et al., 1991 [53]
Myocardial infarction Stroke	Diabetic	Age at diabetes diagnosis, sex, race, smoking status, HbA1c level, SBP, cholesterol ratio, history of IHD, history of CHF, years since diabetes diagnosis	UKPDS Outcomes Model, Clarke et al., 2004 [33]
Mortality, associated with…Myocardial infarction Stroke	Both	Age, sex, SBP, antihypertensive therapy, diabetes, smoking status, MI, atrial fibrillation, LVH	Framingham Heart Study, D'Agostino et al, 1994[54]
Mortality, associated with… Cancers (Except Gallbladder)	Both	Survival rates by year since cancer diagnosis for all cancers	SEER Cancer Statistics Review [5]
Chronic kidney disease	Both	Relative risk adjustment to all-cause mortality	Tonelli et al., 2006 [55]
Congestive heart failure	Both	Age, sex	Schaufelberger et al., 2004 [56]
Diabetes	Diabetic	Time since diabetes diagnosis, MI, stroke, renal failure, or amputation history	UKPDS Outcomes Model, Clarke et al., 2004 [33]
Gallbladder Cancer	Both	Survival rates by year since diagnosis	Trends_in_1_year_survival_report_pdf in I:\Healthcare\Microsimulation Model\Disease Prevention\Literature Search\Literature for Obesity\Mortality and

[5] Relative Survival Rates by Year of Diagnosis
http://seer.cancer.gov/archive/csr/1975_2010/results_merged/topic_survival_by_year_dx.pdf

Outcome	Diabetic/Non-Diabetic Population	Risk Factors Considered	Source
			http://seer.cancer.gov/statfacts/html/pancreas.html [6]
Ischemic heart disease	Both	Age, sex, SBP, smoking status, cholesterol ratio, LVH	Anderson et al., 1991 [53]
Myocardial infarction	Both	Age, sex	Swedish MI registry [57]
Pulmonary embolism	Both	Overall mortality Rate	Carson, et al., 1992 [58]
Stroke	Both	Age, sex	Vemmosa et al., 2000 [59]
All other causes	Both	Age, sex	CDC death tables
Medical expenditures			
All cancers other than gallbladder cancer, liver cancer, multiple myeloma, and thyroid cancer	Both	Sex, cancer type	National Cancer Institute [7]
Gallbladder cancer, liver cancer, multiple myeloma, and thyroid cancer	Both	Sex, cancer type	Yabroff et al., 2008 [60]
Chronic kidney disease	Both	Average annual cost	USRDS [8]
Gallstone disease	Both	Average annual cost by complicated vs. non-complicated disease	Glasgow et al., 2000 [61]
Pneumonia	Both	Average annual cost	Colice et al., 2004 [62]
Pulmonary embolism	Both	Average annual cost	Park et al., 2009 [63]
Osteoarthritis	Both	Average annual cost by sex	Kotlarz et al., 2009 [64]
Gastroesophageal reflux disease	Both	Average annual cost	Bloom et al., 2001 [65]
Chronic back pain	Both	Average annual cost	Crow et al., 2009 [66]
Non-Alcoholic Fatty Liver Disease	Both	Average annual cost	Younoussi et al., 2014 [67]
All other conditions	Both	Age, sex, race/ethnicity, body weight category, hypertension, diabetes, CHF,	Logistic regression with MEPS/NHIS

[6] SEER Stat Fact Sheets: Pancreas Cancer http://seer.cancer.gov/statfacts/html/pancreas.html
[7] Cancer Prevalence and Cost of Care Project: Annualized Mean Net Costs of Care http://costprojections.cancer.gov/annual.costs.html
[8] United States Renal Data System: Chapter 7 Costs of Chronic Kidney Disease http://www.usrds.org/2012/view/v1_07.aspx

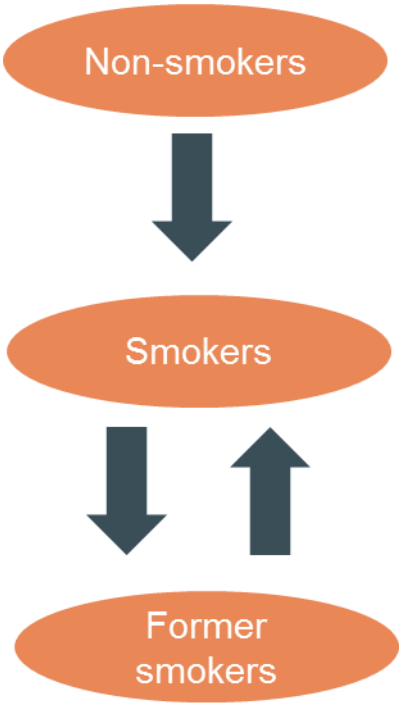
Outcome	Diabetic/Non-Diabetic Population	Risk Factors Considered	Source
		IHD, stroke, MI, blindness, renal failure, insurance, Medicaid	
Last year of life	Both	Average expenditures for the Medicare population	Riley and Lubitz, 2010 [68]
Productivity and income			
Employment	Both	Age, sex, race/ethnicity, body weight category, disease presence	Logistic regression with MEPS/NHIS
Missed work days	Both	Age, sex, race/ethnicity, body weight category, disease presence All conditions except amputation	Poisson regression with MEPS/NHIS
Household income	Both	Age, sex, race/ethnicity, body weight category, disease presence All conditions except amputation Amputation	OLS regression with MEPS/NHIS
SSI disability	Both	Age, sex, race/ethnicity, body weight category, disease presence	Logistic regression with MEPS/NHIS
Quality adjusted life years	Both	All conditions except amputation	Sullivan et al., 2006 [69] Zhang et al., 2012 [70]

Modeling Individual Characteristics

Smoking

3 smoking states will be incorporated in the model – current smoker, ex-smoker, and non-smoker (never smoked). Each individual will be able to transition from non-smokers to current smokers, and between current smokers and ex-smokers, as depicted in the chart below.

Exhibit 11. Model Diagram of Smoking States



Prevalence: The initial prevalent population at time 0 can be derived through the "smoking - cigarette use" dataset of NHANES.[71] Alternatively, it can be derived from BRFSS if there are too many missing data points in the NHANES dataset.

Incidence: For those who have *never regularly smoked*, there's a probability each year that they may become *regular smokers*. Kiefe et al., reported 10-year smoking initiation rate (percentage of baseline nonsmokers or non-regular smokers who smoked regularly at year 10) among 5,115 adults as the following:[72]

- African American female: 7.1%
- African American male: 13.2%
- White female: 3.5%
- White male: 5.1%

This study doesn't report data for other races. In the absence of better data source, it is assumed that the initiation rate for Hispanics is the same as African American population, while the rate for other races is the same as white population. The 10-year initiation rate can then be converted to annual initiation rate in Exhibit 12.

Exhibit 12. Annual smoking initiation rate

Race/Ethnicity	Gender	Annual initiation rate
Hispanics	Female	0.71%
	Male	1.32%
Non-Hispanic Black	Female	0.71%
	Male	1.32%
Non-Hispanic White	Female	0.35%
	Male	0.51%
Non-Hispanic Other	Female	0.35%
	Male	0.51%

Smoking cessation: It is assumed that only those who managed to be tobacco-free for at least 1 year are considered as "former smokers". If the individual relapsed at any given point in the year, (s)he will still be regarded as "current smokers" and thus continue to have the adverse health impacts of smoking.

According to CDC, the percentages of adult cigarette smokers who stopped smoking for more than 1 day in 2012 because they were trying to quit are as follows:[73]

- 42.7% of all adult smokers
- 48.5% smokers aged 18–24 years
- 46.8% smokers aged 25–44 years
- 38.8% smokers aged 45–64 years
- 34.6% smokers aged 65 years or older

However, the large majority of self-started quitters relapsed within a year. Garvey et al. reported that 87.2% relapsed within a year of quit dates.[74] (Exhibit 13) Consequently, those abstainers who stopped smoking for more than a year can be calculated via "% attempts to quit x (1 - % who relapse within a year) ". (Exhibit 14)

Exhibit 13. Incidence Rates of Relapse across 1 Year of Follow-up of 235 Subjects

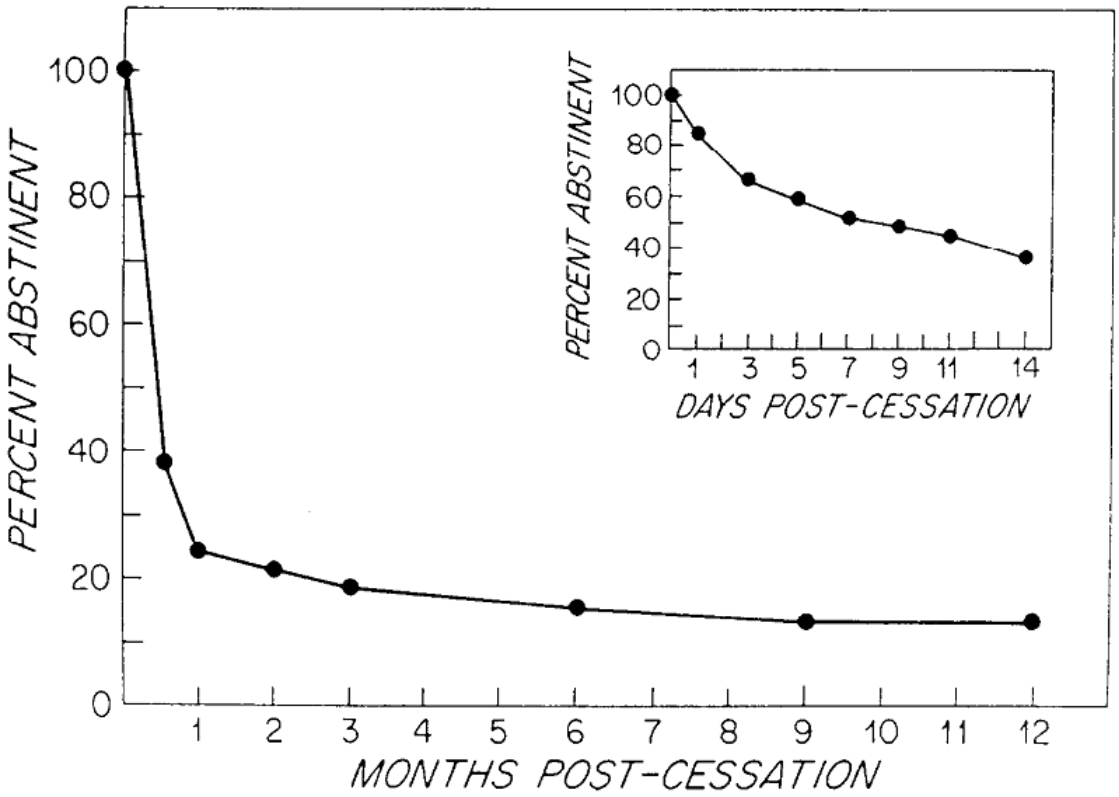


Exhibit 14. Percentage of Smokers who Stopped Smoking for More than a Year

Age group	% quit for more than a year
18-24	6.2%
25-44	6.0%
45-64	5.0%
65+	4.4%

Former smokers revert back to smoking: Swan et al. followed 329 ex-smokers (149 males, 180 females) who had maintained abstinence for at least 3 months prior to intake for 1 year to study the percentage of relapse.[75] During follow up, 33.6% of males and 32.2% of females relapsed.

Key assumptions:

- The impact of second hand smoke is not included in the study
- Those who are forced to quit smoking due to diagnosis of adverse health conditions (e.g. lung cancer, asthma, COPD) will remain non-smokers for the remainder of their life

- Only those who managed to be tobacco-free for at least 1 year are considered as “former smokers”

Alcohol use

The prevalence and magnitude of alcohol use as well as change in drinking behavior over time will be derived from BRFSS or NHANES.

BRFSS has 4 questions regarding drinking behavior:[76]

Exhibit 15. Alcohol Related Variables in BRFSS

Variable name	Question	Note
ALCDAY5	During the past 30 days, how many days per week or per month did you have at least one drink of any alcoholic beverage such as beer, wine, a malt beverage or liquor?	Days in past 30 had alcoholic beverage
AVEDRNK2	One drink is equivalent to a 12-ounce beer, a 5-ounce glass of wine, or a drink with one shot of liquor. During the past 30 days, on the days when you drank, about how many drinks did you drink on the average? (A 40 ounce beer would count as 3 drinks, or a cocktail drink with 2 shots would count as 2 drinks.)	Avg alcoholic drinks per day in past 30
DRNK3GE5	Considering all types of alcoholic beverages, how many times during the past 30 days did you have 5 or more drinks for men or 4 or more drinks for women on an occasion?	Binge drinking
MAXDRNKS	Considering all types of alcoholic beverages, how many times during the past 30 days did you have 5 or more drinks for men or 4 or more drinks for women on an occasion?	Most drinks on single occasion past 30 days

NHANES has 9 variables regarding drink behaviors in the alcohol use dataset (ALQ_G).[77] This dataset will be used to inform model inputs as it has more relevant information compared with BRFSS data. All variables will be part of the individual patient characteristics.

Prevalence of drinking behaviors: ALQ110 will be used to identify drinker (who answered year to the question) and those who never drank (who answered no to the question). ALQ101 will be used to identify current drinker (answered ‘Yes’) and past drink (answer ‘no’ to ALQ101 but ‘yes’ to ALQ110). ALQ120Q, ALQ120U, and ALQ130 can be used to estimate the level of drinking.

Exhibit 16. Alcohol Related Variables in NHANES

Variable name	Question	Note
ALQ101	The next questions are about drinking alcoholic beverages. Included are liquor (such as whiskey or gin), beer, wine, wine coolers, and any other type of alcoholic beverage.In any one year, {have you/has SP} had at least 12 drinks of any type of alcoholic beverage? By a drink, I mean a 12 oz. beer, a 5 oz. glass of wine, or one and half ounces of liquor.	Had at least 12 alcohol drinks/1 yr?
ALQ110	In {your/SP's} entire life, {have you/has he/ has she} had at least 12 drinks of any type of alcoholic beverage?	Had at least 12 alcohol drinks/lifetime?
ALQ120Q	In the past 12 months, how often did {you/SP} drink any type of alcoholic beverage? PROBE: How many days per week, per month, or per year did {you/SP} drink?	How often drink alcohol over past 12 mos
ALQ120U	UNIT OF MEASURE.	# days drink alcohol per wk, mo, yr
ALQ130	In the past 12 months, on those days that {you/SP} drank alcoholic beverages, on the average, how many drinks did {you/he/she} have?	Avg # alcoholic drinks/day - past 12 mos
ALQ141Q	In the past 12 months, on how many days did {you/SP} have {DISPLAY NUMBER} or more drinks of any alcoholic beverage? PROBE: How many days per week, per month, or per year did {you/SP} have {DISPLAY NUMBER} or more drinks in a single day?	# days have 4/5 drinks - past 12 mos
ALQ141U	UNIT OF MEASURE.	# days per week, month, year?
ALQ151	Was there ever a time or times in {your/SP's} life when {you/he/she} drank {DISPLAY NUMBER} or more drinks of any kind of alcoholic beverage almost every day? Display number = 4 for female and 5 for male	Ever have 4/5 or more drinks every day?
ALQ155	For about how many years did {you/SP} drink {DISPLAY NUMBER} or more drinks of any kind of alcoholic beverage almost every day?	How many years did you drink every day?

There's not sufficient data to model the change of drinking behavior over time on an individual level. Hence the model assumes drinking behavior remain constant over time.

The analysis target is to derive an equation to describe how individual's drinking behavior change as a person ages, by gender and race. The variable that describes the drinking behavior is the average unit of alcohol drunk per day.

Modeling Movement in Biometrics

Body Mass Index

Individuals with BMI between 25.0 and 29.9 were categorized as overweight, and individuals with BMI of 30 or above were categorized as obese.[78] A person can move in and out of the overweight and obese categories depending on change in BMI.

In the absence of an intervention that changes BMI, an individual's BMI in the next year (y+1) was modeled as a function of BMI in the current year (y), and the average change in BMI associated with aging for someone in the individual's weight category (Equation 1).

Equation 1. Annual BMI Change

$$BMI_{y+1} = BMI_y + \Delta BMI_{\Delta Age}$$

Longitudinal studies have analyzed body weight change associated with aging, trying to disentangle secular trends (e.g., changes in diet and physical activity level) from weight changes that occur naturally as a person ages. Published studies have primarily tracked people during the 1980s and 1990s—a period characterized by rapid weight gain in the U.S. and other industrialized countries as physical activity levels declined while daily energy intake rose. The CARDIA study started tracking 5,115 adults age 18-30 in 1985, and continued to have 3,499 participants through 2011. Findings suggest that over a 25-year period the average annual weight gain for those still participating in the study was 0.84 kg for black women, 0.74 kg for black men, 0.56 kg for white men, and 0.55 kg for white women, with weight gain most pronounced during early adulthood.[79] Sheehan et al. analyzed longitudinal data from a 20-year follow-up of a nationally representative sample of 5,117 adults in the National Health Examination Follow-up Study (NHEFS).[80] They found lower annual rates of weight gain between 1971 and 1991 relative to the CARDIA study, with average annual increases equating to a 0.149 BMI (0.39 kg) increase for white women and a 0.125 BMI (0.36 kg) increase for white men age 25-35 (Exhibit 17). BMI growth was higher for black adults than for white adults. The rate of weight gain slowed between ages 36 and 60, and for black women there was an average BMI reduction of 0.095/year between ages 48 and 60.

The time period covered by these studies, however, was one in which obesity rates were rising rapidly throughout the U.S. Daily energy intake of adults in the U.S. appears to have declined in recent years, with the decline most noticeable among adults age 20 to 39 with BMI between 25 and 30.[81] Likewise, the prevalence of physical activity appears to have increased in recent years among both men and women.[82] Consequently, for modeling one might expect that weight gain among individuals as they age might be slower than previously observed trends among the U.S. population.

Therefore, we used cross-sectional data in the pooled 2003-2010 NHANES to estimate the anticipated change in body weight associated with aging. We first stratified the NHANES sample by weight category (BMI<30 and BMI≥30), and then compared BMI from adjacent years of age. For each weight category combination we used Ordinary Least Squares (OLS) regression with BMI as the dependent variable and the explanatory variables consisting of a dummy variable for each age (from 20 through 85) and control variables for non-Hispanic black,

Hispanic, and male status. Control variables for other races were not found to be statistically significant. The difference in coefficients for adjacent ages was used as a proxy for the average change in BMI as a person ages. Because of small sample size for some age groups, we smoothed the BMI transition estimates using a 5-year moving average (2 years before to two years after the age). Individuals with BMI between 25 and 30 were assigned a weight change that equaled the sum of half the weight gain that a normal individual of the same age would receive, plus half the weight gain that an obese individual of the same age would receive. This methodology proved more accurate in validation exercises than having a separate weight change regression for those with BMI between 25 and 30. This approach assumed that people who are normal weight will experience an aging effect on body weight that is different than people who are overweight or obese. This assumption is consistent with published data on weight gain trends among adults in Canada, which found that that already obese men and women tend to gain less weight each year than adults who start the year overweight, and these adults in turn report less annual weight gain than a population that starts the year in the normal weight category.[83]

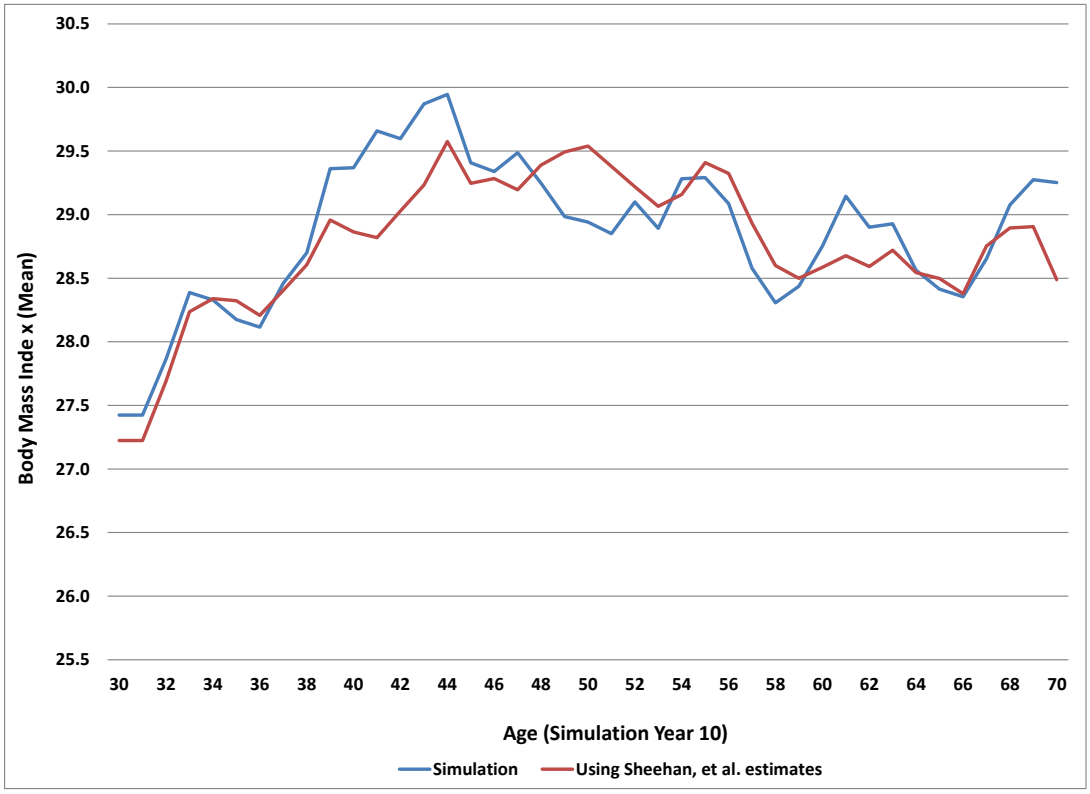
Exhibit 17 presents annual change in BMI by sex and body weight category, with the data aggregated by age group for comparison to the Sheehan et al. findings. This analysis suggests that each year men and women who are not obese tend to have slight BMI gain annually through about age 60. For those in their 60s and early 70s there is a slight BMI reduction, on average, each year.

Exhibit 17. Summary of Average Annual BMI Change

	Females				Males			
	25-35	**36-47**	**48-60**	**61-74**	**25-35**	**36-47**	**48-60**	**61-74**
NHANES analysis								
BMI <25	0.171	0.048	0.071	(0.013)	0.121	0.027	0.024	0.010
BMI 25-29.9	0.148	0.045	0.040	(0.015)	0.139	0.030	(0.055)	(0.074)
BMI ≥30	0.125	0.043	0.009	(0.018)	0.158	0.034	(0.133)	(0.157)
Overall	0.149	0.045	0.038	(0.016)	0.139	0.031	(0.064)	(0.082
Sheehan et al.[18]								
White	0.149	0.088	0.019	(0.111)	0.125	0.074	0.016	(0.093)
Black	0.206	0.095	(0.095)	0.137	0.174	0.080	(0.080)	0.116

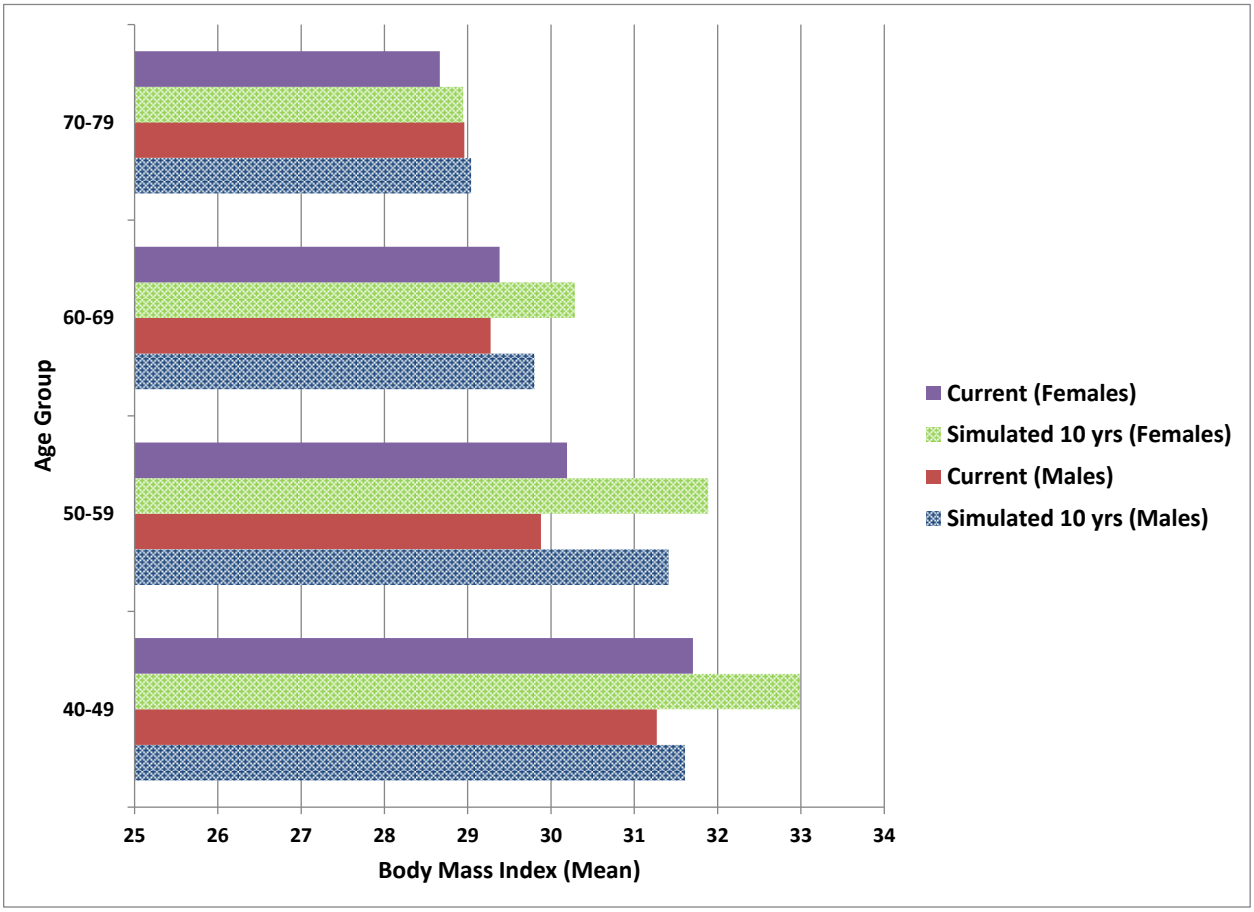
As BMI is a key input to other clinical measures, disease onset, and adverse medical events, we conducted several validation and sensitivity analyses on BMI movement associated with aging. One analysis was to project future body weight over a 10-year period for each adult in the U.S. population (regardless of diabetes or prediabetes status). For the population age 20 to 60, we projected their weight change over 10 years and then calculated their average weight for each age 30 to 70 (Exhibit 18). The differences in population average weight using the NHANES-based versus Sheehan-based weight change were a fraction of one BMI unit.

Exhibit 18. Average Simulated BMI 10 Years into the Future



Another validation activity was to simulate weight over 10 years, and then compare average simulated BMI among the prediabetic adult population with average BMI for the current population with prediabetes (Exhibit 19). Because of small sample sizes for individual ages, the results are shown by 10-year age band for the population age 40 to 79. Projected average BMI tracked closely with actual average BMI for the age 60 to 69 population. Across all age groups the projected average BMI was about 0.13 BMI lower than current average BMI. Simulated BMI was higher, on average, than BMI observed in the current prediabetic population especially for females. Any over-prediction (or under-prediction) in BMI or any of the other patient risk factors or adverse medical events will affect both the intervention and non-intervention scenarios, and any bias caused by over (under) prediction will largely be cancelled out when comparing outcomes between the intervention and non-intervention scenarios.

Exhibit 19. Actual versus Forecasted Average BMI by Age



Smoking status will also affect BMI change. If a person quits or picks up smoking in a specific year, smoking-induced BMI change will overwrite the change caused by aging. Within a year of smoking cessation, BMI will increase by approximately 1.67 (standard error 0.19) and 3.15 (standard error 0.34) units for male and female quitters, respectively.[1]

Blood Pressure

The hypertension indicator in the model is initiated in the first year of the simulation if the person's NHANES blood pressure readings meet the threshold for hypertension (SBP≥140mmHg or DBP≥90mmHg), if the person indicates he or she is taking antihypertensive medicine, or if the person indicates having been told by a health professional that he or she has hypertension.[84] During the simulation, hypertension is indicated if a person's SBP or DBP readings reach the above thresholds in that year. If in the following year their SBP or DPB fell below the hypertensive range, hypertension would no longer be indicated.

The link between hypertension and cardiovascular disease risk has been well established.[85-92] Annual change in blood pressure level was modeled separately for the population without diabetes and the population who had experienced diabetes onset. Systolic blood pressure (SBP) tends to have a stronger association with disease

onset than diastolic blood pressure (DBP), and in disease onset equations from the UKPDS Outcomes Model, only SBP is included as a risk factor.[93, 94-96] For the non-diabetes population, both SBP and DBP were modeled, as published equations for disease onset included both measures (with SBP used more frequently).[85-92]

For people without diabetes, we modeled annual change in SBP as a function of aging and change in BMI. Neter et al. examined 25 randomized clinical trials and estimated that a 1kg loss in body weight was associated with a 1.05 mmHg reduction in SBP.[8] To model the relationship between aging and SBP, while holding BMI constant, we used OLS regression with the pooled NHANES file to fit separate trend lines for men and women (Equations 2 and 3). Model validation activities, discussed later, suggested that that the model was over-predicting growth in SBP among the population under age 40 and under predicting among the population over age 40. Therefore, model calibration suggested reducing the age coefficients by one standard error for the under 40 population and increasing the age coefficients by one half a standard error for the population over age 40. This calibration adjustment led to a much better fit in hypertension incidence rates when comparing model outputs to published statistics.

Equation 2: Annual SBP Change Associated with Aging (Males)

$$SBP_{male} = 118.11+0.2147age +0.0032age^2$$

Equation 3: Annual SBP Change Associated with Aging (Females)

$$SBP_{female} = 107.16+0.5027age + 0.0034age^2$$

For the population with diabetes, the modeling of SBP was based on equations from the UKPDS Outcomes Model.[33]

To model change in DBP due to aging, we derived equations based on NHANES data using OLS regression techniques. One equation could not sufficiently explain the statistical variation between male and female populations, and thus a separate equation was derived for each gender. The dependent variable was DBP. Explanatory variables included a continuous BMI variable and dummy variables for each year of age, whether the individual was taking anti-hypertensive medication, race and Hispanic ethnicity. Explanatory variables were selected clinically and statistically. The aforementioned variables were shown to be significantly correlated with DBP in explorative regression analysis. Additional literature review revealed good clinical face validity for the chosen variables. Comparison of adjacent age coefficients provides estimates of how DBP changes with age (controlling for BMI). A trend line was fit to age coefficients (Equations 4 and 5):

Equation 4. Annual DBP Change Associated with Aging (Males)

$$DBP_{male} = 0.5881+0.0177age - 0.0025age^2+ 0.00003age^3$$

Equation 5. Annual DBP Change Associated with Aging (Females)

$$DBP_{female} = 0.3808 + 0.0268age - 0.0021age^2 + 0.00002age^3$$

Whether or not an individual with hypertension was on anti-hypertensive therapy is used as a predictor for stroke and CHF. To estimate the probability that an individual with hypertension is on an anti-hypertensive therapy, logistic regression analysis was conducted on the pooled 2003-2010 NHANES sample. Separate regressions were run for males and females where the dependent variable was whether or not an individual was on a therapy, and the explanatory variables were age, BMI, race, and insurance status. The odds ratios associated with these regressions are presented in Exhibit 20.

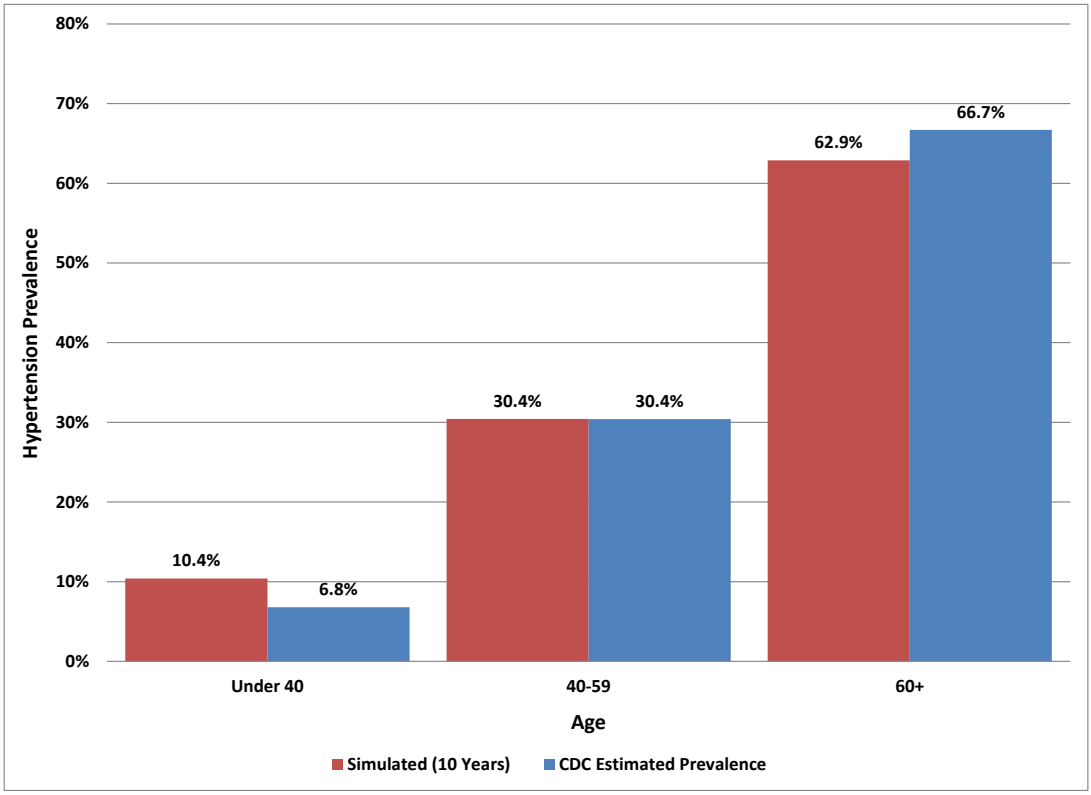
Exhibit 20. Logistic Regression Results for Predicting Anti-Hypertensive Therapy

	Males		Females	
Parameter	**Odds Ratio**	**95% C.I.**	**Odds Ratio**	**95% C.I.**
Age	1.054	1.044, 1.064	1.067	1.057, 1.078
BMI	1.023	1.001, 1.046	1.035	1.015, 1.054
Hispanic	0.592	0.421, 0.832	0.741	0.525, 1.044
Non-Hispanic Black	0.886	0.640, 1.226	1.085	0.784, 1.502
Non-Hispanic Other	0.898	0.434, 1.859	0.825	0.435, 1.565
Insured	1.664	1.271, 2.178	1.783	1.364, 2.331
Diagnostics				
Sample size	2,046		2,314	
Akaike information criterion	1,491		1,516	
Percent concordant	71.7%		74.4%	

For validation, we simulated hypertension prevalence 10 years into the future for a representative sample of all U.S. adults regardless of diabetes or prediabetes status. We then compared simulated prevalence estimates with current prevalence rates reported by the Centers for Disease Control and Prevention by age group (Exhibit 21).[9] For the population under age 40 the simulated prevalence rates are about 3.6 percentage points higher than current hypertension prevalence reported by CDC; the rates are identical for the age 40-59 population, and the simulated rates are 3.8 percentage points lower than CDC-reported rates for the age 60 and older population.

[9] http://www.cdc.gov/nchs/data/databriefs/db107.pdf

Exhibit 21. Hypertension Prevalence Comparison

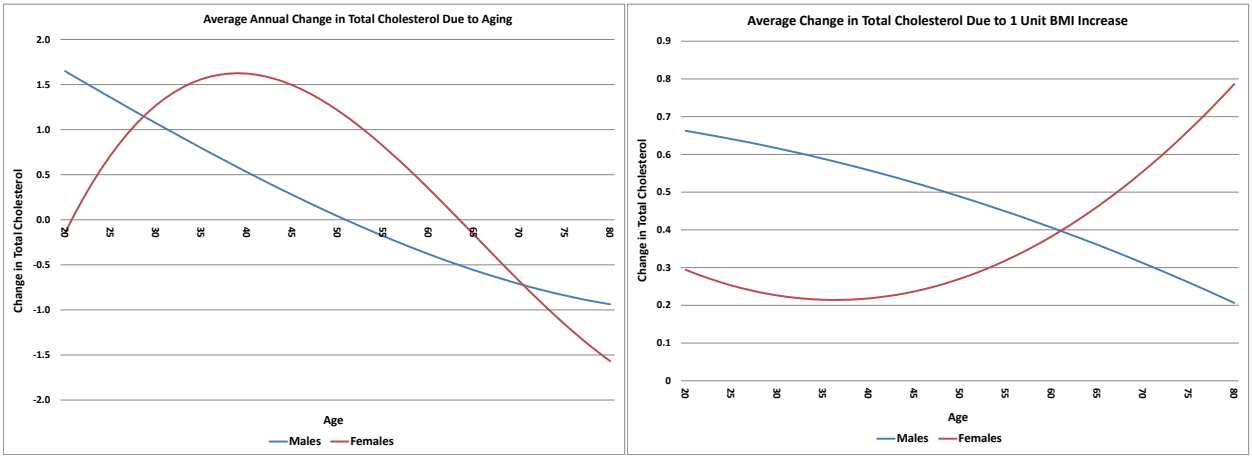


Cholesterol

Cholesterol is a risk factor for multiple cardiovascular conditions, and was modeled separately for the populations with and without diabetes. Once an individual experienced diabetes onset, then cholesterol ratio (total cholesterol divided by HDL cholesterol) was modeled as an input to published equations from the UKPDS Outcomes Model.[33] For the population without diabetes, we modeled total cholesterol, cholesterol ratio, and HDL cholesterol (which is a risk factor for CHF and also used to calculate cholesterol ratio), by modeling cholesterol change due to age and BMI separately, using different equations for men and women.

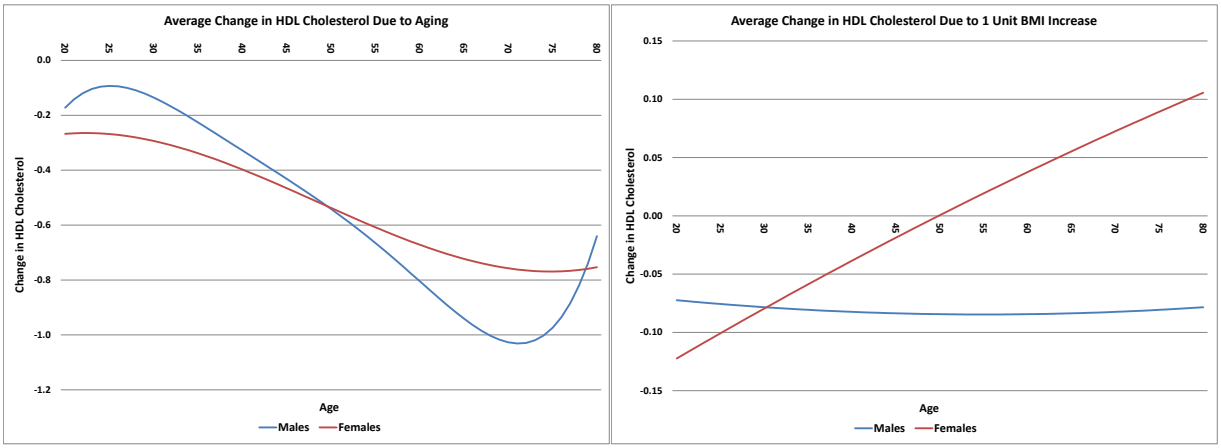
The change in total cholesterol due to both aging and change in BMI come from work by Wilson et al. based on the Framingham Heart Study.[14] The authors reported 8-year change in total cholesterol stratified by age group, as well as the age-adjusted effect of a 1-unit BMI change on 8-year change in total cholesterol. These 8-year changes were first converted to annual changes by dividing by 8, using the simplifying assumption that the annual effect was equal across the interval. Because the trend was non-linear across ages, a polynomial trend line was then fit to these annual estimates and used to estimate the changes in total cholesterol due to aging and BMI changes (Exhibit 22). The total annual change in total cholesterol is the sum of the aging and BMI effect, and the relationships are quite different for men versus women.

Exhibit 22. Average Changes in Total Cholesterol



HDL cholesterol was modeled similarly to total cholesterol using findings from Wilson et al. [14] The changes in HDL cholesterol associated with aging and changing BMI for men and women are illustrated in Exhibit 23.

Exhibit 23. Average Changes in HDL Cholesterol



Hemoglobin A1c

HbA1c was chosen as the main measure of glucose control for multiple reasons. One, the pooled NHANES sample had higher response rate and consequently a larger sample of people who had valid measures of HbA1c compared to fasting plasma glucose (FPG) level (19,505 responses vs. 9,396). Two, the UKPDS Outcomes Model was the primary source for prediction equations to model diabetes sequelae and these equations used HbA1c. Three, a literature search yielded more articles using HbA1c as an input than using FPG or 2-hour oral glucose

tolerance test (OGTT). For people who develop diabetes, the equation used to model subsequent changes in HbA1c came from the UKPDS Outcomes Model.[33]

There is a paucity of literature around the way that HbA1c levels change over the course of an individual's life. To inform this parameter, an analysis of the mean change in body weight and mean change in HbA1c levels from the CONQUER trial was used. The CONQUER trial was a randomized, double-blind, placebo control trial that enrolled individuals with BMI of 27-45 in the U.S. between 2007 and 2009. Our analysis indicated that over the 56 week period of trial an average change of 1kg in weight was associated with a 0.071% change in HbA1c levels.

Heinza et al. indicated that annual rate of change in HbA1c levels for those who develop diabetes and those who do not develop diabetes only differs significantly in the year before diabetes onset.[12] In the year before onset, a roughly 0.41 jump in HbA1c levels was observed. A separate article by Heinza indicates that the annual incidence rate of diabetes among those with HbA1c levels of 6.0 to 6.4 was 12.92%.[13] As such, each year every individual in this 6.0-6.4 range has a 12.92% probability of a 0.41 increase in HbA1c level (versus the increase calculated from the change in BMI).

An HbA1c cutoff of 6.5% was used to indicate clinical diabetes, while a fasting plasma glucose (FPG) cutoff of 7 mmol/L was used.[97] Once an individual experienced diabetes onset, they continued to be categorized as having diabetes even if they subsequently reduced their HbA1c or FPG levels below the threshold. Validation activities involving HbA1c, described later, include comparisons of simulated diabetes incidence to published estimates.

Fasting Plasma Glucose

Our analysis of NHANES data has shown that using HbA1c alone to identify a population with diabetes or prediabetes tends to skew toward diagnosis of an older segment of said population. Therefore, FPG was also included in the model and used to identify individuals. Similar to HbA1c, our literature search did not yield any usable published parameters to model the way that FPG changes over the course of an individual's life. Danaei et al. conducted a systematic review of health examination surveys and epidemiological studies focused on FPG and diabetes.[9] While they did not focus on investigating the factors that can predict an individual's change in FPG, they did conduct a regression analysis that estimated FPG from HbA1c. For some individuals in the NHANES sample with an HbA1c value but no FPG value the prediction equation uses HbA1c, age and gender to predict an FPG value. The prediction equation is shown below.

$$FPG_T = 1.034 + 0.84*HbA1c_T + 0.034*YearPost2005 + 0.0051*Age - 0.19*Female$$

$$\Delta FPG_T = 0.0391 + 0.84*\Delta HbA1c_T$$

Modeling Disease and Adverse Events

Alzheimer's Disease

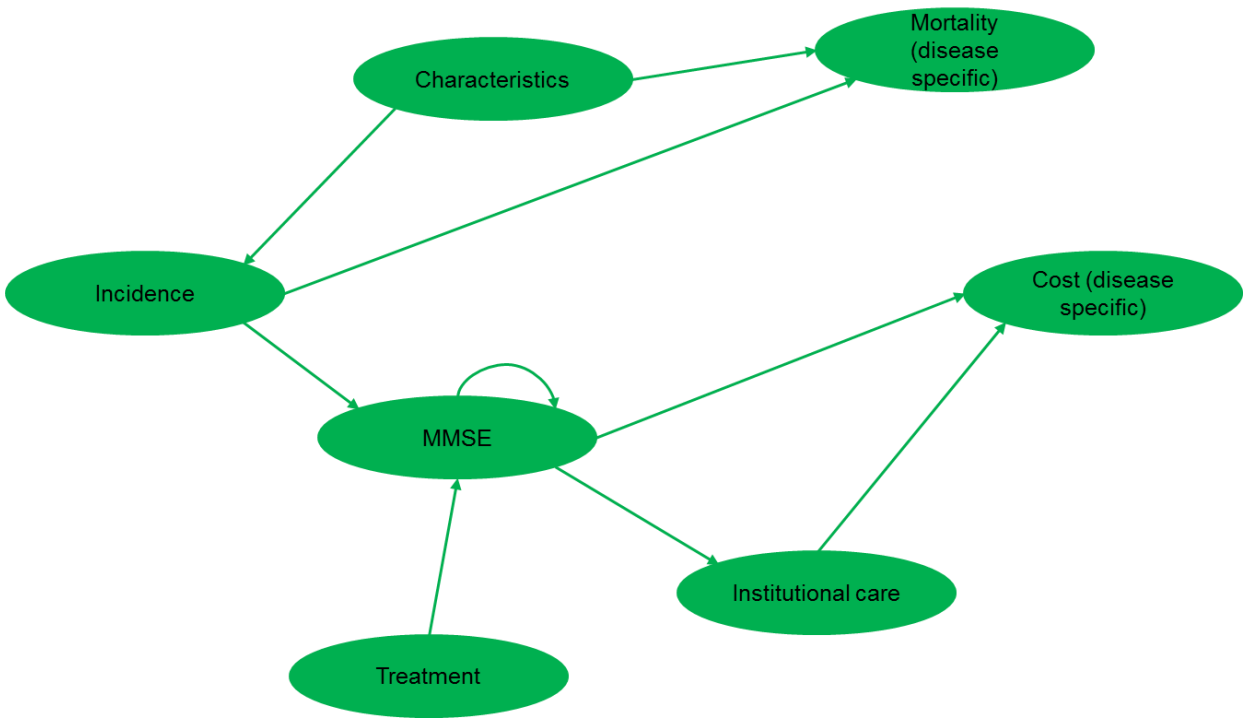
The modeling of Alzheimer's disease (AD) will follow a similar structure as the NICE HTA submission of donepezil by Eisai/Pfizer in 2010.[98] In the submission the disease is characterized by MMSE (Mini-Mental State Examination) scores.[99]

Exhibit 24. MMSE Scores and Severity of AD

MMSE range	AD severity
21-26	Mild
10-20	Moderate
<10	Severe

The simulation of the disease is based on the progression of MMSE over time with or without treatment (Exhibit 25).

Exhibit 25. Influence diagram of AD



Initial prevalence: Many epidemiology studies found AD to be more prevalence in women than in men. The prevailing explanation for this is that on average women have longer life spans than men and are thereby more

likely to reach an age of high risk for AD. There is no evidence that one gender is more likely to develop dementia at any given age.[100]

96% of all AD patients are age 65 and older.[101] In 2006 there were only 200,000 AD patients who are younger than age 65 (prevalence rate 7/100,000). Due to this extremely low prevalence we assume only those aged 65 and older can get AD.

The prevalence of dementia by age group and race is depicted in Exhibit 26. The source didn't report any data on the ethnic group "Non-Hispanic Other". To be conservative we assume it has the same prevalence as "White" population, which has the lowest known prevalence of all races. Because AD accounts for an average of 70% of all dementia cases,[101] the prevalence of AD can be calculated in Exhibit 27.

Exhibit 26. Proportion of people aged 65 or older with dementia[101,102]

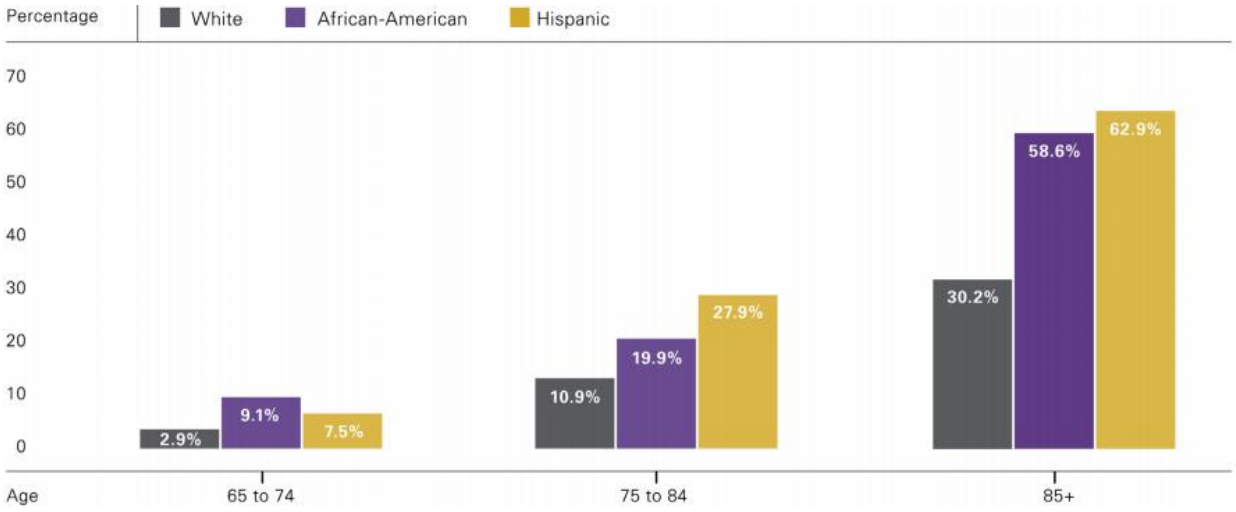


Exhibit 27. Prevalence of AD by age and race

Age group	Race/ethnicity	Prevalence
65-74	Hispanic	5.3%
	Non-Hispanic white	2.0%
	Non-Hispanic black	6.4%
	Non-Hispanic other	2.0%
75-84	Hispanic	19.5%
	Non-Hispanic white	7.6%
	Non-Hispanic black	13.9%
	Non-Hispanic other	7.6%
85+	Hispanic	44.0%
	Non-Hispanic white	21.1%
	Non-Hispanic black	41.0%
	Non-Hispanic other	21.1%

Because the progression of AD is highly correlated with age, it is assumed that younger prevalent population also has milder disease. A MMSE score will be randomly generated for each age group. Age group 65-74 will be assigned a randomly generated MMSE score between 21 and 26 (inclusive, equal probability for each score). By the same token, age group 75-84 will be randomly assigned a score between 10-20, and age group 85+ will be between 1-10.

Incidence: It was projected that in 2014, there will be approximately 59,000 new cases among people aged 65 to 74 years (incidence rate 224/100,000), 172,000 new cases among people aged 75 to 84 years (incidence rate 1,260/100,000), and 238,000 new cases among people aged 85 years and older (Incidence rate 3,887/100,000).[103] New AD cases are assumed to have the mildest disease (MMSE 26).

Disease progression and treatment effect: Because AD is irreversible, MMSE will decline continuously after disease occurrence. The annual rate of MMSE decline with and without treatment (donepezil) is as follows:[98]

> Annual decline in MMSE = Tx_effect + norm(0,0.5) - 0.429PM1 – 0.004PM2 + 0.1415PM3 – 0.079PrevMMSEChange + 0.0747Ageorig

Among the variables, norm(0,0.5) is a standard normal distribution with a standard deviation of 0.5. This represents the random variation in treatment effects among individuals. Tx_Effect is a constant with the value being 2.4671 for treated and 0 for untreated. PM1, PM2 and PM3 are the individual's previous MMSE score partitioned over the scale of MMSE. PM1= min(PrevMMSE,9), PM2=max(0,min(PrevMMSE-9,9)), PM3=max(0,min(PrevMMSE-18,12)). PrevMMSEChange is the individual's last known MMSE decline. Ageorig is the age at baseline (age of disease incidence for those developed the disease during the course of simulation, or age at time 0 for those came into the model with AD).

The % population under treatment is unclear and thus needs to be calibrated. Calibration target is the total annual direct medical cost attributable to AD in the US, which is estimated to be $218.6billion (2015 USD).[100]

Mortality: Bowne et al. followed up 327 newly diagnosed AD patients for a median of 3.3 years and compared their mortality rate with a comparable community population.[104] The reported RR of death for every 5-point increase in MMSE is 1.4 (95% CI: 1.2-1.7). To give more granularity we derived the RR of death for every point of increase in MMSE to be 1.4^(1/5)=1.07 with the assumption that an AD patient with an MMSE score of 26 (mildest case) has the same mortality as the general population.

Because mortality *among* AD patients is different from mortality *due to* AD, AD-specific death can be calculated by subtracting all-cause death from death *among* AD patients.

> Death due to AD = All cause death for AD patients – All cause death a community population

For example, for someone with an MMSE score of 20, the RR of death *due to* AD is 1.07^(26-20)-1=0.50. The probability of dying *due to* AD is 0.50 * all-cause mortality from the life table. (See appendix. "Non-Hispanic Other" population will use the national life table for males and females)

Cost: Cost drivers of AD include community based care and institutionalized care. The percentage of people in community based or institutional care were reported to be as follows:[99]

Exhibit 28. Community-based Care and Institutional Care by MMSE Score

MMSE score	Severity scale	Home (%)	Institutional care (%)
25–30	Mild	87.1	12.9
20–24	Mild to moderate	74.4	25.6
15–19	Moderate	61.7	38.3
10–14	Moderate to severe	49.0	51.0
0–9	Severe	30.0	70.0

The annual direct medical cost of community based care and institutional care is calculated by Alzheimer's Association as follows:[100]

Exhibit 29. Annual direct medical cost of AD by setting

table 8 Average Annual Per-Person Payments for Health Care and Long-Term Care Services, Medicare Beneficiaries Age 65 and Older, with and without Alzheimer's Disease and Other Dementias and by Place of Residence, in 2013 Dollars

	Beneficiaries with Alzheimer's Disease and Other Dementias by Place of Residence			Beneficiaries without Alzheimer's Disease and Other Dementias
Payment Source	Overall	Community-Dwelling	Residential Facility	
Medicare	$21,095	$18,787	$24,319	$8,005
Medicaid	10,771	237	25,494	561
Uncompensated	290	417	114	328
HMO	1,058	1,642	241	1,543
Private insurance	2,407	2,645	2,074	1,619
Other payer	964	174	2,067	153
Out of pocket	9,970	3,370	19,196	2,431
Total*	**46,669**	**27,465**	**73,511**	**14,772**

*Payments from sources do not equal total payments exactly due to the effect of population weighting. Payments for all beneficiaries with Alzheimer's disease and other dementias include payments for community-dwelling and facility-dwelling beneficiaries.

Created from unpublished data from the Medicare Current Beneficiary Survey for 2008.[155]

The increased cost compared to those without AD is directly related to the disease. Consequently, AD-specific cost can be calculated as follows:

- Annual direct medical cost for community dwelling patients: ($27,465-$14,772)*(444.65/425.13)=$13,276. The allocation of this cost to different settings (I/P, O/P, Rx, etc.) will the derived from a generic analysis on MEPS data.
- Annual direct medical cost for institutionalized patients: ($73,511-$14,772)*(444.65/425.13)=$**61,436**

Because all AD patients are over 65 years old, it is assumed they incur no absenteeism cost. The indirect burden of AD is mainly caused by the absenteeism of family members who provide care to the *community-dwelling* patient.

The number of AD patients was estimated to be approximately 5 million in 2014, who collectively received 17.7 billion hours of unpaid care from family and other unpaid caregivers.[100] This translates into 3,540 hours of unpaid care per patient per year. Each hour of unpaid care is valued at $13.02 per hour (inflated from 2013 cost)[100], resulting in a total unpaid care giver cost of 3,540*$13.02= $46,090 per year (2015 cost).

Key assumptions:

- Because the prevalence of AD is 0.007% in the population younger than 65, we assume only those aged 65 and older can get AD
- By the same token, we assume AD patients incur no absenteeism cost.
- Only those living in community incurs caregiver absenteeism cost
- Because the progression of AD is highly correlated with age, older prevalent populations are assumed to have more severe disease than younger prevalent population
- An AD patient with an MMSE score of 26 has the same mortality rate as the general population

Amputation

Amputation was only modeled for the population with diabetes. The model assumes that amputations not attributable to diabetes (e.g., from trauma) have equal probability among the diabetes and non-diabetes populations. Incidence of amputation among a diabetes population come from the UKPDS Outcomes model and used a Weibull model.[33]

Asthma

The simulation of asthma is centered on control status and its associated risk of exacerbations, as shown in Exhibit 30. Model control status is defined by widely accepted GINA guideline.[10] (Exhibit 31) All asthma cases will have a definitive duration, which will be calibrated so that the overall prevalence of asthma remain constant over time.

[10] Global Initiative for Asthma (GINA), Pocket Guide for Asthma Management and Prevention, 2010, http://www.ginasthma.org/local/uploads/files/GINA_Pocket_2010a_1.pdf, accessed Oct 30, 2015

Exhibit 30. Influence diagram for Asthma

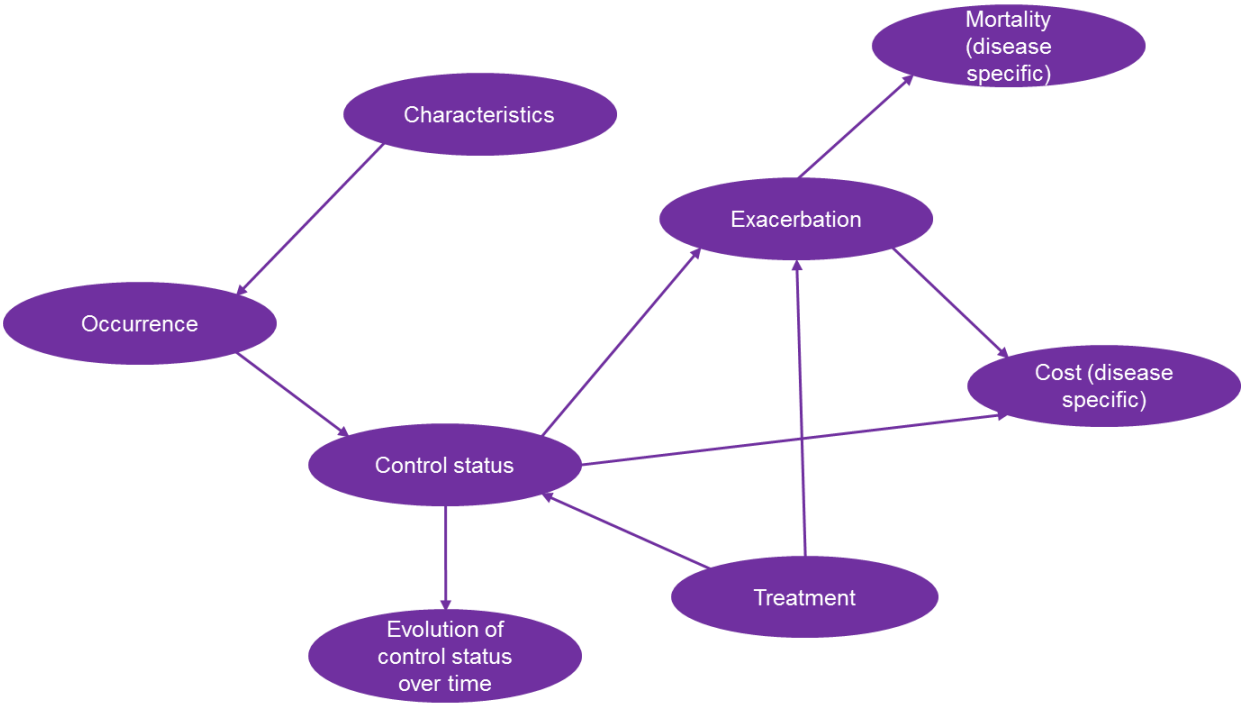


Exhibit 31. Definition of asthma control status by GINA guideline

Figure 2. LEVELS OF ASTHMA CONTROL			
A. Assessment of current clinical control (preferably over 4 weeks)			
Characteristic	**Controlled (All of the following)**	**Partly Controlled (Any measure present)**	**Uncontrolled**
Daytime symptoms	None (twice or less/week)	More than twice/week	Three or more features of partly controlled asthma*†
Limitation of activities	None	Any	
Nocturnal symptoms/awakening	None	Any	
Need for reliever/ rescue treatment	None (twice or less/week)	More than twice/week	
Lung function (PEF or FEV_1)‡	Normal	<80% predicted or personal best (if known)	

Prevalence: the initial prevalence of asthma and asthma history can be derived from BRFSS questions “Ever told had asthma” and “Still have asthma”.

The distribution of control state is available from Thomas et al's study[11] on outpatients in EU5 and USA. It is reported that the percentage of controlled, partly controlled, and uncontrolled is 49%, 32%, 20%, respectively.

Incidence: Winter et al. reported the analytical method to estimate asthma incidence for adults based on BRFSS Asthma Call-back Survey (ACBS) data.[12] Number of incidence cases (numerator) can be estimated from those who answered 'yes' to BRFSS's "life time asthma" question and also said they were told to have asthma within the past 12 months in ACBS.

The analysis was done on pooled ACBS population from 2011-2013[13] (all entries with missing values have been removed) using logistic regression with age, gender, race/ethnicity and BMI as independent variables. The final equation is in Exhibit 32.

Exhibit 32. Regression equation to predict asthma incidence

	Maximum Likelihood Estimate	Standard Error	p
Intercept	-3.9322	0.0652	<.0001
Male	-0.1586	0.0421	0.0002
Age groups			
Age 18-34	-0.0796	0.1006	0.4288
Age 35-44	-0.1492	0.1017	0.1423
Age 45-54	0.149	0.0748	0.0463
Age 55-64	Comparison group		
Age 65+	0.0698	0.0637	0.2732
Race/Ethnicity			
Hispanic	0.0443	0.1102	0.6873
Non-Hispanic Black	-0.0597	0.1103	0.5886
Non-Hispanic other	0.025	0.1113	0.8226
Non-Hispanic White	Comparison group		
Weight category			
Normal weight	-0.0886	0.0565	0.1171
Over weight	-0.0212	0.0539	0.6946
Obese	Comparison group		

The distribution of control status in the incidence population is assumed to be the same as in the prevalence population above.

[11] Thormas M et al.The Asthma Control TestTM (ACT) as a predictor of GINA guideline-defined asthma control: analysis of a multinational cross-sectional survey, Prim Care Resp J, 2009

[12] Winer, RA, et al., Asthma Incidence among Children and Adults: Findings from the Behavioral Risk Factor Surveillance System Asthma Call-back Survey—United States, 2006–2008, Journal of Asthma, 49: 16-22, 2012

[13] Centers for Disease Control and Prevention, http://www.cdc.gov/brfss/acbs/index.htm, July 2, 2015, accessed Nov 2, 2015

Natural course of the disease: Based on the GOAL study, Bateman ED reported in 2010 ERS congress the following *weekly* transition diagram between controlled, partly controlled, not controlled, and exacerbation, summarized in Exhibit 33 and Exhibit 34.[14] These patients are all treated with current standard of care for asthma (usual dose of inhaled corticosteroids). Simulation of state transition and exacerbation will be conducted on a weekly basis to match the data source.

Treatment effect: Bateman et al. reported no statistical difference between treatments in terms of the change in control status.[14] The primary difference between treatment arms is the risk of exacerbation. Based on these findings it is assumed that the transition probability between treated and untreated arms is the same. Based on an analysis of 16,941 patients in HMO environment, inhaled steroids can generally reduce the probability of exacerbation by half (RR 0.5, CI 0.4-0.6).[15] Consequently, the risk of exacerbation can be converted for the untreated population (Exhibit 33).

Exhibit 33. Weekly Transition Matrix between Asthma States (Treated And Untreated Populations)

To: / From:	Controlled	Partly controlled	Uncontrolled
Controlled	74.3%	19.8%	5.9%
Partly controlled	7.8%	72.8%	19.4%
Uncontrolled	2.0%	14.7%	83.3%

Exhibit 34. Weekly Risk of Exacerbation

	Treated Population	Untreated Population
Controlled	0.14%	0.28%
Partly controlled	0.27%	0.53%
Uncontrolled	0.79%	1.58%

Note: unweighted average exacerbation rate from a population treated with ICS+SABA, ICS/LABA+SABA, and budesonide/formoterol

Mortality: Ivanova et al.[16] studied a population of moderate to severe persistent asthma, and reported that 34.2% had at least 1 asthma exacerbation and 1.9% had at least 1 patient stay. The probability of being hospitalized among those with exacerbation is 1.9%/34.2% = 5.6%

The probability of death during hospitalization is 3.1% according to Lowhagen et al.[17]

[14] Bateman ED, et al., Overall asthma control: the relationship between current control and future risk, J Allergy Clin Immunol, 2010, 125(3)
[15] Donahue, JG, et al., Inhaled Steroids and the Risk of Hospitalization for Asthma, JAMA, 1997
[16] Ivanova, JI, et al., Effect of asthma exacerbations on health care costs among asthmatic patients with moderate and severe persistent asthma, J Allergy Clin Immunol, 2012
[17] Lowhagen O, Ekstrom L, Holmberg S, Wennerblom B, Rosenfeldt M. Experience of an emergency mobile asthma treatment programme. Resuscitation 1997;35:243–247.

Cost: Direct cost of asthma includes routine care cost (Rx and regular outpatient visits) and exacerbation cost. Barnett et al reported the annual cost of asthma by setting as follows:[18]

Exhibit 35. Cost of Asthma by Setting (2009 USD)

TABLE II. Estimates from 2-step GLM models of the per-person incremental costs of asthma by expenditure category (2009$)

	Total expenditure	Prescription medication	Office-based visits	Emergency department visits	Outpatient visits	Inpatient visits
2002	4,193* (0 to 5,276)	1,583* (1,314 to 1,916)	646* (485 to 814)	143 (90 to 197)	201 (73 to 351)	985 (488 to 1,573)
2003	3,413* (0 to 4,142)	1,867* (1,563 to 2,207)	636* (480 to 810)	113 (67 to 166)	220 (66 to 369)	329† (21 to 758)
2004	3,623* (0 to 4,880)	1,779* (1,517 to 2,113)	585* (394 to 803)	74‡ (35 to 121)	63‡ (−35 to 178)	68‡ (−383 to 545)
2005	2,745* (2,012 to 3,748)	1,567* (1,335 to 1,929)	483* (303 to 689)	155 (97 to 228)	153 (24 to 313)	−18† (−411 to 398)
2006	2,582* (2,012 to 3,258)	1,490* (1,295 to 1,803)	515* (381 to 696)	66 (31 to 104)	167 (48 to 307)	349 (−13 to 752)
2007	3,856* (2,821 to 5,153)	1,861* (1,571 to 2,230)	632* (399 to 931)	124 (64 to 195)	143 (39 to 262)	1,137 (406 to 1,961)
Pooled sample	3,259* (2,912 to 3,676)	1,680* (1,557 to 1,827)	581* (505 to 661)	110 (90 to 133)	151 (98 to 211)	446‡ (236 to 655)

Notes: The incremental cost of asthma is estimated from the 2-step model (step 1, logit; step 2, GLM with gamma distribution and log link), with both steps including age groups, married status, minority race, region, level of education, female sex, poverty, uninsured status, and Charlson comorbidity index as covariates. Asthma was positive and significant in all logit models. Ninety-five percent bootstrap CIs are presented in parentheses.
*Asthma was positive and significant at 1% in GLMs
†Asthma was negative and significant at 1% in GLMs.
‡Asthma was negative and significant at 5% in GLMs.

People with asthma need to visit outpatient unit, ED or get hospitalized when suffering from exacerbation. It is therefore assumed that in Exhibit 35, prescription medication and office-based visits are routine care while ED/outpatient/inpatient visits are due to exacerbation.

Routine care cost can thus be summarized in the following table:

Exhibit 36. Asthma Routine Care Cost (2015 USD)

	Prescription medication	Office-based visits (other cost)
Annual cost	$1,989	$688

It was estimated that 80% of exacerbations require outpatient visits only, and that US patients made 14.7 million outpatients visits a year due to exacerbation.[19] The total number of exacerbations can thus be calculated as 14.7/80% = 18.375 million. Considering there are 22.2 million asthma patients in the US, the average number of exacerbations a patient has in a year is 18.375/22.2 = 0.83.

Average ED, outpatient, and inpatient cost per exacerbation can thus be calculated as (using inpatient cost as an example): Average annual cost of inpatient cost/0.83.

[18] Barnett, SB, et al., Cost of asthma in the United States: 2002-2007, J Allergy Clin Immunol, 2011
[19] Dougherty, RH, Fahy, JV, Acute exacerbations of asthma: epidemiology, biology and the exacerbation prone phenotype, Clin Exp Allergy, 2009

Exhibit 37. Asthma exacerbation cost (per case, 2015 USD)

	ED visit	Outpatient visit	Inpatient visit
Annual cost	$157	$215	$636

In terms of indirect cost, Barnett et al. found asthma to be associated with 2.62 missing work days every year. [18]

Assuming each outpatient and ED visit is associated with 0.5 lost work day. The average length of ED visit and hospitalization due to asthma is reported to be 2.8 days. [20]The distribution of exacerbation-related resource use is as follows:

Exhibit 38. Exacerbation Related Absenteeism (Per Case)

Resource use	% of exacerbation	Lost works days
Outpatient visit	80%	0.5
ED visit	14.4%	0.5
ED visit and hospitalization	5.6%	2.8
Weighted average	-	**0.63**

Since an average asthma patient in the US will have 0.83 exacerbations every year, average number of lost work days due to exacerbation is 0.83*0.63=0.53 days.

In summary, asthma related absenteeism is 2.62-0.53=2.09 days without exacerbation, and 0.63 days per exacerbation.

Key assumptions:

- It is assumed that only death due to asthma only occurs to those who are hospitalized
- Treatment is assumed not to affect transition probabilities of control status. Treatment efficacy is in reduced probability of exacerbation
- Exacerbation cost covers outpatient visits, ED visits and hospitalization

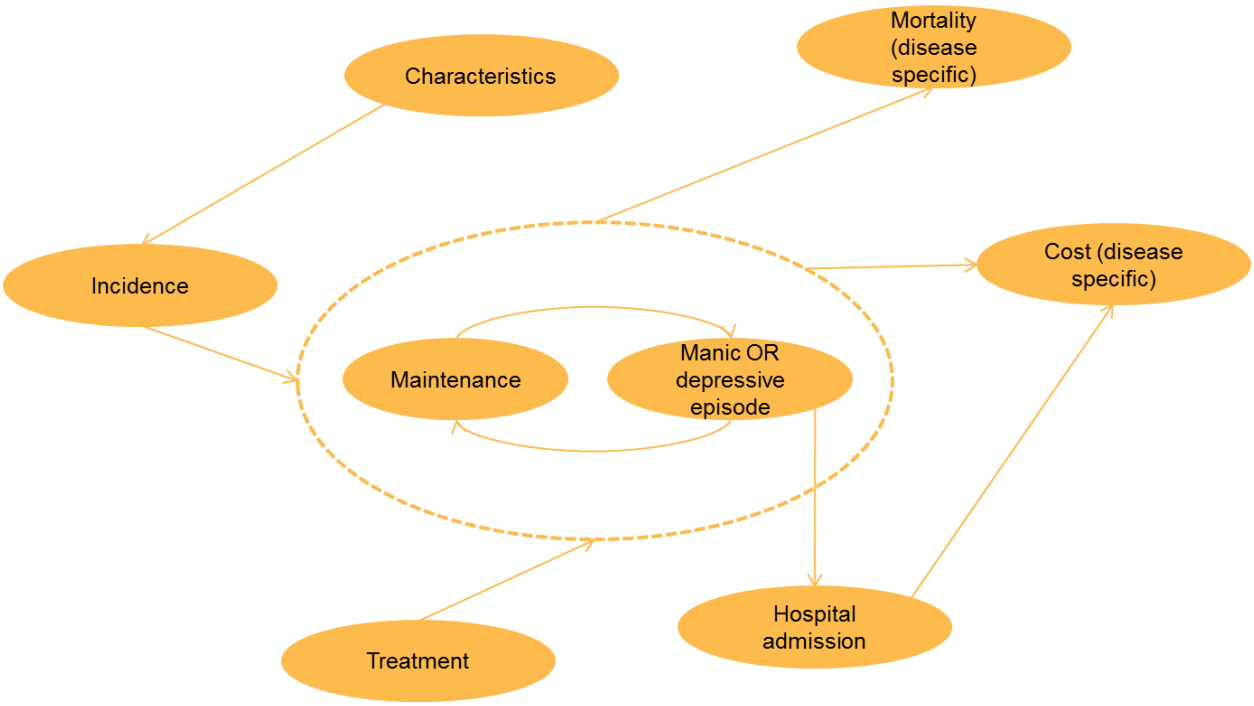
Bipolar Disorder (BD)

Bipolar disorder does not currently have a cure, and will be modeled as a life-long disease in the DPMM. At present, treatment for the condition focuses on managing mood swings and associated symptoms in order to decrease the frequency and severity of episodes of depression and mania.[21] The simulation of BD will focus on maintaining condition stability as shown in Exhibit 39 .

[20] Stanford, R. et al., The cost of asthma in the emergency department and hospital, Am J Respir Crit Care Med, 199, Vol 160, pp 211-215
[21] Bipolar Disorder- Treatment http://www.nhs.uk/conditions/bipolar-disorder/pages/treatment.aspx

There are 2 main subtypes of BD – type I and type II. Type I BD is characterized by manic episodes while type II is defined by a pattern of depressive episodes. It is thusly assumed that type I BD patients start with a manic episode while type II patients start with a depressive episode.

Exhibit 39. Influence diagram for bipolar disorder



Prevalence: The initial prevalence of BD in 2007 can be obtained from the National Health Interview Survey (2007) [22] that asked "Have you EVER been told by a doctor or other health professional that you had bipolar disorder?" (Variable name: BIPDIS) [23]

Merikangas KR et al also provided lifetime and 12-month prevalence of the condition in their 2007 paper (Exhibit 40).[24] According to this study, 0.6/(0.6+0.8)=42.9% of the prevalence population have BP-I and the other 57.1% have BP-II.

[22] http://www.cdc.gov/nchs/nhis/quest_data_related_1997_forward.htm
[23] ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2007/English/qadult.pdf
[24] Merikangas KR et al. Lifetime and 12-month Prevalence of Bipolar Spectrum Disorder in the National Comorbidity Survey Replication_Arch Gen Psych_2007

Exhibit 40. Lifetime and 12 Month Prevalence and Age of Onset of Bipolar Disorder %

Table 1. Lifetime and 12-Month Prevalence and Age at Onset of *DSM-IV*/CIDI Bipolar Disorder in the 9282 Respondents

	Any BPD	BP-I	BP-II	Subthreshold BPD
Prevalence, mean (SD)				
Lifetime	4.4 (24.3)	1.0 (13.2)	1.1 (10.6)	2.4 (23.3)
12 mo	2.8 (18.9)	0.6 (9.2)	0.8 (9.9)	1.4 (15.1)
Age at onset, y*				
Mean (SE)	20.8 (11.8)	18.2 (11.6)	20.3 (9.7)	22.2 (12.6)
IQR†	12.6-24.9	12.3-21.2	12.1-24.0	13.0-28.3

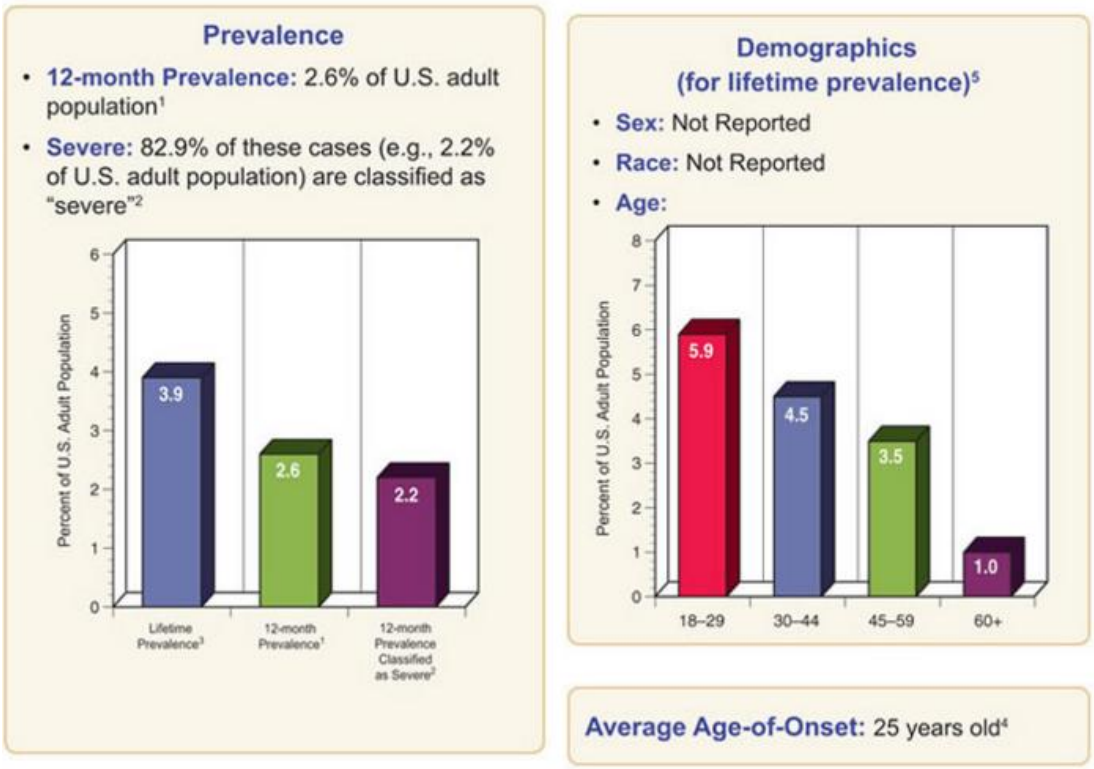
Abbreviations: BPD, bipolar disorder; BP-I, *DSM-IV* bipolar I disorder; BP-II, *DSM-IV* bipolar II disorder; CIDI, Composite International Diagnostic Interview; IQR, interquartile range.

*Retrospectively reported age at onset of the first manic/hypomanic or major depressive episode. The means differ significantly across the 3 BPD subgroups at the $P=.05$ level using a 2-sided test ($\chi^2_2=7.8$; $P=.02$).

†The range between the 25th and 75th percentiles on the age-at-onset distribution.

For data verification purpose, publically available prevalence data on a population level is scarce, with the below exhibit from the National Institute of Health among the most recently reported statistics for the USA, from 2005. Due to the ambiguity of “lifetime”, “lifetime prevalence” is not suitable for modeling use. But it provides a data point that the modelers can use to verify prevalence numbers produced by the model.

Exhibit 41. NIMH Statistics, 2005



Incidence: In a study by Kroon et al. a longitudinal electronic record of 800,000 patients in the Netherlands was analyzed for the primary outcome of interest, Bipolar I or II disorder defined according to DSM-IV criteria. Age and gender specific incidence rates (IRs) were calculated by dividing the total number of incident cases by the total number of person years at risk, per calendar year. [25] The analysis was done on a population older than 15 years, over time frame 1996 – 2007. Overall incidence of bipolar disorder was found to be 6.2 per 100,000 person years (95% CI: 5.7-8.3). For modeling purposes we will be using the overall incidence of the condition (Exhibit 42).

Exhibit 42. Overall Annual Incidence Rates for Bipolar Disorder

Bipolar Disorder	Incidence Rate
BP-I	4.3/100,000 PY (95% CI: 3.4–5.5), **69**% of incidence with BP-I
BP-II	1.9/100,000 PY (95% CI: 1.3–2.7), **31%** of incidence with BP-II

[25] Kroon JS et al. Incidence rates and risk factors of bipolar disorder in the general population: a population-based cohort study. Bipolar Disorders_2013

Exhibit 43. Incidence Rates by Age Group and Gender (final model inputs)

Age groups	Female IR/ 100,000 person years			Male IR/ 100,000 person years		
	Overall	BP-I*	BP-II*	Overall	BP-I*	BP-II*
15- 24	8.4	5.8	2.6	7.1	4.9	2.2
25 – 34	4.2	2.9	1.3	5.9	4.1	1.8
35 – 44	8.2	5.6	2.5	7.3	5.0	2.3
45 -54	13.0	9.0	4.0	10.6	7.3	3.3
55 – 64	5.7	3.9	1.8	9.6	6.6	3.0
65 – 74	5.1	3.5	1.6	1.4	1.0	0.4
75+	0.0	0.0	0.0	2.2	1.5	0.7

*Estimated from the % BP-I and % BP-II in Exhibit 42

Course of Disease: Soares et al. performed a systematic review that analyzed the clinical and cost effectiveness of pharmacological and/or psychosocial interventions, and in doing so reported the relapse rates for patients who had a previous manic or depressive episode, by treatment intervention.[26] The relapse rates for patients on placebo will be used to model the natural course of the disease. (Exhibit 44) As the data assumes the probability of a relapse based on whether the previous episode was depressive or manic, we will use BP-I and BP-II incidence rates to indicate what the previous episode is, as BP-I is characterized by more manic episodes, and BP-II is characterized by more depressive episodes.

Exhibit 44. Probability of Relapse after Depressive and Manic Episodes (untreated population)

Probability of relapse	Previous acute depressive episode	Previous acute manic episode
All	0.80 (0.62-1.0)	0.57 (0.46-0.69)
Depressive episode	0.62 (0.46-0.77)	0.18 (0.11-0.27)
Manic episode	0.18 (0.08-0.32)	0.38 (0.29-0.48)

Treatment effect: In their systematic review, Soares et al examined the effects of multiple therapies for preventing relapses in bipolar disorder.[27] Lithium has been the standard of care for bipolar disorder and thus will be used to represent treatment effect in the model.

[26] Soares-Weiser K, Bravo Vergel Y, Beynon S, Dunn G, Barbieri M, Duffy S, et al. A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder. Health Technol Assess 2007;11(39).

[27] Soares-Weiser K, Bravo Vergel Y, Beynon S, Dunn G, Barbieri M, Duffy S, et al. A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder. Health Technol Assess 2007;11(39).

Exhibit 45. Efficacy of Common Bipolar Disorder Medications

TABLE 80 *Results of the evidence synthesis: probability of relapse for patients with pretrial acute depressive episode (Analysis 1) and pretrial acute manic episode (Analysis 2)*[a]

	Analysis 1			Analysis 2		
	Posterior mean	**2.5% CrI**	**97.5% CrI**	**Posterior mean**	**2.5% CrI**	**97.5% CrI**
Type of relapse: all						
Lithium	0.46	0.37	0.56	0.27	0.22	0.32
Placebo	0.80	0.62	1.0	0.57	0.46	0.69
Divalproex/valproate	0.42	0.26	0.61	0.29	0.22	0.38
Imipramine	0.64	0.37	0.95	0.64	0.44	0.83
Lamotrigine	0.50	0.27	0.78	0.42	0.26	0.61
Olanzapine	0.58	0.40	0.75	0.23	0.16	0.31
Carbamazepine	0.84	0.51	1.0	0.66	0.30	1.0
Lithium + imipramine	0.43	0.24	0.68	0.37	0.21	0.57
Type of relapse: depression						
Lithium	0.38	0.29	0.47	0.07	0.05	0.10
Placebo	0.62	0.46	0.77	0.18	0.11	0.27
Divalproex/valproate	0.31	0.17	0.49	0.05	0.03	0.09
Imipramine	0.29	0.13	0.50	0.05	0.02	0.12
Lamotrigine	0.33	0.15	0.55	0.06	0.02	0.13
Olanzapine	0.55	0.37	0.72	0.14	0.08	0.21
Carbamazepine	0.64	0.38	0.92	0.23	0.07	0.62
Lithium + imipramine	0.28	0.12	0.49	0.05	0.02	0.11
Type of relapse: mania						
Lithium	0.08	0.04	0.13	0.20	0.15	0.24
Placebo	0.18	0.08	0.32	0.38	0.29	0.48
Divalproex/valproate	0.10	0.04	0.19	0.23	0.16	0.32
Imipramine	0.34	0.15	0.59	0.59	0.39	0.77
Lamotrigine	0.17	0.06	0.32	0.36	0.21	0.52
Olanzapine	0.03	0.01	0.06	0.08	0.05	0.12
Carbamazepine	0.24	0.05	0.57	0.43	0.17	0.76
Lithium + imipramine	0.14	0.05	0.30	0.31	0.16	0.51

[a] Marginal posterior distributions estimated on the log-odds scale, under the assumption that the relative treatment effect is additive to the (lithium) baseline.

Exhibit 46. Probability of Relapse after Depressive and Manic Episodes (treated population)

Probability of relapses	Previous acute depressive episode	Previous acute manic episode
All	0.46 (0.37-0.56)	0.27 (0.22-0.32)
Depressive episode	0.38 (0.29-0.47)	0.07 (0.05-0.10)
Manic episode	0.07 (0.04-0.13)	0.20 (0.15-0.24)

Mortality: Roshanaei-Moghaddam et al. report that patients diagnosed with bipolar spectrum experience increased premature mortality, with possible underlying causes including unhealthy lifestyle, biological factors, adverse pharmacologic effects and disparities in health care.[28] Only suicide risk is consdered in the model as death due to other clinical causes have already been captured by the modelign of other conditions.

Tondo et al.[29] reported suicide rates specifically related to bipolar disorder.

- Suicide rates in bipolar disorder patients average 0.4% per year, nearly 28 times higher than the international base rate of 0.0143% per year

From the Soares et al. paper, it was reported that lithium reduces suicides rates by 80%.[27] National statistics from 2005 indicate that 48.8% of those with bipolar disorder get treated, so the assumption is that this population will have their suicide risk (0.4% per year, according to Tondo et al.) reduced by 80%. For the remaining 51.2% who are untreated, the suicide rate (x%) has been calculated using the following equation:

$$x\%*51.2\% + x\%*(1-80\%)*48.8\% = 0.4\%$$

$$X= 0.66\%$$

In conclusion, annual mortality due to suicide for treated and untreated bipolar cases are 0.66% and 0.13%, respectively.

Cost:

Bipolar disorder is noted as the most expensive of the behavioral health illnesses.[30] However, according to 2005 statistics, less than half are receiving treatment. Cost for treated patients and overall patients can be extracted from literature. The cost of untreated cases can then be "backed out" following Exhibit 48 through the following equation:

Cost of treated case * treated % + cost of untreated case * untreated % = Cost of an average BP case

[28] Roshanaei-Moghaddam et al. Premature Mortality From General Medical Illnesses Among Persons With Bipolar Disorder: A Review

[29] Tondo, L., Suicidal behavior in dipolar disorder: risk and prevention, CNS Drugs, 2003; 17(7): 491-511

[30] Peele et al, Insurance expenditures on bipolar disorder clinical and parity implications, Am j Psychia, 2003

Exhibit 47. Percent of Bipolar Patients Receiving Treatment

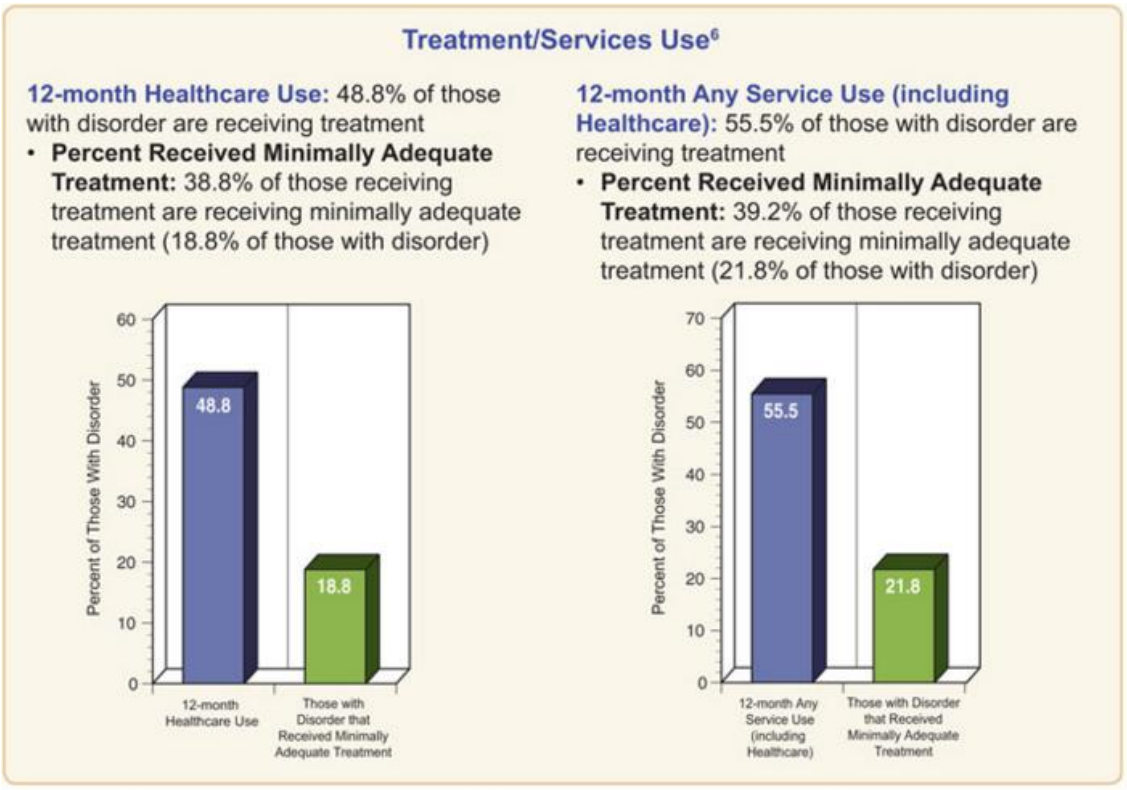
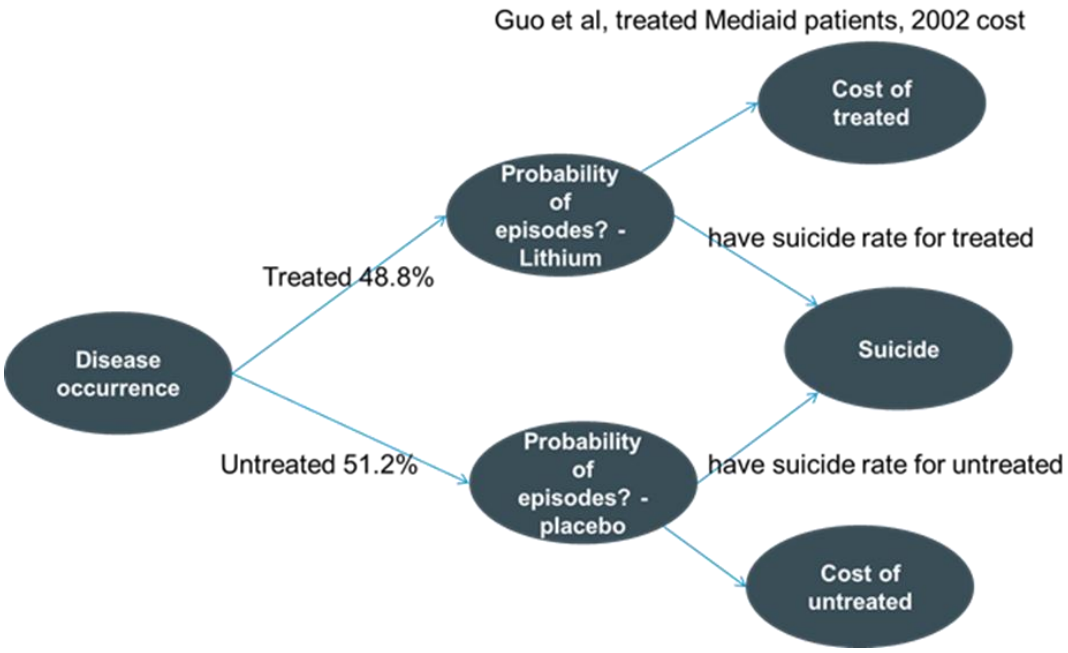


Exhibit 48. General Calculation Flow of Cost and Mortality Rates



Cost for treated patients: In the model, the cost per person receiving treatment for bipolar disorder is sourced from Guo et al.'s work, which analyzed the costs related to Medicaid patients with bipolar disorder.[31] A major caveat is that the data is dated, as it presents costs in 2002 dollars. However, the report breaks down the costs associated with the disorder according to different settings, and goes a step further and estimates that in patients with bipolar disorder and comorbid conditions, only 30% of the monthly mean costs incurred ($22,110) are directly associated with the disorder, which brings the mean cost per patient per year to $6,633 in 2002 dollars.

Alternatively, Brook et al.[32] report that the cost per year for an indivdual in an employee sponsored health plan is $9,983 (2001 cost). Again, this cost is for an individual who is receiving treatment.

The model cost input for treated patients will be the average from the 2 sources (inflated to 2015 dollars), which is $13,300.

Cost of an average patient: Surprisingly few studies have been done for the overall economic burden of bipolar disorder in more recent years. One study that is often referenced is Wyatt et al.'s work on the

[31] Guo J et al_Treatment costs related to bipolar disorder and comorbid conditions among medicaid patients with bipolar disorder_Pysch Serv_2007

[32] Brook RA et al. Incurring Greater Health Care Costs: Risk Stratification of Employees With Bipolar Disorder. Prim Care Companion J Clin Psychiatry. 2006; 8(1): 17–24.

economic burden of manic-depressive disorder, which reports the table below.[33] The same paper estimated that there were approximately 2.5 million individuals with bipolar disorder in the 1991, making the overall per person cost $18,084. Using data from official sources[34], the Medical CPI was used to calculate what the 1991 total value reported above would be in 2015 dollars, for an amount of $47,570.

Exhibit 49. Costs of Bipolar Disorder in a Medicaid Population

Table 2

Utilization and costs of treatment related to bipolar disorder and to comorbid conditions among 13,471 Medicaid patients with bipolar disorder[a]

Variable	Charges for treatment ($)	Reimbursed treatment ($)	Number[b]	Cost per unit	% of total cost	% related to bipolar disorder	% related to comorbid conditions
Inpatient care	128,294,262	66,645,686	140,998	473	34.9	11.7	23.2
Emergency room	26,712,789	11,944,764	180,297	66	6.3	.6	5.7
Outpatient	62,835,250	29,825,263	122,430	244	15.6	2.8	12.8
Mental health services	6,678,019	15,867,739	46,298	343	8.3	3.6	4.7
Physician visits	39,250,019	20,755,645	396,309	52	10.9	1.9	8.9
Laboratory tests	1,241,812	575,367	14	.3	.1	.2	
Other medical services	42,876,172	20,755,278	315,472	66	10.9	1.2	9.7
Prescriptions	34,490,891	24,416,370	568,093	43	12.8	8.1	4.7
Total	362,379,214	190,786,112			100.0	30.0	70.0
Mean cost per patient-year[c]	22,110	11,641					

[a] Costs in 2002 dollars. Percentages are based on reimbursed costs.
[b] Refers to number of hospital days, emergency room visits, discrete mental health services, physician visits, laboratory tests, discrete prescriptions, and other medical services.
[c] Mean reimbursed cost per patient year was calculated as [(total cost/total number of patients)/mean enrollment months] × 12.

[33] Wyatt RJ et al. An economic evaluation of manic-depressive disorder. Soc Psychiatry Psychiatr Epidemiol. 1995 Aug; 30(5): 213–219.
[34] US Bureau of Labor Statistics http://www.bls.gov/cpi/

Exhibit 50. Overall Cost of Bipolar Disorder

Costs of manic-depressive illness-rounded totals in millions[a]

Direct costs			
Treatment-related	Total inpatient costs	$2,350 million	
	Total outpatient costs	$300 million	
	Total nursing home, intermediate, domiciliary care costs	$2,980 million	
	Medication	$130 million	
	Substance abuse	$720 million	
	Shelters	$80 million	
Non-treatment-related	Total crime (includes jails/prisons)	$2,260 million	
	Suicide	$190 million	
	Research/Training	$50 million	
Subtracted from direct costs	Transfer costs	$1,300 million	
Total Direct Costs			$7,570 million
Indirect costs			
	Lost productivity homemakers	$3,150 million	
	Lost productivity institutions	$2,860 million	
	Lost productivity suicide	$7,840 million	
	Lost family productivity	$6,220 million	
	Lost compensation	$17,570 million	
Total indirect costs			$37,630 million
1991 Total (direct and indirect)			$45,210 million

[a]All figures, including totals, are rounded from the original figures

Cost of untreated patients: When calculating the cost for untreated bipolar disorder patients, the cost of untreated was "backed out" following the flow Exhibit 48.

$13,300*48.8% + cost for untreated * 51.2% = $47,570

Cost for untreated = $80,234

- Summary of disease cost: The cost for treated and untreated cases of BPD has been calculated in the steps above. Exhibit 50 provides the percentage breakdown of direct costs into inpatient, outpatient, and Rx costs, assuming untreated patients will NOT accumulate Rx cost.

Exhibit 51. Disease cost of Bipolar Disorder (2015 USD)

Type of patients	Inpatient cost	Outpatient	Rx	Total
Treated	$11,243	$1,435	$622	$13,300
Untreated	$71,151	$9.083	0	$80,234

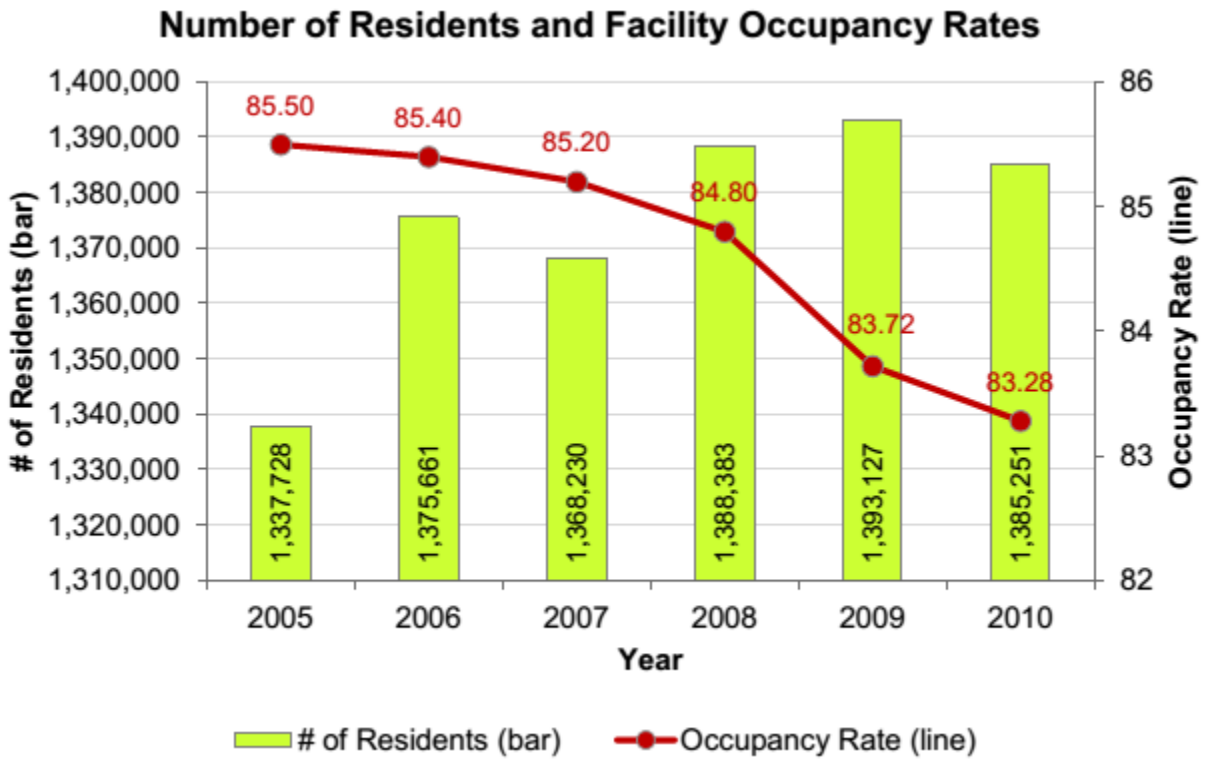
Long term care: Total number of nursing home residents in 2005 was 1.34 million. (Exhibit 52)[35] In another report, 5,299 (0.53%) out of 996 thousand newly admitted nursing home residents had bipolar disorder in 2005. [36] So the total number of bipolar disorder patients admitted to nursing home was 1.34million*0.53%= 7,100.

Total number of bipolar disorder patients was about 8.3 million (295.5 million population * 2.8% prevalence). This means 0.086% (7,100/8.3 million) of all existing bipolar patients are admitted to nursing home each year. The cost of nursing home is the same as for Alzheimer's disease, which is $61,436/year.

[35] Harrington, C., Nursing facilities, staffing, residents, and facility deficiencies, 2005 through 2010, Department of social & behavior sciences, University of California San Francisco, October 2011

[36] Fullerton, CA, Trends in mental health admissions to nursing homes, 1999-2005, Psychiatry services, Vol.60, No.7, July 2009

Exhibit 52. Number of Nursing Home Residents 2005-2010[35]



Missed work days: Hirschfeld reported an annual number of missed work days to be 49.5 per worker. [37] Assuming the number of missed work days is linearly correlated with the number of relapses, then the relative reduction of relapse rate is the same as reduction in missed work days. The number of missed work day for treated and untreated cases can then be "backed out" following the same approach as above (Exhibit 48)

Relative reduction due to treatment (lithium) is 0.52. (Exhibit 45) Suppose x is the number of missed workdays for untreated patients, the following equation holds:

$$x*51.2\% + x*0.52*48.8\% = 49.5$$

$$x=64.6$$

Exhibit 53. Number of Missed Work Days per Year for Treated and Untreated Bipolar Disorder Patients

Type of patients	Proportion	Missed work days
Treated	48.8%	33.6
Untreated	51.2%	64.6
Overall	100%	49.5

[37] Hirschfeld R et al, Bipolar Disorder—Costs and Comorbidity, Am J Man Care_2005

Benchmarking: "The costs per person associated with bipolar disorder have been estimated to be more than twice that of unipolar depression, making it one of the most expensive behavioral healthcare challenges."[38]

Key assumptions

- We know the prevalence of BP-I and BP-II. If BP-I, we assume the patient will start with a manic episode. If BP-II, we assume patient will start with depressive episode
- Natural course of relapse will depend on the previous episode (manic or depressive, dictated by whether individual has BP-I or BP-II)
- Our model will use Lithium as the default treatment, as it appears to be the standard of care, first line treatment for the condition
- Separating the condition out into BD I and BD II will be a future refinement
- Missed work days is linearly correlated with the number of relapses (manic or depressive)

Cancers

Incidence: Incidence of all cancers in the DPMM is modeled by following the same methodology, outlined in this section. For a microsimulation model with individual patient characteristics, the ideal method to model cancer incidence is through risk equations. However, our literature search yielded no usable published risk prediction models, as the risk prediction models that were identified were deemed unusable either due to a lack of input availability in our NHANES population (family history, genetic markers, etc.) or issues of external validity (Asian and Pacific Islander populations vs. U.S. populations).

Therefore, our approach to model the relationship between BMI and cancer risk was to reconcile estimates of cancer risk ratios by BMI, smoking status (where a link has been found in the literature), alcohol consumption (where a link has been found in the literature), and patient demographics (age, sex, and race/ethnicity), combining the following data sources:

(1) Published cancer risk ratios by BMI, smoking status, and alcohol consumption (sources for each cancer are detailed in the following sections)
(2) Estimates from NHANES of the proportion of the population by personal characteristics (where population is defined by age group, sex, and race/ethnicity); and
(3) Population cancer incidence probabilities by demographic.

CDC's SEER database provides incidence rates by 5-year age band, sex, and race for each cancer modeled. For each age, sex, and race combination we know the following relationship holds:

[38] http://www.sunovion.com/news/LAT307-13_Bipolar_Depression_Fact_Sheet.pdf

$$Incidence_{Overall} = Incidence_{Reference\ Group} \times \frac{Reference\ group\ population}{total\ population} + Incidence_{Reference\ Group} \times Relative\ Risk_{Group\ 1} \frac{Group\ 1\ population}{total\ population} + Incidence_{Reference\ Group} \times Relative\ Risk_{Group\ 2} \frac{Group\ 2\ population}{total\ population} + Incidence_{Reference\ Group} \times Relative\ Risk_{Group\ 3} \frac{Group\ 3\ population}{total\ population} \dots$$

The only unknown parameter in this equation is the incidence rate in the reference group, so we were able to solve for this rate giving us the reference group's incidence rate for each age, sex, and race combination.

For BMI, because we only have relative risks associated with BMI groups, the next step was to estimate a continuous relationship of relative risk vs. BMI. We used regression analysis to relate a person's estimated annual cancer risk (based on demographics and weight group) and BMI (controlling for demographics) to fit a non-linear curve relating cancer risk to BMI. This approach allowed us to estimate how cancer risk might change if a person loses body weight but remains within a body weight category.

For each cancer we ran three regressions with relative risk as the dependent variable. In each case BMI (from 18-40) was an independent variable and either log of BMI, BMI^2, or BMI^3 was another independent variable. The resulting equations from the regressions were then plotted against the actual RR plots over BMI. The equation with the highest adjusted R squared was chosen in all cases other than those at which the regression fit led to divergent trends at the extremes of the BMI spectrum. The one exception was for breast cancer, where Green et al., conducted a regression analysis aimed at estimating a functional relationship between BMI and relative risk of breast cancer incidence.[40] Our analysis used these equations to generate an individual's relative risk of breast cancer in lieu of the methodology outlined above.

No such analysis was needed for smoking or alcohol behavior. As these readily fall into discrete categories (i.e. current, former, or never smoker) the relative risk adjustments are able to be applied directly to the baseline risk.

Having calculated the baseline reference group's incidence for each condition, having estimated an equation to predict the relative risk for BMI, and having identified the smoking/alcohol related relative risks it was then possible to calculate a disease incidence probability for each person on the basis of their age, sex, race, BMI, alcohol consumption, and smoking status at time T. The following sections detail the relative risks associated with BMI, smoking, and alcohol groupings used in the model. Where a relationship between smoking and/or alcohol was included the source(s) outlining the link are detailed. A lack of link with alcohol consumption or smoking detailed below means that the literature search did not return sufficient evidence of a link for inclusion (not all cancers are related to smoking or drinking).

In the sections that follow, identification of cancers that have a link with smoking come from work by the International Agency for Cancer Working Group (IACWG).[39] Links (or lacks thereof) with alcohol come from a meta-analysis by Boffetta and Hashibe.[40] In the interest of space, these are not explicitly mentioned in the sections that follow.

Breast Cancer: The relative risks for the BMI risk adjustment for breast cancer (only in women) come from an analysis of published studies relating BMI to breast cancer risk by Green, et al.[41] While evidence of a link between alcohol and breast cancer incidence has been found, the evidence does not support such a link with smoking. The relative risks used to reflect differential breast cancer risk are summarized in Exhibit 54.

Exhibit 54. Relative Risk Adjustments in Breast Cancer

Breast Cancer		
BMI	**Pre-menopausal Women**	**Post-menopausal Women**
<22.5	0.96	0.85
22.5-24.9	1	1
25-27.4	0.93	1.1
27.5-29.9	0.99	1.21
30+	0.79	1.29
g/day of alcohol	**All Women**	
<0.1	1	
0.1-4.9	1.06	
5-9.9	1.15	
10-19.9	1.22	
20-29.9	1.2	
30+	1.51	

Note: the unit of drinking is “averaged across all days in a year”, **NOT** “averaged across only those days that alcohol is consumed”. (the same below) The latter definition was used by NHANES.

Cervical Cancer: The relative risks for the BMI risk adjustment for cervical cancer (only in women) come from an analysis of the UK Million Women Study, as detailed by Reeves, et al.[42] While evidence of a link between

[39] Vineis, P., M. Alavanja, P. Buffler, E. Fontham, S. Franceschi, Y. T. Gao, P. C. Gupta, A. Hackshaw, E. Matos, J. Samet, F. Sitas, J. Smith, L. Stayner, K. Straif, M. J. Thun, H. E. Wichmann, A. H. Wu, D. Zaridze, R. Peto, and R. Doll. "Tobacco and Cancer: Recent Epidemiological Evidence." *JNCI Journal of the National Cancer Institute* 96.2 (2004): 99-106. Web.
[40] Boffetta, Paolo, and Mia Hashibe. "Alcohol and Cancer." *The Lancet Oncology* 7.2 (2006): 149-56. Web.
[41] Green LE, Dinh TA, Smith RA. An estrogen model: the relationship between body mass index, menopausal status, estrogen replacement therapy, and breast cancer risk. Computational and mathematical methods in medicine 2012;2012.
[42] Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. Bmj 2007;335:1134.

smoking and cervical cancer incidence has been found, the evidence does not support such a link with alcohol consumption. The relative risks used to reflect differential cervical cancer risk are summarized in Exhibit 55.

Exhibit 55. Relative Risk Adjustments in Cervical Cancer

Cervical Cancer	
BMI	**RR**
<22.5	0.9
22.5-24.9	1
25-27.4	0.94
27.5-29.9	0.79
30+	1.02
Smoking Status	
Never Smoked	1
Former Smoker	1.5
Current Smoker	2.05

Colorectal Cancer: The relative risks for the BMI risk adjustment for colorectal cancer come from a meta-analysis of studies on the topic by Moghaddam, et al.[43] While evidence of a link between alcohol and colorectal cancer incidence has been found, the evidence does not support such a link with smoking. The relative risks used to reflect differential colorectal cancer risk are summarized in Exhibit 56.

Exhibit 56. Relative Risk Adjustment in Colorectal Cancer

Colorectal Cancer		
BMI	**Men**	**Women**
<25	1	1
25-30	1.16	1.03
30+	1.4	1.07
Drinking Status	**Both Sexes**	
Non-Drinker	1	
Light (≤12.5g/day)	1	
Moderate (12.6-49.9g/day)	1.21	
Heavy (50+g/day)	1.52	

[43] Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiology Biomarkers & Prevention 2007;16:2533-2547.

Endometrial Cancer: The relative risks for the BMI risk adjustment for endometrial cancer (only in women) come from an analysis of the Australian National Endometrial Study by Nagle, et al.[44] The literature did not support a link between drinking or smoking and endometrial cancer risk. The relative risks used to reflect differential endometrial cancer risk are summarized in Exhibit 57.

Exhibit 57. Relative Risk Adjustments in Endometrial Cancer

Endometrial Cancer	
BMI	**RR**
<25	1
25-29.9	1.47
30-34.9	2.66
35-39.9	4.39
>=40	7.98

Esophageal Cancer: The relative risks for the BMI risk adjustment for esophageal cancer come from an analysis of incident esophageal and gastric cancers in Connecticut, New Jersey, and Washington by Chow, et al.[45] The literature supported links between both smoking and drinking and risk of developing esophageal cancer. The relative risks used to reflect differential esophageal cancer risk are summarized in Exhibit 58.

[44] Nagle CM, Marquart L, Bain CJ et al. Impact of weight change and weight cycling on risk of different subtypes of endometrial cancer. European Journal of Cancer 2013;49:2717-2726.
[45] Chow WH, Blot WJ, Vaughan TL et al. Body mass index and risk of adenocarcinomas of the esophagus and gastric cardia. *Journal of the National Cancer Institute* 1998;90:150-155.

Exhibit 58. Relative Risk Adjustments in Esophageal Cancer

Esophageal Cancer		
	Usual BMI	RR
Men	<23.12	1.0
	23.12-25.08	1.5
	25.09-27.31	2.0
	>27.32	3.0
Women	<21.95	1.0
	21.95-24.12	0.8
	24.13-27.43	2.1
	>27.44	2.6
Smoking Status		
Never Smoker	1	
Former Smoker	2.03	
Current Smoker	2.5	
Drinking Status		
Non-Drinker	1	
Light (≤12.5g/day)	1.38	
Moderate (12.6-49.9g/day)	2.62	
Heavy (50+g/day)	5.54	

Gallbladder Cancer: The relative risks for the BMI risk adjustment for gallbladder cancer come from a meta-analysis by Larsson, et al.[46] The literature did not support a link between drinking or smoking and gallbladder cancer risk. The relative risks used to reflect differential gallbladder cancer risk are summarized in Exhibit 59.

Exhibit 59. Relative Risk Adjustments in Gallbladder Cancer

Gallbladder Cancer		
BMI	Men	Women
<30	1	1
30+	1.35	1.88

[46] Larsson SC, Wolk A. Obesity and the risk of gallbladder cancer: a meta-analysis. *Br J Cancer* 2007;96:1457-1461.

Kidney Cancer: The relative risks for the BMI risk adjustment for kidney cancer come from an analysis of the NIH-AARP Diet and Health Study, as detailed by Adams, et al.[47] While evidence of a link between smoking and kidney cancer incidence has been found, the evidence does not support such a link with alcohol consumption. The relative risks used to reflect differential kidney cancer risk are summarized in Exhibit 60.

Exhibit 60. Relative Risk Adjustments in Kidney Cancer

Kidney Cancer		
BMI	**Men**	**Women**
<18.5	1.37	1.7
18.5-<22.5	1	1
22.5-<25	1.15	1.11
25-<27.5	1.43	1.57
27.5-<30	1.64	1.6
30-<35	1.87	2.16
35+	2.47	2.59
Smoking Status		
Never Smoker	1	
Former Smoker	1.19	
Current Smoker	1.67	

Leukemia: The relative risks for the BMI risk adjustment for leukemia come from a meta-analysis conducted by Larsson, et al.[48] While evidence of a link between smoking and leukemia incidence has been found, the evidence does not support such a link with alcohol consumption. The relative risks used to reflect differential leukemia risk are summarized in Exhibit 61.

[47] Adams, K. F., M. F. Leitzmann, D. Albanes, V. Kipnis, S. C. Moore, A. Schatzkin, and W.-H. Chow. "Body Size and Renal Cell Cancer Incidence in a Large US Cohort Study." *American Journal of Epidemiology* (2008): n. pag. Web.
[48] Larsson SC, Wolk A. Overweight and obesity and incidence of leukemia: A meta–analysis of cohort studies. *International journal of cancer* 2008;122:1418-1421.

Exhibit 61. Relative Risk Adjustments in Leukemia

Leukemia	
BMI	**RR**
<25	1
25-<30	1.14
30+	1.39
Smoking Status	
Never Smoker	1
Former Smoker	1.33
Current Smoker	1.46

Liver Cancer: The relative risks for the BMI risk adjustment for liver cancer come from a meta- analysis of incident liver cancer studies by Larsson, et al.[49] The literature supported links between both smoking and drinking and risk of developing liver cancer. The relative risks used to reflect differential liver cancer risk are summarized in Exhibit 62.

Exhibit 62. Relative Risk Adjustments in Liver Cancer

Liver Cancer		
BMI	**Both Genders**	
<25	1	
25-30	1.17	
30+	1.89	
Smoking Status	**Men**	**Women**
Never Smoker	1	1
Former Smoker	1.72	1.35
Current Smoker	5.37	1.7
Drinking Status	**Both Genders**	
Non-Drinker	1	
Moderate (<3 drinks/day)	0.91*	
Heavy (3+ drinks/day)	1.16	

*not statistically significant

[49] Larsson, S. C., and A. Wolk. "Overweight, Obesity and Risk of Liver Cancer: A Meta-analysis of Cohort Studies." *Br J Cancer British Journal of Cancer* (2007): n. pag. Web.

Lung Cancer: The literature did not support a link between BMI or alcohol consumption and the risk of incident lung cancer. Peto, et al. investigated the excess risk of lung cancer attributable to smoking in their case-control analysis of UK data.[50] This study informs the risk in the DPMM and the findings are summarized in Exhibit 63.

Exhibit 63. Relative Risk Adjustments in Lung Cancer

Lung Cancer		
No Association with BMI		
Smoking Status	**Men**	**Women**
Never Smoker	1	1
Former Smoker	16.9	14.3
Current Smoker	33.3	20

Multiple Myeloma: The relative risks for the BMI risk adjustment for multiple myeloma come from an analysis of the Nurses' Health Study and Health Professionals Follow-up Study by Birmann, et al.[51] The literature did not support a link between drinking or smoking and multiple myeloma risk. The relative risks used to reflect differential multiple myeloma risk are summarized in Exhibit 64.

Exhibit 64. Relative Risk Adjustments in Multiple Myeloma

Multiple Myeloma		
BMI	**Men**	**Women**
<22	1	1
22-<25	1.3	1.1
25-<30	1	1.6
30+	2.4	1.2

Non-Hodgkin's Lymphoma: The relative risks for the BMI risk adjustment for non-Hodgkin's Lymphoma come from an analysis of the NIH-AARP Diet and Health Study by Lim, et al.[52] The literature did not support a link between drinking or smoking and non-Hodgkin's Lymphoma risk. The relative risks used to reflect differential non-Hodgkin's Lymphoma risk are summarized in Exhibit 65.

[50] Peto, R. "Smoking, Smoking Cessation, and Lung Cancer in the UK since 1950: Combination of National Statistics with Two Case-control Studies." Bmj 321.7257 (2000): 323-29. Web.
[51] Birmann BM, Giovannucci E, Rosner B, Anderson KC, Colditz GA. Body mass index, physical activity, and risk of multiple myeloma. *Cancer Epidemiology Biomarkers & Prevention* 2007;16:1474-1478.
[52] Lim U, Morton LM, Subar AF et al. Alcohol, smoking, and body size in relation to incident Hodgkin's and non-Hodgkin's lymphoma risk. *American journal of epidemiology* 2007;166:697-708.

Exhibit 65. Relative Risk Adjustments in Non-Hodgkin's Lymphoma

Non-Hodgkins Lymphoma	
BMI	**RR**
18.5–24.9	1
25–29.9	1.05
30–34.9	1.07
>=35	1.29

Ovarian Cancer: The relative risks for the BMI risk adjustment for ovarian cancer (for women only) come from an analysis of the UK Million Women Study by Reeves, et al.[53] The literature did not support a link between drinking or smoking and ovarian cancer risk. The relative risks used to reflect differential ovarian cancer risk are summarized in Exhibit 66.

Exhibit 66. Relative Risk Adjustments in Ovarian Cancer

Ovarian Cancer	
BMI	**Women**
<22.5	0.98
22.5-24.9	1
25-27.4	0.99
27.5-29.9	1.13
30+	1.12

Pancreatic Cancer: The relative risks for the BMI risk adjustment for pancreatic cancer come from a pooled analysis of pancreatic cancer studies by Genkinger, et al.[54] While evidence of a link between smoking and pancreatic cancer incidence has been found, the evidence does not support such a link with alcohol consumption. The relative risks used to reflect differential pancreatic cancer risk are summarized in Exhibit 67.

[53] Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. Bmj 2007;335:1134.
[54] Genkinger JM, Spiegelman D, Anderson KE et al. A pooled analysis of 14 cohort studies of anthropometric factors and pancreatic cancer risk. *International journal of cancer* 2011;129:1708-1717.

Exhibit 67. Relative Risk Adjustment in Pancreatic Cancer

Pancreatic Cancer		
BMI	**Men**	**Women**
<21	1.19	1.15
21-22.9	1	1
23-24.9	1.07	1.08
25-29.9	1.09	1.29
≥30	1.5	1.46
Smoking Status	**Men**	**Women**
Never Smoker	1	1
Former Smoker	1.6	1.3
Current Smoker	4.4	2.8

Prostate Cancer: The relative risks for the BMI risk adjustment for prostate cancer (for men only) come from an analysis of the NIH-AARP Diet and Health Study by Wright, et al.[55] The literature did not support a link between drinking or smoking and prostate cancer risk. The relative risks used to reflect differential prostate cancer risk are summarized in Exhibit 68.

Exhibit 68. Relative Risk Adjustments in Prostate Cancer

Prostate Cancer	
BMI	**RR**
<25	1
25-<30	1
30-<35	0.97
35-<40	0.84
40+	0.65

[55] Wright, Margaret E., Shih-Chen Chang, Arthur Schatzkin, Demetrius Albanes, Victor Kipnis, Traci Mouw, Paul Hurwitz, Albert Hollenbeck, and Michael F. Leitzmann. "Prospective Study of Adiposity and Weight Change in Relation to Prostate Cancer Incidence and Mortality." Cancer 109.4 (2007): 675-84. Web.

Stomach Cancer: The relative risks for the BMI risk adjustment for stomach cancer come from a meta-analysis of stomach cancer studies by Chen, et al.[56] While evidence of a link between smoking and stomach cancer incidence has been found, the evidence does not support such a link with alcohol consumption. The relative risks used to reflect differential stomach cancer risk are summarized in Exhibit 69.

Exhibit 69. Relative Risk Adjustments in Stomach Cancer

Stomach Cancer		
BMI	**Men**	**Women**
<25	1	1
25-30	1.01	0.99
30+	1.12	1.04
Smoking Status	**Men**	**Women**
Never Smoker	1	1
Former Smoker	1.74	0.94*
Current Smoker	1.98	1.78

*not statistically significant

Thyroid Cancer: The relative risks for the BMI risk adjustment for thyroid cancer come from an analysis of the NIH-AARP Diet and Health Study by Leitzmann, et al.[57] The literature did not support a link between drinking or smoking and thyroid cancer risk. The relative risks used to reflect differential thyroid cancer risk are summarized in Exhibit 70.

Exhibit 70. Relative Risk Adjustments in Thyroid Cancer

Thyroid Cancer		
BMI	**Men**	**Women**
18.5 to 24.9	1	1
25 to 29.9	1.53	1.11
≥ 30	1.89	1.1

[56] Chen, Y., L. Liu, X. Wang, J. Wang, Z. Yan, J. Cheng, G. Gong, and G. Li. "Body Mass Index and Risk of Gastric Cancer: A Meta-analysis of a Population with More Than Ten Million from 24 Prospective Studies." Cancer Epidemiology Biomarkers & Prevention 22.8 (2013): 1395-408. Web.

[57] Leitzmann MF, Brenner A, Moore SC et al. Prospective study of body mass index, physical activity and thyroid cancer. International journal of cancer 2010;126:2947-2956.

Absenteeism: Absenteeism of cancers is modeled via a regression equation on MEPS missed work days data. Due to the relative small sample size of each specific cancers, all cancers are grouped under one variable "any cancer" in the regression equation.

Treatment Effect: Treatment effect(s) will be expressed as a relative risk adjustment to disease mortality.

Cost: The modeling of cancer costs has not changed and is detailed in the original technical report.

Mortality: The modeling of cancer mortality has not changed and is detailed in the Technical Appendix.

Key Assumptions

- Relative Risks from older studies or studies outside of the US appropriately represent US incidence patterns.
- The curve fitting exercise for the relative risk of BMI on cancer incidence appropriately estimates the effect of BMI on cancer risk.

Chronic Obstructive Pulmonary Disease

There's no cure for chronic obstructive pulmonary disease (COPD) and thus it will be modeled as a life-long condition in the DPMM. Disease progression can be slowed, however, via medical treatment. [58] Disease severity will be characterized by GOLD severity stage, a widely used criteria developed by Global Initiative for Chronic Obstructive Lung Disease.[59]

Exhibit 71. Definition of GOLD Severity Stage[59]

Table 3. Classification of Severity of Airflow Limitation in COPD (Based on Post-Bronchodilator FEV_1)

In patients with $FEV_1/FVC < 0.70$:		
GOLD 1:	Mild	$FEV_1 \geq 80\%$ predicted
GOLD 2:	Moderate	$50\% \leq FEV_1 < 80\%$ predicted
GOLD 3:	Severe	$30\% \leq FEV_1 < 50\%$ predicted
GOLD 4:	Very Severe	$FEV_1 < 30\%$ predicted

The model influence diagram is shown in Exhibit 72. The simulation will primarily focus on the deterioration of the lung function (i.e. $FEV_1\%$) and the associated increase in cost and mortality per GOLD severity stage. [62]

[58] National Institute of Health, How is COPD treated, http://www.nhlbi.nih.gov/health/health-topics/topics/copd/treatment, July 31, 2013, accessed Oct 29, 2015

[59] Global Initiative for Chronic Obstructive Lung Disease, Pocket Guide to COPD diagnosis, management, and prevention, http://www.goldcopd.org/uploads/users/files/GOLD_Pocket_2015_Feb18.pdf, 2015, accessed Oct 29, 2015

Exhibit 72. Influence Diagram for COPD

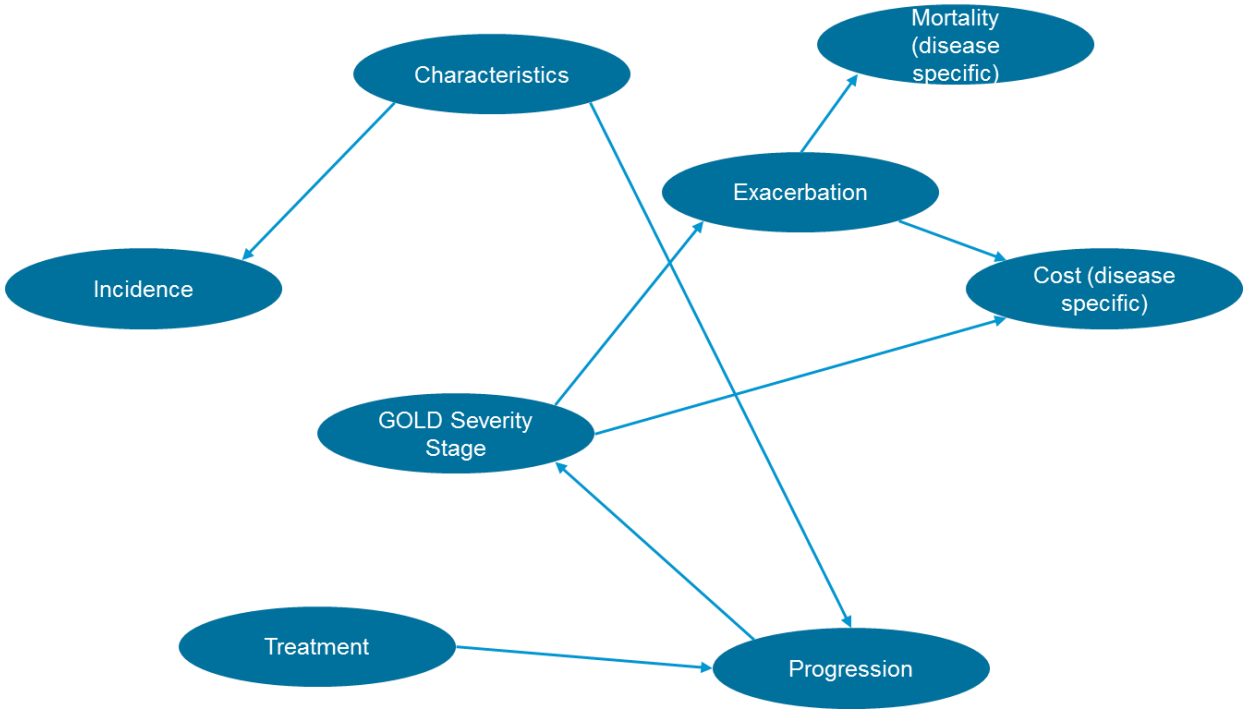
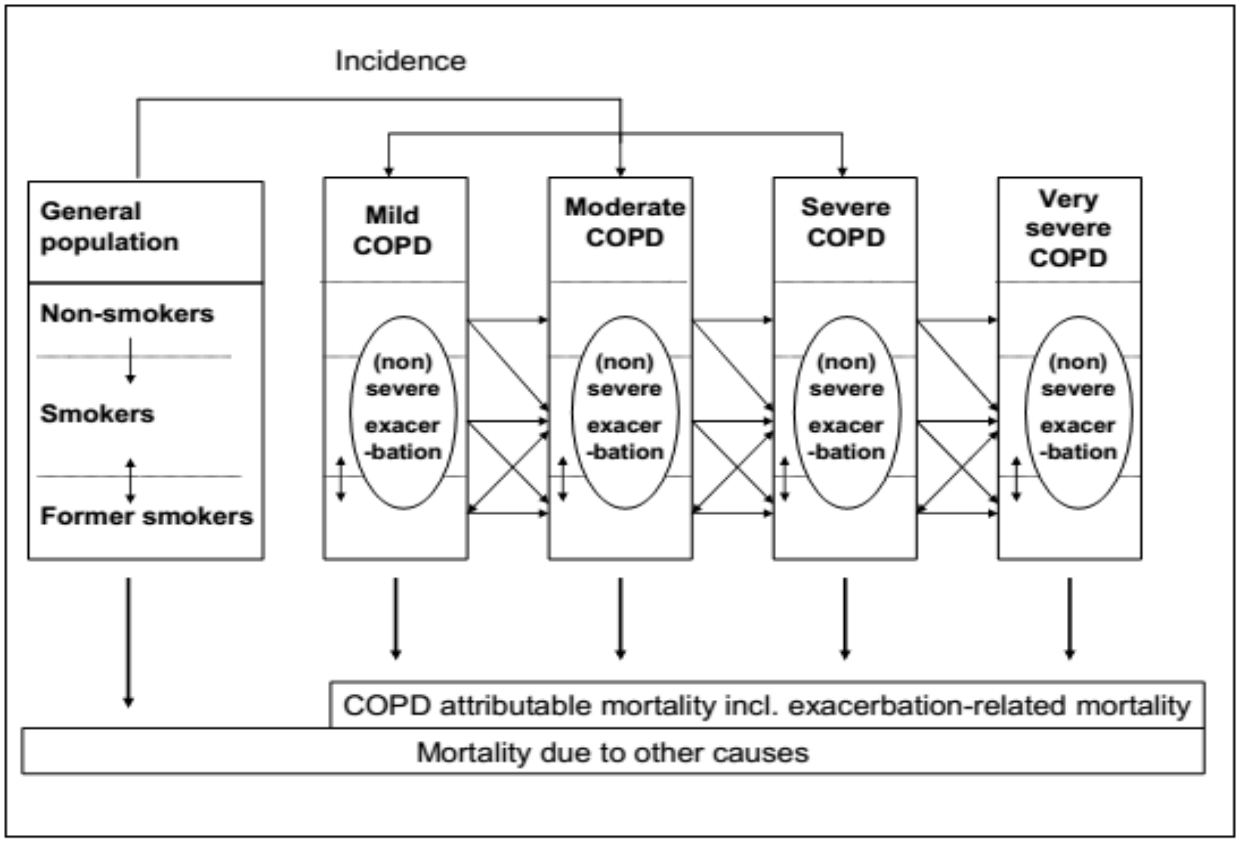


Exhibit 73. Schematic of a COPD Model



Prevalence: The question in BRFSS "Have you ever been told by a doctor or health professional that you have COPD, emphysema, or chronic bronchitis?" will be used to determine whether an individual has COPD. Both emphysema and chronic bronchitis are subtypes of the disease.

It was estimated that of all diagnosed COPD patients in 2000, 27% had mild, 55% moderate, 15% severe and 3% very severe COPD.[60] The average FEV_1% for the severity states can be found in Exhibit 78. Synthesizing both pieces of evidence the distribution of FEV_1% for the prevalence cohort is as follows:

Exhibit 74. FEV_1% Distribution in the COPD Prevalence Cohort

Percentage of prevalence cohort	Mean FEV_1%	Severity
27%	90	Mild
55%	65	Moderate

[60] Hoogendoorn M, et al. A dynamic population model of disease progression in COPD. 2005, Eur Respir J, 26(2): 223-233

15%	42	Severe
3%	23	Very severe

Incidence: US department of Health and Human Services reported relative risks of smokers and former smokers to get COPD as follows.[61]

Exhibit 75. Relative Risks of Smokers and Former Smokers to Get COPD

	Males			Females		
	Never smoker	Current smoker	Former smoker	Never smoker	Current smoker	Former smoker
0-4	1	1.9	1.7	1	1.7	1.5
5-9	1	1.9	1.7	1	1.7	1.5
10-14	1	1.9	1.7	1	1.7	1.5
15-19	1	1.9	1.7	1	1.7	1.5
20-24	1	1.9	1.7	1	1.7	1.5
25-29	1	1.8	1.6	1	1.6	1.4
30-34	1	2.0	1.8	1	1.8	1.6
35-39	1	3.2	2.7	1	2.6	2.3
40-44	1	5.1	4.1	1	4.0	3.4
45-49	1	7.5	5.9	1	5.6	4.8
50-54	1	9.8	7.9	1	7.7	6.4
55-59	1	11.0	9.6	1	9.9	7.7
60-64	1	11.7	10.3	1	11.3	8.3
65-69	1	12.2	9.5	1	11.2	8.2
70-74	1	12.5	8.5	1	10.0	7.4
75-79	1	12.2	7.7	1	8.1	6.5
80-84	1	11.0	7.0	1	6.1	5.4
85+	1	9.1	6.4	1	4.5	4.1

[61] U.S. Department of Health and Human Services. The health consequences of smoking: a report of the Surgeon General. 2004, U.S. Department of Health and Human Services, centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health promotion, Office on Smoking and Health, Atlanta

A Dutch source is used to inform the baseline *non-smoker* incidence in each age group.[62] The manuscript reported the total incidence in each age group and the pro-portion of smokers, non-smokers, and former smokers, which can be sued to back-calculate the baseline risk for non-smokers using the following equation:

> *Total incidence =% non-smoker * non-smoker risk + % former smoker * former smoker risk + % current smoker * current smoker risk*

Based on the formula above, the baseline non-smoker incidence is calculated and summarized in Exhibit 76.

Exhibit 76. Annual Incidence of COPD for Never-Smokers

Age	Males	Females
0-4	0	0
5-9	0	0
10-14	0	0
15-19	0	0
20-24	0	0
25-29	0	0
30-34	0	0
35-39	0.05%	0.06%
40-44	0.03%	0.04%
45-49	0.04%	0.05%
50-54	0.04%	0.08%
55-59	0.06%	0.08%
60-64	0.08%	0.10%
65-69	0.10%	0.13%
70-74	0.15%	0.15%
75-79	0.17%	0.15%
80-84	0.20%	0.19%
85+	0.22%	0.21%

Distribution of severity for the incidence cohort was reported to be 40% mild, 55% moderate, 4% severe and 0.1% very severe.[60]

Disease progression: The gradual decline in lung function is modeled via the following random effect model.[60] We can also calculate the corresponding GOLD severity stage per the definition in Exhibit 71.

Exhibit 77. Lung Function Decline Over Time

	β-Coefficient	SE
Intercept	-20.9546	18.4636
Year	0.2394	0.2473
Smoking cessation	14.3188	1.0216
Gender	7.3174	4.44
Age	1.1132	0.7312
Baseline FEV_1% predicted	1.3646	0.2282
Year*smoking cessation	0.4556	0.05597
Year*gender	-0.1562	0.03543
Year*age	-0.03144	0.00332
Year*baseline FEV_1% predicted	0.006027	0.001933
Smoking cessation*gender	1.7297	0.2029
Smoking cessation*baseline FEV_1% predicted	-0.1242	0.01092
Gender*age	-0.4038	0.1694
Gender*baseline FEV_1% predicted	0.02723	0.01347
Age*baseline FEV_1% predicted	-0.01818	0.009069
Age^2	-0.01213	0.007189
Age^2*smoking cessation	-0.00086	0.000143
Age^2*gender	0.004299	0.001674
Age^2*baseline FEV_1% predicted	0.000197	0.000089

Exacerbation: Exacerbation is defined by an increase in use of health care resources ("event based definition") and severe exacerbation is defined by a hospitalization. Hoogendoorn et al. reported the probability of both types of exacerbations by GOLD stage. (Exhibit 78)

Exhibit 78. Estimated Annual Exacerbation Frequency by GOLD Stage[62]

GOLD stage	Mean FEV_1% predicted at start	Total exacerbations: event-based definition	Severe exacerbations
I, Mild COPD	90	0.82 (0.46-1.49)	0.11 (0.02-0.56)
II, Moderate COPD	65	1.17 (0.93-1.50)	0.16 (0.07-0.33)
III, Severe COPD	42	1.61 (1.51-1.74)	0.22 (0.20-0.23)
IV, Very severe COPD	23	2.10 (1.51-2.94)	0.28 (0.14-0.63)

Since total exacerbation includes severe exacerbation, the incidence for moderate and severe exacerbations is listed in the following table:

Exhibit 79. Annual Incidence for Moderate and Severe Exacerbations

GOLD stage	Moderate exacerbation	Severe exacerbation
Mild	0.71 (0.44-0.93)	0.11 (0.02-0.56)
Moderate	1.01 (0.91-1.17)	0.16 (0.07-0.33)
Severe	1.39 (1.31-1.51)	0.22 (0.20-0.23)
Very severe	1.82 (1.37-2.31)	0.28 (0.14-0.63)

Treatment effect: As reflected in Exhibit 77, the most effective measure to slow the progression of the disease is smoking cessation. Hypothetical scenario can be modeled on a future medication that may slow the progression.

Mortality: It is assumed only severe exacerbation may lead to death. The case fatality rate due to severe exacerbation is, on average, 15.6%.[62] This rate corresponds to the average age of COPD population which is 69 years. Age is also a significant predictor of mortality.[63] The probability to die after a hospitalization increased with 4.1% per year increase in age (RR=1.041, 95% CI: 1.037-1.045) For each year below 69 years, the case fatality decreased by 4.1%, and vice versa.

Cost: There is a strong correlation between COPD severity and direct medical cost. Hilleman et al estimated the direct cost of COPD by severity as follows.[64]

[62] Hoogendoorn, M., et al., Comparing the cost-effectiveness of a wide range of COPD interventions using a stochastic, dynamic, population model for COPD, European Respiratory Journal, 2010
[63] Cazzola M, MacNee W, Martinez FJ, et al. Outcomes for COPD pharmacological trials: from lung function to biomarkers. Eur Respir J. 2008;31 (suppl 2):416-69.
[64] Hilleman, DE., Pharmacoeconomic Evaluation of COPD, Chest, Vol 118, No 5, 2000

Exhibit 80. Annual Medical Cost of COPD Reported in Hilleman et al. (Year 2000 USD)

Cost category	Stage I (FEV1 ≥50%)	Stage II (50%>FEV1 ≥35%)	Stage III (FEV1<35%)	Classification in the DPMM
Initial drug acquisition cost	$299	$529	$634	Rx
Add-on drug acquisition cost	$213	$191	$132	Rx
Oxygen therapy	0	$699	$2,012	Outpatient
Lab test	$345	$493	$610	Outpatient
Clinic visit	$82	$148	$171	Outpatient
Emergency department	$62	$319	$483	ED
Hospitalization	$680	$2,658	$6,770	Inpatient
Total cost	**$1,681**	**$5,037**	**$10,812**	

The disease staging system used in Hilleman is different from GOLD severity grade and thus needs to be adapted. Based on FEV_1%, stage III can be mapped to "Very severe COPD" per GOLD criteria. By the same token, stage II is "severe COPD" and stage I is "mild or moderate COPD".

Combining the information above, direct cost of COPD by GOLD severity is as follows (inflated to 2015 cost). This includes both maintenance cost and exacerbation cost.

Exhibit 81. Annual Cost of COPD – Including both Maintenance and Exacerbation Cost (2015 USD)

Cost category	Mild to moderate	Severe	Very Severe
Rx	$873	$1,228	$1,306
Outpatient	$728	$2,285	$4,762
Inpatient	$1,159	$4,532	$11,542
ED	$106	$544	$823
Total cost	**$2,866**	**$8,588**	**$18,434**

Based on the definition of exacerbation (increased resource use than normal) we assume a moderate exacerbation only involves ED visit while severe exacerbation involves hospitalization. Dalal et al reported the year 2008 cost of COPD-related ED visit to be $647 (SD: $445, n=24,617), simple admission $7,242 (SD: 7,987, n=42,734), and complex admission $20,757 (SD: $41,370, n=4,142)[65] If we use weighted average cost of admission as the cost of severe exacerbation, we can calculate the annual 2015 cost of exacerbation in Exhibit 82.

[65] Dalal, AA, et al., Costs of COPD exacerbations in the emergency department and inpatient setting, Respiratory Medicine, 2011:105, 454-460

Exhibit 82. Cost of COPD Exacerbation (2015 USD)

Cost category	Moderate exacerbation	Severe exacerbation
Rx	0	0
Outpatient	0	0
Inpatient	0	$10,303
ED	$790	$790
Total cost	$790	$11,093

The weighted annual average cost of exacerbation per year can be calculated based on the data given in Exhibit 79 and Exhibit 82. Maintenance cost of COPD is calculated by subtracting annual exacerbation cost from the total cost. (Exhibit 83)

Exhibit 83. Cost of COPD Maintenance (2015 USD)

Cost category	Mild	Moderate	Severe	Very Severe
Rx	$873	$873	$1,228	$1,306
Outpatient	$728	$728	$2,285	$4,762
Inpatient	$0	$0	$2,265	$8,657
ED	$0	$0	$0	$0
Total cost	$1,601	$1,601	$5,778	$14,725

Wacker et al. reported a large-scale absenteeism study in a German population in 2015.[66] The average number of sick days for GOLD severity 1-4 is 27.4, 26.5, 28.3, and 39.0, respectively. The number of sick days for GOLD grade 2 is less than in grade 1, which is likely due to statistical variation. We adjusted the number of sick days for grade as follows for better clinical validity.

Exhibit 84. Absenteeism Due to COPD by GOLD Stage

GOLD severity grade	Ave number of sick days per year
1	27.4
2	27.4
3	28.3
4	39.0

[66] Wacker, ME, et al.,

As for indirect costs, Gaurisco et al reported that there is 9% reduction in the likelihood of being employed along with a 4% increase in the probability of collecting social security disability insurance.[67]

Long-term care: it was reported in 2004 that 10% of all COPD patients are enrolled in a skilled nursing facility at any given time. [68] The chance of getting admitted to a LTC (Long term care) facility is associated with disease severity, with more severe disease associated with a higher likelihood of getting LTC. Recall in Exhibit 74 we have 3% of prevalence patients with "very severe" disease and 15% with "severe" disease. If we can allocate the 10% patients in LTC between "Very severe" and "severe" disease, we'll be able to calculate the rate of enrolling in LTC by severity. A reasonable assumption to make is that all 3% very severe patients are in LTC, and that the remaining 7% are from "severe" category. We then get the following:

Exhibit 85. % Patients in LTC Facility by Severity

Severity	% Patients in LTC Facility
Mild	0%
Moderate	0%
Severe	47% (=7%/15%)
Very severe	100%

Key assumptions:

- Some non-US but high quality data are used to inform disease progression and incidence
- Only severe exacerbation of COPD may lead to death
- Definition of exacerbation is based on the increased higher-than-normal resource use. Moderate exacerbation involves only an ED visit while severe exacerbation involves hospitalization
- All patients with "very severe" COPD are in some type of long-term care facility

Chronic Kidney Disease

CKD was defined as an estimated glomerular filtration rate less than 60 mL/min/1.73m^2. CKD risk was modeled using an algorithm developed by Kshiragar et al. using data from the Atherosclerosis Risk in Communities Study and the Cardiovascular Health Study.[17] Although the authors report a "Best-Fitting Categorical Model," we used the "Simplified Categorical Model" because not all of the risk inputs for the former model were in our simulation population dataset. The equation predicts 10-year risk for CKD, which we converted to annual risk by assuming equal probability of incidence in each of the 10 years. Diabetes is one of the variables in the risk function, and this equation was used to model CKD incidence for both the diabetes and non-diabetes populations. For the population over age 65, published annual incidence estimates were higher for the group as a whole than the maximum rate specified for the highest risk score group (defined in the article as ≥4% compared to 4.4% in the general population).[17] As such, the annual probability of incident CKD in this age group for the non-diabetes

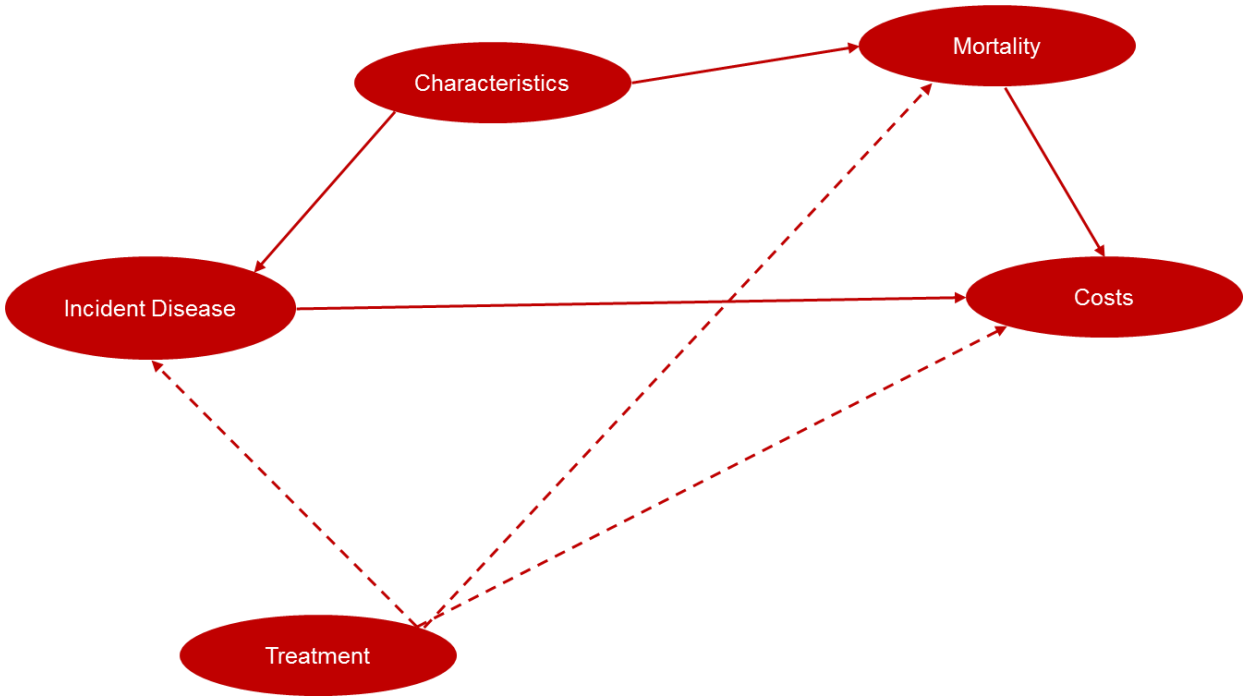
[67] Gaurisco, et al., The clinical and economic burden of chronic obstructive pulmonary disease in the USA, ClinicoEconomics and Outcomes Research, 2013:5 235-245
[68] Pleasants, RA, Chronic obstructive pulmonary disease in long term care, Annals of Long Term Care, Vol 17, Issue 3, 2009

population was defined as 4.4%, and for the elderly population with diabetes this probability was scaled with a relative risk adjustment of 2.1 based on the work of Hsu et al.[18]

Congestive Heart Failure

Congestive Heart Failure (CHF) is included in the DPMM as a chronic condition. As depicted in Exhibit 86, the modeling of CHF is based on disease occurrence and the resulting medical resource use and mortality. While disease severity classifications, such as the New York Heart Association's functional classification and the American Heart Association/American College of Cardiology classification exist, the available information in the DPMM population and published literature does not permit disease modeling based on disease classification.

Exhibit 86. Influence Diagram for CHF



Incidence: The equations used to model incidence of CHF are based on data from the Framingham Heart Study using a Cox proportional-hazards regression analysis.[69] Separate regressions were run for males and females. The predicted outcome measure is 10-year risk, which we converted to annual risk assuming equal risk across the 10 years. These equations are shown in Exhibit 87 below.

[69] D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM et al. General cardiovascular risk profile for use in primary care the Framingham Heart Study. Circulation 2008; 117(6):743-753.

Exhibit 87. Risk Equations for Incident CHF

Men* (10-year Baseline Survival: So(10) = 0.88936)				
Variable	**Beta****	**p-value**	**Hazard Ratio**	**95% CI**
Log of Age	3.06117	<.0001	21.35	(14.03, 32.48)
Log of Total Cholesterol	1.12370	<.0001	3.08	(2.05, 4.62)
Log of HDL Cholesterol	-0.93263	<.0001	0.40	(0.30, 0.52)
Log of SBP if not treated	1.93303	<.0001	6.91	(3.91, 12.20)
Log of SBP if treated	1.99881	<.0001	7.38	(4.22, 12.92)
Smoking	0.65451	<.0001	1.92	(1.65, 2.24)
Diabetes	0.57367	<.0001	1.78	(1.43, 2.20)

Women* (10-year Baseline Survival: So(10) = 0.95012)				
Variable	**Beta****	**p-value**	**Hazard Ratio**	**95% CI**
Log of Age	2.32888	<.0001	10.27	(5.65, 18.64)
Log of Total Cholesterol	1.20904	<.0001	3.35	(2.00, 5.62)
Log of HDL Cholesterol	-0.70833	<.0001	0.49	(0.351, 0.691)
Log of SBP if not treated	2.76157	<.0001	15.82	(7.86, 31.87)
Log of SBP if treated	2.82263	<.0001	16.82	(8.46, 33.46)
Smoking	0.52873	<.0001	1.70	(1.40, 2.06)
Diabetes	0.69154	<.0001	2.00	(1.49, 2.67)

Treatment effect: Treatment effect(s) will be expressed as a relative risk adjustment to mortality.

Mortality: Many studies have outlined the mortality risk from CHF. The DPMM models mortality risk attributable to CHF based on data from a Scottish registry using a Cox proportional hazards model.[70] Crude case-fatality rates are reported, and are used to approximate the baseline hazard in the model (for the referent group, aged <55) via a conversion function where the hazard rate (HR)is equal to the negative natural log of survival probability at time (t), divided by time (t). This is shown below.

$$HR = -\ln(Survival_t)\,/t$$

Following this approximation of the baseline hazard function (not reported in the work by MacIntyre, the proportional hazards reported for age group, sex, and comorbidities are used to calculate an individual's risk of CHF related mortality at a given time since diagnosis. The case fatality rates and regression results from MacIntyre's work can be found in Exhibit 88 below.

Exhibit 88. Case Fatality Rates and Cox Proportional Hazards Regression Coefficients

		Case-Fatality Rate, % (95% CI)				Median Survival, y
	n	30 d	1 y	5 y	10 y	
All patients	66 547	19.89 (19.89–19.89)	44.52 (44.52–44.52)	76.52 (76.52–76.52)	87.64 (87.64–87.64)	1.43 (1.40–1.46)
Age, y						
<55	3 765	10.41 (9.43–11.39)	24.2 (24.19–24.21)	46.75 (46.73–46.77)	57.98 (57.96–58.00)	7.25 (6.54–7.96)
55–64	8 664	13.39 (12.66–14.12)	32.43 (32.42–32.44)	62.97 (62.96–62.98)	78.15 (78.14–78.16)	3.07 (2.91–3.23)
65–74	18 635	18.23 (17.68–18.78)	41.61 (41.60–41.62)	73.64 (73.63–73.65)	86.49 (86.48–86.50)	1.71 (1.64–1.79)
75–84	24 267	22.18 (21.65–22.71)	48.89 (48.88–48.90)	82.19 (82.19–82.19)	92.47 (92.47–92.47)	1.06 (1.02–1.10)
>84	11 216	25.86 (25.06–26.66)	56.08 (56.07–56.09)	87.99 (87.98–88.00)	95.47 (95.46–95.48)	0.66 (0.62–0.71)

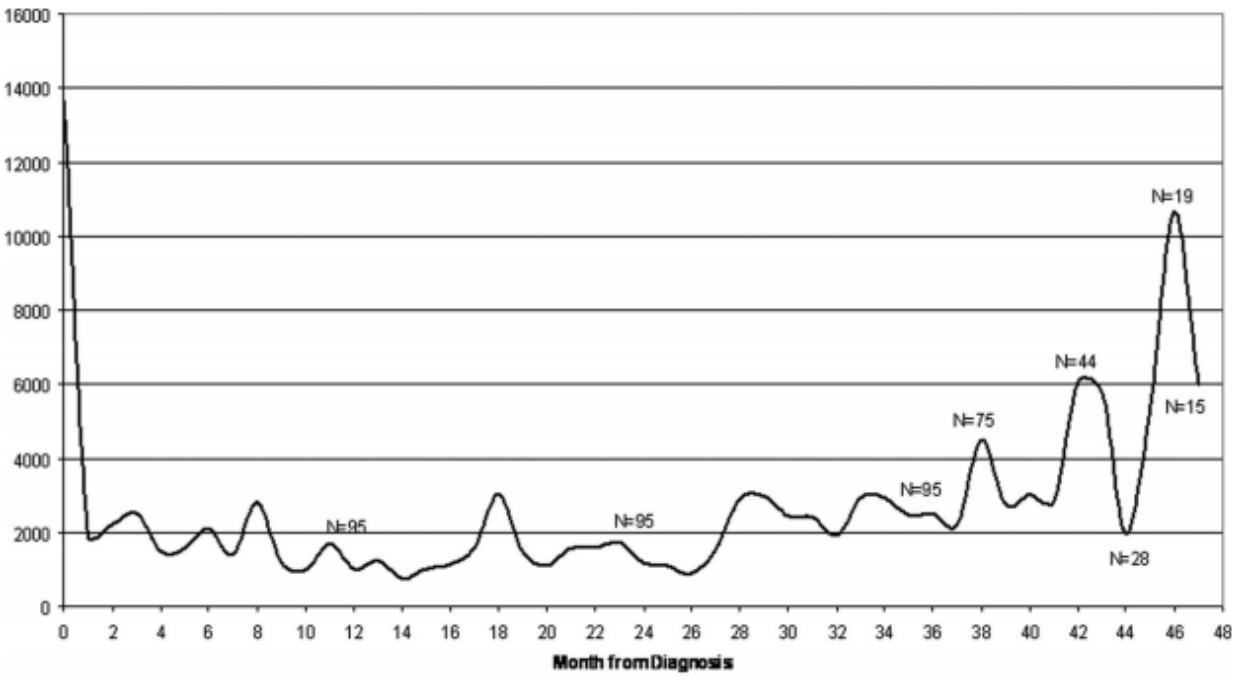
[70] MacIntyre K, Capewell S, Stewart S, Chalmers J, Boyd J, Finlayson A et al. Evidence of Improving Prognosis in Heart Failure : Trends in Case Fatality in 66 547 Patients Hospitalized Between 1986 and 1995. Circulation. 2000;102(10):1126-1131.

	Logistic Regression at 30 d (Odds Ratios)				Cox Proportional Hazards (30 d to End of Study Period)	
	n	Men	n	Women	Men	Women
Age, y						
<55	2 450	1	1 195	1	1	1
55–64	5 154	1.39 (1.19–1.62)	3 216	1.14 (0.94–1.40)	1.68 (1.55–1.81)	1.81 (1.62–2.03)
65–74	9 587	2.01 (1.74–2.32)	8 355	1.59 (1.33–1.91)	2.48 (2.30–2.66)	2.42 (2.17–2.69)
75–84	9 709	2.77 (2.40–3.20)	13 677	1.92 (1.61–2.29)	3.32 (3.09–3.57)	3.32 (2.99–3.69)
>84	3 011	3.72 (3.19)	7 843	2.35 (1.96–2.81)	4.36 (4.02–4.74)	4.35 (3.91–4.84)
Deprivation category						
1 (least deprived)	4 855	1.0	5 357	1.0	1.0	1.0
2	5 832	1.17 (1.06–1.29)	6 402	1.08 (0.99–1.18)	1.03 (0.98–1.08)	1.03 (0.98–1.08)
3	5 874	1.18 (1.07–1.30)	6 632	1.00 (0.91–1.10)	1.07 (1.01–1.12)	0.99 (0.94–1.04)
4	6 125	1.24 (1.12–1.37)	7 165	1.13 (1.03–1.23)	1.11 (1.05–1.16)	1.06 (1.02–1.11)
5 (most deprived)	7 225	1.26 (1.14–1.38)	8 730	1.11 (1.02–1.21)	1.10 (1.05–1.16)	1.06 (1.02–1.11)
Prior admission						
AMI	5 379	0.78 (0.72–0.85)	4 468	0.86 (0.79–0.94)	1.13 (1.09–1.18)	1.17 (1.12–1.22)
Arthritis	1 054	1.03 (0.89–1.20)	2 266	0.96 (0.86–1.07)	1.03 (0.95–1.12)	1.16 (1.10–1.23)
Atrial fibrillation	951	0.68 (0.56–0.82)	1 152	0.76 (0.64–0.89)	0.96 (0.87–1.05)	1.13 (1.04–1.22)
Cancer	1 715	1.44 (1.29–1.61)	1 546	1.47 (1.31–1.65)	1.44 (1.35–1.53)	1.36 (1.27–1.45)
Cerebrovascular disease	1 706	1.17 (1.04–1.32)	1 874	0.97 (0.87–1.09)	1.31 (1.23–1.40)	1.33 (1.26–1.41)
Coronary heart disease	4 073	0.92 (0.84–1.01)	3 212	0.86 (0.78–0.95)	1.04 (1.00–1.09)	1.01 (0.96–1.06)
Chronic renal failure	242	1.58 (1.17–2.12)	207	1.16 (0.82–1.63)	2.12 (1.80–2.50)	1.58 (1.32–1.88)
Diabetes	742	1.17 (0.98–1.41)	982	0.96 (0.81–1.13)	1.55 (1.41–1.70)	1.50 (1.38–1.62)
Hypertension	336	0.78 (0.57–1.06)	399	0.76 (0.58–1.01)	1.14 (0.99–1.31)	1.20 (1.06–1.36)
Peripheral vascular disease	1 343	1.36 (1.19–1.55)	906	1.19 (1.02–1.40)	1.42 (1.32–1.53)	1.48 (1.36–1.61)
Respiratory disease	2 417	1.26 (1.14–1.39)	2 289	1.06 (0.95–1.18)	1.38 (1.30–1.45)	1.37 (1.30–1.45)

Cost: The direct medical costs of CHF are modelled based on the lifetime costs observed in a cohort of patients in Olmstead, MN as detailed by Dunlay, et al.[71] Average annual costs are calculated based on the monthly cost breakdown shown in Exhibit 89. In the first year of diagnosis patients will accrue the costs for the average first twelve months. In all subsequent years, other than the year in which a patient dies from CHF, the patient will accrue the average costs of months 13-36. If the patient dies from CHF he or she accrues the annual cost from months 37-48.

[71] Dunlay, S. M., N. D. Shah, Q. Shi, B. Morlan, H. Vanhouten, K. Hall Long, and V. L. Roger. "Lifetime Costs of Medical Care After Heart Failure Diagnosis." Circulation: Cardiovascular Quality and Outcomes 4.1 (2010): 68-75. Web.

Exhibit 89. Average Monthly Cost of CHF



Absenteeism due to CHF, MI, or stroke is accounted for in one regression equation. The equation will be able to separately predict the number of missed workdays due to each of the 3 conditions.

Absenteeism will be modeled via a regression analysis on the MEPS. The dependent variable in the Poisson regression will be the number of missed workdays in a year the individual experienced and independent variables will include demographics (i.e. age, sex, race, etc.), socioeconomic characteristics (i.e. insurance status/type, annual income, etc.), biometrics (i.e. BMI, SBP, cholesterol ratio, etc.), and disease status (dummy variables for all modeled conditions). For each individual who has been simulated to be employed in a given cycle this equation will be used with his or her characteristics to predict the number of missed workdays in that year.

Treatment effects can be applied in this approach in two ways, both broadly working within a risk-reduction framework.

- If it is assumed that a treatment will decrease the number of missed workdays by a set amount, that amount can be applied to the total predicted count of missed workdays in the cycle.
- If, however, the assumption is that the treatment reduces missed workdays due to a certain condition by a percentage, the risk reduction can be applied directly to the coefficient in the estimation of the total number of missed workdays. For example, a treatment may reduce the total number of missed workdays of patients by 2 (modeled by subtracting 2 from the total predicted days missed for each

patient) or it may half the risk attributable to the disease (modeled by halving the exponentiated coefficient of that disease).

The relative reduction in absenteeism due to treatment will be synchronized with the relative reduction in MI events, to reflect the correlation between reduction in MI event and decrease in absenteeism.

Key assumptions:

- Scottish mortality rates are an appropriate approximation of US rates
- Case-fatality rates serve as a valid approximation of the baseline hazard function for a Cox proportional hazards estimation of mortality risk
- Time with disease is an appropriate surrogate for disease progression as expressed by mortality and costs

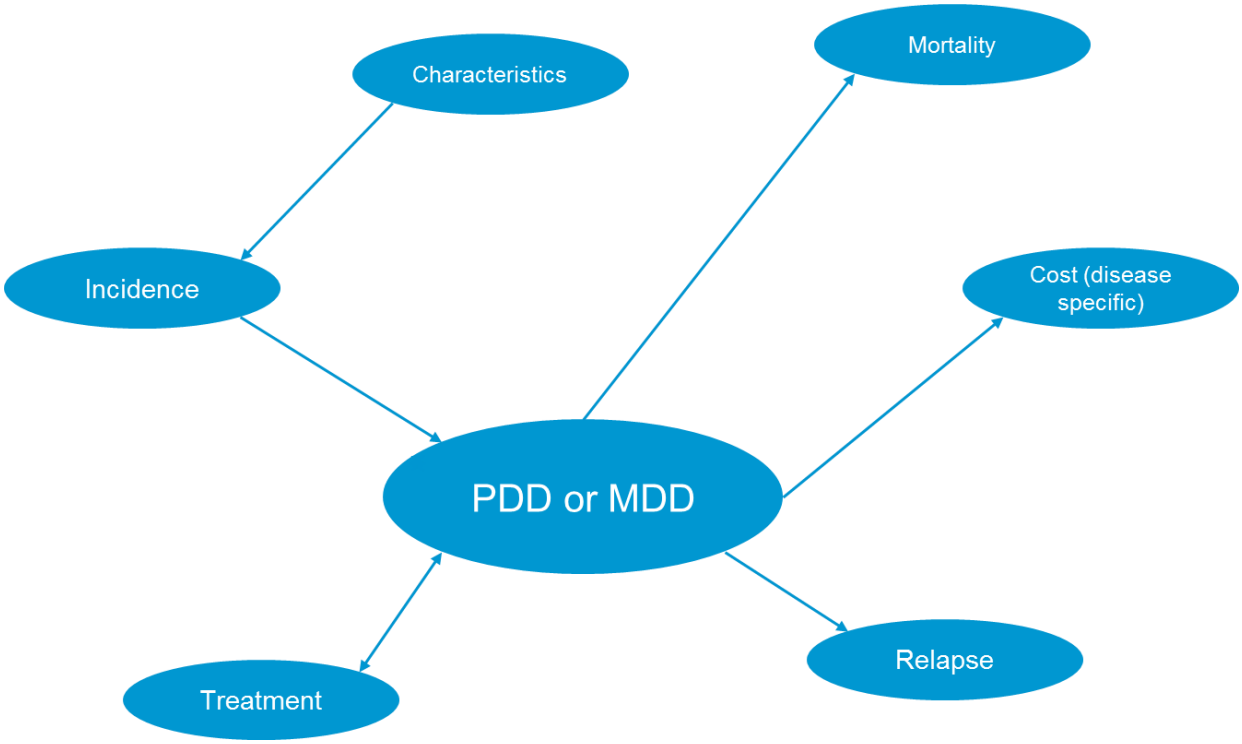
Depressive Disorder

In 2010, the United States spent $135 billion on mental health treatment, or about 5.6% of the national health care spending. Unlike overall health spending, the vast majority of behavioral health services are publicly funded. Medicaid, currently the largest source of financing for behavioral health services in the nation, covers a quarter of all expenditures.[72]

The modeling of depression includes major depressive disorder (MDD) which has symptoms lasting ≥2 weeks, and persistent depression disorder (PDD), which is characterized by depressive symptoms often lasting for ≥2 years without remission. The definition of PDD covers that of chronic major depressive disorder, dysthymia, and long-term depression.[73] MDD episode will be modeled as an event because the majority of MDD ends within a year. PDD is a life-long condition with much longer episodes and relapses.

[72] SAMSHA Spending estimates project, 2010.
[73] Coryell, W, Depressive disorders, http://www.merckmanuals.com/professional/psychiatric-disorders/mood-disorders/depressive-disorders, 2013, accessed Nov 23, 2015

Exhibit 90. Influence Diagram 0f Depression



Initial prevalence: The prevalence of depressive disorder can be determined via the patient health questionnaire (PHQ-8) dataset of BRFSS.[74] A PHQ-8 score of 0 to 9 is defined as no depression while a score of 10 to 24 points is defined as depression.[75] NIH reported the prevalence of PDD and MDD to be 1.5%[76] and 6.7%[77] among US adults, respectively. Consequently the proportion of the prevalent population that has PDD and MDD can be calculated to be 18.3% and 81.7%, respectively.

Incidence: The incidence of MDD among those without PDD or MDD was already part of the DPMM. The logic is detailed below.

[74] Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009; 114:163--73.
[75] http://www.cdc.gov/mentalhealthsurveillance/documents/11_gsps_user_guide_appendix_e.pdf, accessed November 30, 2015
[76] http://www.nimh.nih.gov/health/statistics/prevalence/dysthymic-disorder-among-adults.shtml, accessed November 30, 2015
[77] http://www.nimh.nih.gov/health/statistics/prevalence/major-depression-among-adults.shtml, accessed November 30, 2015

- Derive baseline annual risk of depression for male, not type-II/III obese, non-smoker and non-diabetic, from NHANES
- Extract the following odd ratios from published literature

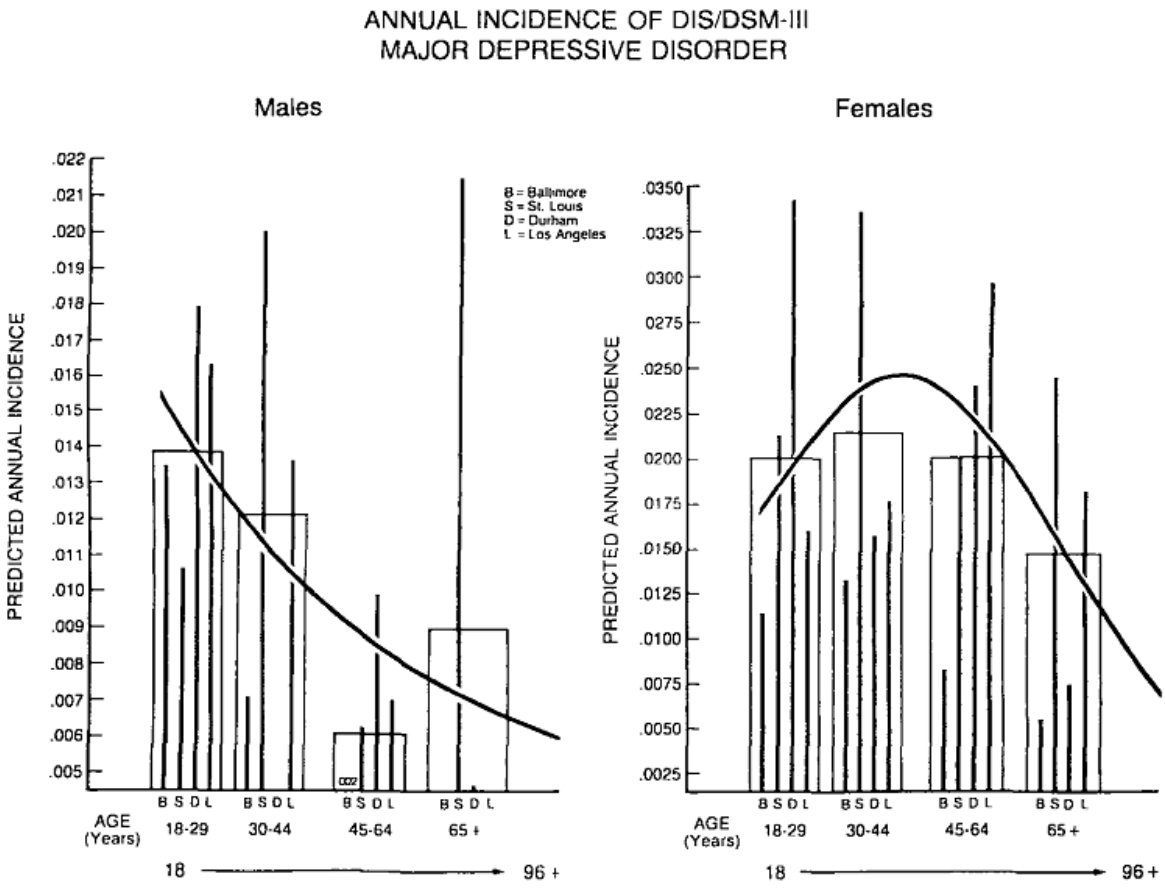
Condition	OR	CI	Source
Female	2.62	1.76-2.32	Onyike et al. Is obesity associated with major depression? American J Epid 2003
Type-II obese	1.9	0.79-4.6	Same as above
Type-III obese	4.63	2.06-10.42	Same as above
Smoking	2.24	1.32-3.81	Same as above
Diabetes	1.24	1.09-1.40	Nouwen et al., Type 2 diabetes mellitus as a risk factor for the onset of depression: a systematic review and meta-analysis, the European Depression in Diabetes (EDID) Research Consortium, 2010
Alcohol abuse	Not significant at 0.01	-	Onyike et al. Is obesity associated with major depression? American J Epid 2003

- Methodology of converting odds ratio to relative risk, using smoking population as an example
    1. Convert the baseline risk of depression to odds: baseline odds of depression = risk / (1- risk)
    2. Calculate the odds of depression among smokers: odds of depression (smoker) = baseline odds of depression * odds ratio of smoking (2.24)
    3. Convert the odds of depression among smokers to risk: risk of depression (smoker) = odds of depression (smoker)/(1+odds of depression(smoker))
- Calculate the risk of depression by multiplying baseline risk with risk ratios.

By def inition, incidence calculated this way includes both MDD and PDD. Rubio et al. reported the 12-month prevalence of PDD within the population with depressive disorder was 26.5%.[78]

The incidence rate will be applied to individuals of all ages. According to Eaton, et al.,[79] [79]major depressive disorder was observed in people from all age groups (Exhibit 91).

Exhibit 91. Incidence of Major Depressive Disorder in the Overall Population



Natural course of the disease: The majority of MDD episodes end within a year, and thus will be modeled as an event. Eaton et al. reported the median duration of MDD episodes to be 8-12 weeks. [80] Rubio et al. reported that the duration of longest MDD episode is 0.39 years.[78]

[78] Rubio, et al., Epidemiology of chronic and non-chronic major depressive disorder: results from the national epidemiologic survey on alcohol and related conditions, Depression and anxiety, 2011
[79] Eaton, WW, et al., The incidence of specific DIS/DSM-III mental disorders: data from the NIMH epidemiologic catchment area program, Acta Psychiatr Scand, 1989:79:163-178
[80] Eaton, WW, Natural history of Diagnostic Interview Schedule/DSM-IV major depression. The Baltimore Epidemiologic Catchment Area follow-up, Arch Gen Psychiatry, 1997, 54(11), 993-9

As mentioned above, PDD is a life-long condition with much longer episodes and relapses.

Klein et al. reported the Kaplan-Meier curve for time to recovery from a PDD episode and time to relapse after recovery, as follows: [8181]

Exhibit 92. Time to Recovery (Left) and Time to Relapse (Right) for PDD

FIGURE 1. Kaplan-Meier Survival Analysis of Time to Recovery From Dysthymic Disorder in 82 Patients Over a 10-Year Follow-Up Period

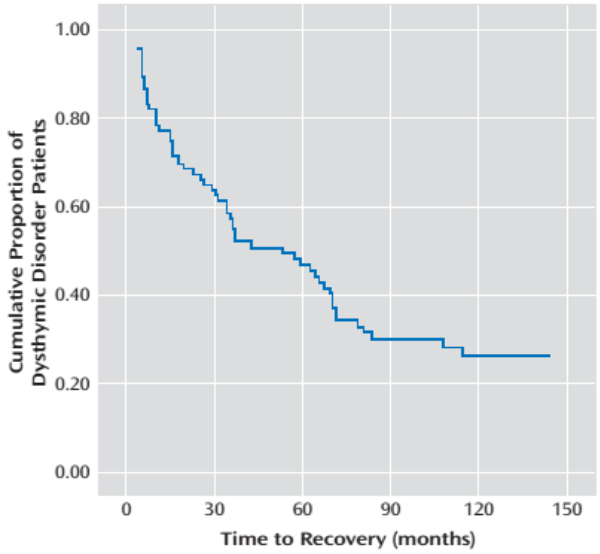
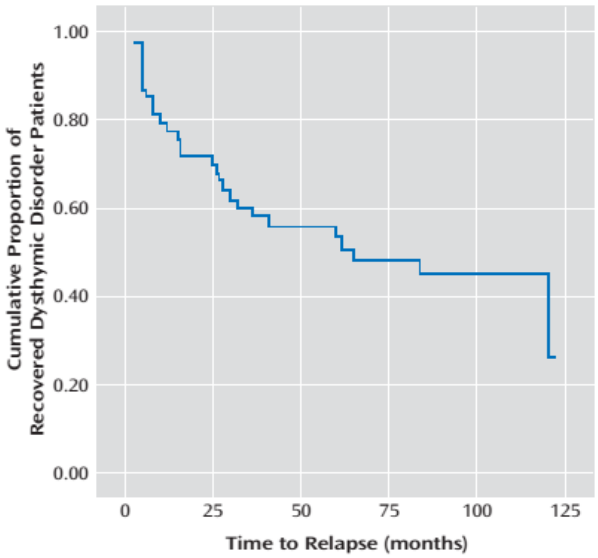


FIGURE 2. Kaplan-Meier Survival Analysis of Time to Relapse of Dysthymic Disorder in 53 Patients Who Recovered From Dysthymic Disorder Over a 10-Year Follow-Up Period



The annual probability of recovering from a PDD episode can thus be estimated to be 0.15 per year (average estimated probability using data from 30, 60, 90, and 120 months).

By the same token, the annual probability of relapsing after recovery is 0.12 (average estimated probability using data from 25, 50, 75, and 100 months). This is based on a naturalistic population that closely mimics the general population with depressive disorder, with some patients receiving medication while others don't.

long-term diagnostic 'stability' of either PDD or MDD is strong. In other words, once diagnosed, patients are far more likely to stay in PDD or MDD than to cross into the other type. Klein et al. found the odds of exhibiting a chronic depressive course were 14 times greater for patients with dysthymic disorder than for patients with nonchronic major depressive disorder

($p<0.001$), and the odds of having a nonchronic depressive course were 12 times greater for patients with nonchronic major depressive disorder than for patients with dysthymic disorder ($p<0.001$).[81]

[81] Klein DN, et al., Ten-year Prospective follow-up study of the naturalistic course of dysthymic disorder and double depression, AM J psychiatry, 2006; 163:872-880

Mortality: Depression likely lead to an unhealthy lifestyle which in turn may cause increased probability of death due to CVD and other physical illnesses. The higher incidence of physical illnesses among patients with depression is also incorporated implicitly in the modeling of those physical illnesses. The only cause of death directly associated with depression is suicide.

The evidence on suicide rate is scarce as it is neither feasible nor ethical to carry out double-blind studies on suicide reduction. For an insured and treated population, Simon et al. reported the suicide rate to be 118/100,000 person years for men (95% CI: 66-170), and 36/100,000 person years for women (95% CI: 18-54),.[82] The study didn't find any correlation between age and suicide rate in patients with depression. Even though an upward trend in suicide rate is observed among older patients, this was likely due to an higher suicide rate among older population overall (healthy or unhealthy).

In long-term follow up on untreated depression, 550 suicides per 100,000 person years were recorded.[83] This number is extrapolated into untreated suicides for men and women using the following method:

- Women are more likely to have depression than men. In a long-term follow-up study, [82] female patients recorded 44,242 person-years while male patients recorded 16,938 person-years. This means women accounts for 72% of depression person-years while men accounts for the remaining 28%.
- Assuming the relative risk of men committing suicide is the same between treated and untreated population, the relative risk is 118/36=3.28
- Suppose the suicide rate for untreated women is X/100,000 person years, the following equation holds:

$$X*72\% + X*3.28*28\% = 550$$

Solve the equation for X = 336

In summary, the suicide rate for treated and untreated patients with active MDD or PDD episodes are the following:

Exhibit 93. Suicide Rate for Treated and Untreated Active Depression Episodes

	Male (per 100,000 person years)	Female (per 100,000 person years)
Treated	118	36
Untreated	1,100	336

When patients are not in active PDD or MDD episodes, they have the same suicide rate as the general population.

[82] Simon, GE, Vonkorff, M, Suicide mortality among patients treated for depression in an insured population, Am J Epi, 1998, Vol. 147, No.2

[83] Coppen A, Lithium in unipolar depression and the prevention of suicide, The journal of clinical psychiatry, 2000:61 Suppl 9:52-56

Treatment effect: Since the rate of recovery and relapse is summarized from a naturalistic population, it already reflects current landscape for treatment effect and % treated. Treatment effect on suicide rate is detailed in the "mortality" section. According to a government website, 50-75% (average 62.5%) of patients with mental illnesses are untreated in the US.[84] The treated rate is thus set to be 37.5%. This constitutes the base case of the model.

Treatment effect for better pharmaceuticals developed in the future will be expressed as a relative term that increases the probability of recovery and reduces the risk of relapse following a recovery.

Cost: Greenberg et al. [85] researched the direct and indirect cost of MDD in the US, and reported cost burden in different settings.

The duration of MDD episode was reported to be 8-12 weeks (median duration) in a 1997 study,[80] 8 weeks (median duration, range 2- 520 weeks) in an adolescent population,[86] and 20 weeks (longest duration) in a 2011 study. [78] To be conservative we used the longest average duration of MDD episode (20 weeks) for calculations here.

Exhibit 94. Direct and Indirect Cost for MDD Episodes

Cost driver	Cost in 2015 USD (inflated from 2012 USD)
Rx	$11,832
Inpatient	$5,227
Outpatient	$10,820
ED	$173
Other	$1,620
Missed work days	12.3 (=31.9*20/52)

There is scarce cost data for PDD in the US, and thus we need to use MDD data as a proxy to estimate PDD cost assuming monthly cost for MDD and PDD are the same. Per month cost can be calculated using the duration of MDD, which is listed in the following table.

[84] State Government of Oklahoma, https://www.ok.gov/odmhsas/documents/suicide%20infographic.pdf, accessed Dec 4, 2015
[85] Greenberg, PE, et al., The economic burden of adults with major depressive disorder in the US (2005 and 2010), J Clin Psychiatry, 76:2 2015
[86] Lewinsohn, PM, et al., Major depression in community adolescents: age at onset, episode duration, and time to recurrence, J Am Acad Adolesc Psychiatry, 1994

Exhibit 95. Monthly Cost for PDD Episodes

Cost driver	Cost in 2015 USD (inflated from 2012 USD)
Rx	$2,528
Inpatient	$1,117
Outpatient	$2,312
ED	$37
Other	$346
Missed workdays	2.7 (=31.9/12)

As noted in earlier in the document, presenteeism costs will be assessed as roughly 3 times the cost of absenteeism following previous MDD literature estimates.

- Long term care: Total number of nursing home residents in 2005 was 1.34 million. (Exhibit 52)[35] In another report, 154 thousand (15.5%) out of 996 thousand newly admitted nursing home residents had depression in 2005.[36] So the total number of depression patients admitted to nursing home was 1.34million*15.5%= 207.7 thousand.

  Total number of depression patients was about 24.2 million (295.5 million population * 1.5%+6.7% prevalence). This means 0.86% of all existing depression patients are admitted to nursing home each year. The cost of nursing home is the same as for Alzheimer's disease, which is $61,436/year.

Key assumptions:

- Patients with PDD will remain on its course and not cross into the natural course of MDD, and vice versa
- Suicide rate is the same in episodes of PDD and MDD. When patients are not in active PDD or MDD episodes, they have the same suicide rate as the general population.
- The relative risk of men committing suicide is the same between treated and untreated population
- Monthly cost of MDD or PDD episodes is the same

Ischemic Heart Disease

The predictive equation used to model the annual probability of IHD incidence came from the Framingham Offspring study (n=4,780) with a 24 year follow-up.[32] This study used an accelerated failure time survival model based on a Weibull distribution. Due to differences in definition of IHD between the UKPDS Outcomes model and Framingham study, the Framingham equation was used to calculate risk of incident IHD for both the non-diabetes and the diabetes populations. The published equation contains an indicator for diabetes—thus accounting for the increased risk of IHD attributable to diabetes.

Left Ventricular Hypertrophy

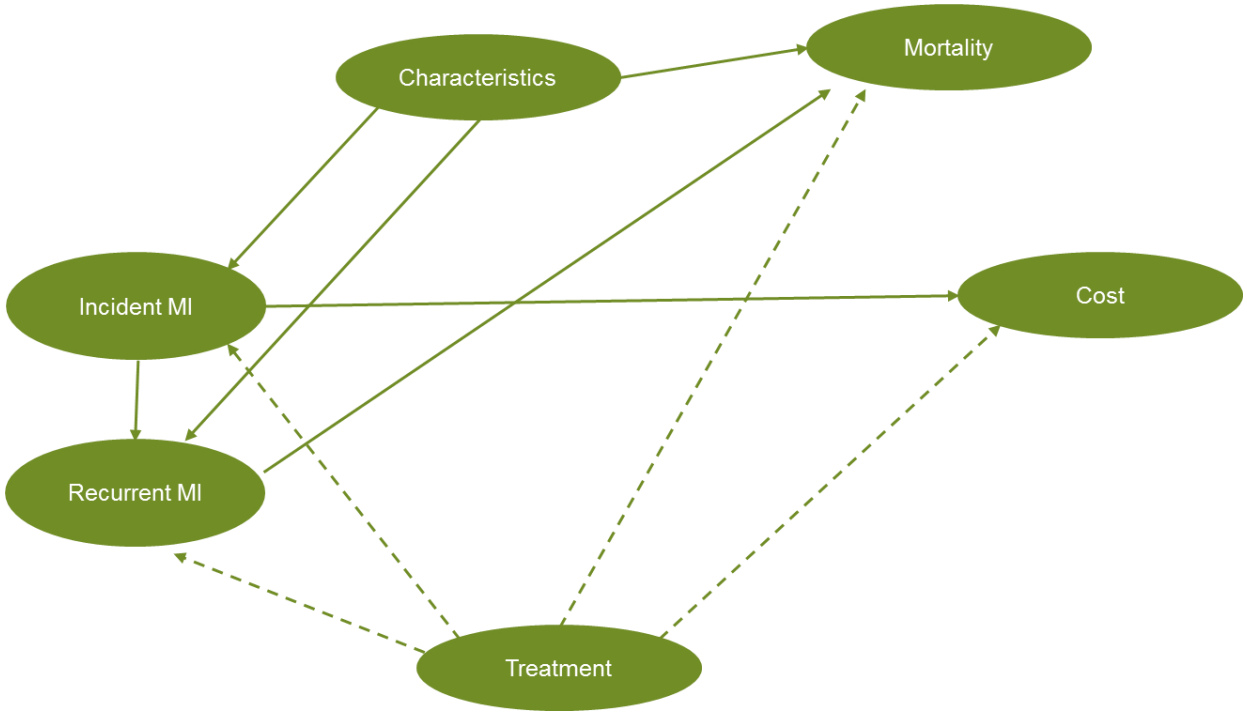
LVH was only modeled as a risk factor for other conditions. Prediction equations for LVH came from an analysis of an employed population in New York consisting of 639 patients.[27] The authors used logistic regression to

generate equations to estimate the probability of incidence of LVH. Different equations, with slightly different specifications, were generated for men and women. The male equation has SBP, DBP, age, and BMI as risk factors, while the female equation includes SBP, age, BMI, and race.

Myocardial Infarction (MI)

Myocardial infarction (MI) is included in the DPMM as an acute condition. As depicted in Exhibit 96, the modeling of MI is based on disease occurrence and the resulting medical resource use and mortality. Risk of subsequent (recurrent) MI is modeled separately from first MI. Excess mortality risk from subsequent MI's is modeled, though costs are assumed to be equivalent between first and recurrent MI's.

Exhibit 96. Influence Diagram for MI



First MI Incidence: Annual incidence of myocardial infarction among the population with diabetes comes from the UKPDS Outcomes Model and is based on a Weibull model.[87] For the non-diabetes population, the equation comes from published analysis of the Framingham Heart Study.[88] This study used a non-proportional hazards Weibull accelerated failure time model to predict the probability of event incidence. This equation used left

[87] Clarke PM, Gray AM, Briggs A, Farmer AJ, Fenn P, Stevens RJ et al. A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68). Diabetologia 2004; 47(10):1747-1759.

[88] Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. American Heart Journal 1991; 121(1):293-298.

ventricular hypertrophy (LVH), SBP, and cholesterol ratio as risk factors, in addition to age, sex, and smoking status. These parameters are displayed in Exhibit 97, below.

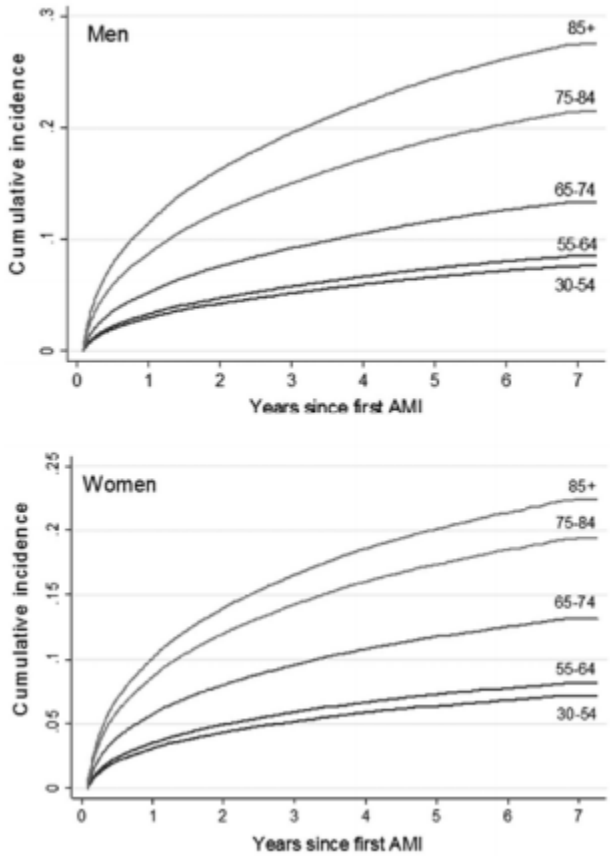
Exhibit 97. Risk Equation for First Incident MI

Diabetics		Non-Diabetics	
Functional form	Weibull	Functional form	Weibull
λ	−4.977	λ	3.064
ρ	1.26	ρ	-0.8584
AGE	0.055	FEMALE	10.5109
FEMALE	−0.826	LN(AGE)	0.7965
AC	−1.312	LN(AGE)*FEMALE	-5.4216
SMOK	0.346	(LN(AGE)^2)*FEMALE	0.7101
HBA1C	0.118	LN(SBP)	-0.6623
SBP	0.101	SMOKING	-0.2675
LN (TOTAL/HDL)	1.19	LN(TOTAL/HDL)	-0.4277
IHD	0.914	LVH*MALE	-0.1588
CHF	1.558		

Recurrent MI: MI recurrence is modelled based on English data recorded from 2004-2010.[89] Rates of subsequent MI by age and sex are displayed in Exhibit 98 below.

[89] Smolina, K., F. L. Wright, M. Rayner, and M. J. Goldacre. "Long-Term Survival and Recurrence After Acute Myocardial Infarction in England, 2004 to 2010." *Circulation: Cardiovascular Quality and Outcomes* 5.4 (2012): 532-40. Web.

Exhibit 98. Annual Risk of Recurrent MI



Treatment effect: Treatment effect(s) will be expressed as a relative risk adjustment to disease incidence and/or mortality.

Mortality: The modeling of MI mortality remains the same. Data on mortality within the first 365 days of an incident myocardial infarction came from the Swedish Socialstyrelsen Registry, with rates reported by sex for 5 year age bands.[90] This applies to both first and recurrent MI, and has been extensively detailed elsewhere.

Cost: The direct medical costs of MI come from regression analysis with the 2006-2010 Medical Expenditure Panel Survey (MEPS) Full Year Consolidated Data File and Medical Conditions File. We used a generalized linear model (GLM) with gamma distribution and log link to reflect the skewed distribution of annual medical expenditures. The dependent variable was total annual medical expenditures, while age, sex, insurance type, and modeled diseases were the independent variables. All medical cost estimates were converted to 2013 dollars using the medical component of the consumer price index.

[90] Socialstyrelsen. Swedish Health and Welfare Statistical Databases: AMI Statistics. Socialstyrelsen [serial online] 2013.

Absenteeism due to CHF, MI, or stroke is accounted for in one regression equation. The equation will be able to separately predict the number of missed workdays due to each of the 3 conditions.

Absenteeism will be modeled via a regression analysis on the MEPS. The dependent variable in the Poisson regression will be the number of missed workdays in a year the individual experienced and independent variables will include demographics (i.e. age, sex, race, etc.), socioeconomic characteristics (i.e. insurance status/type, annual income, etc.), biometrics (i.e. BMI, SBP, cholesterol ratio, etc.), and disease status (dummy variables for all modeled conditions). For each individual who has been simulated to be employed in a given cycle this equation will be used with his or her characteristics to predict the number of missed workdays in that year.

Treatment effects can be applied in this approach in two ways, both broadly working within a risk-reduction framework.

- If it is assumed that a treatment will decrease the number of missed workdays by a set amount, that amount can be applied to the total predicted count of missed workdays in the cycle.
- If, however, the assumption is that the treatment reduces missed workdays due to a certain condition by a percentage, the risk reduction can be applied directly to the coefficient in the estimation of the total number of missed workdays. For example, a treatment may reduce the total number of missed workdays of patients by 2 (modeled by subtracting 2 from the total predicted days missed for each patient) or it may half the risk attributable to the disease (modeled by halving the exponentiated coefficient of that disease).

The relative reduction in absenteeism due to treatment will be synchronized with the relative reduction in MI events, to reflect the correlation between reduction in MI event and decrease in absenteeism.

Key assumptions:

- English mortality rates and recurrent incidence rates are an appropriate approximation of US rates

Obstructive Sleep Apnea

Viner et al. reported the risk of obstructive sleep apnea (OSA), defined as apnea-hypopnea Index >10, among a group of patients suspected to have OSA.[105] Most of these patients were obese and experienced loud snoring. The risk equation is as follows:

Risk of OSA = exp(x)/ (1+exp(x))

Where x = -10.5132 + 0.9164*sex + 0.0470*age + 0.1869*BMI +1.932* loud snoring (sex=1 for male, 0 for female)

Prevalence of loud snoring in a general population is reported by Philips et al, in the following table. [106]

Exhibit 99. Prevalence of Loud Snoring in General Population

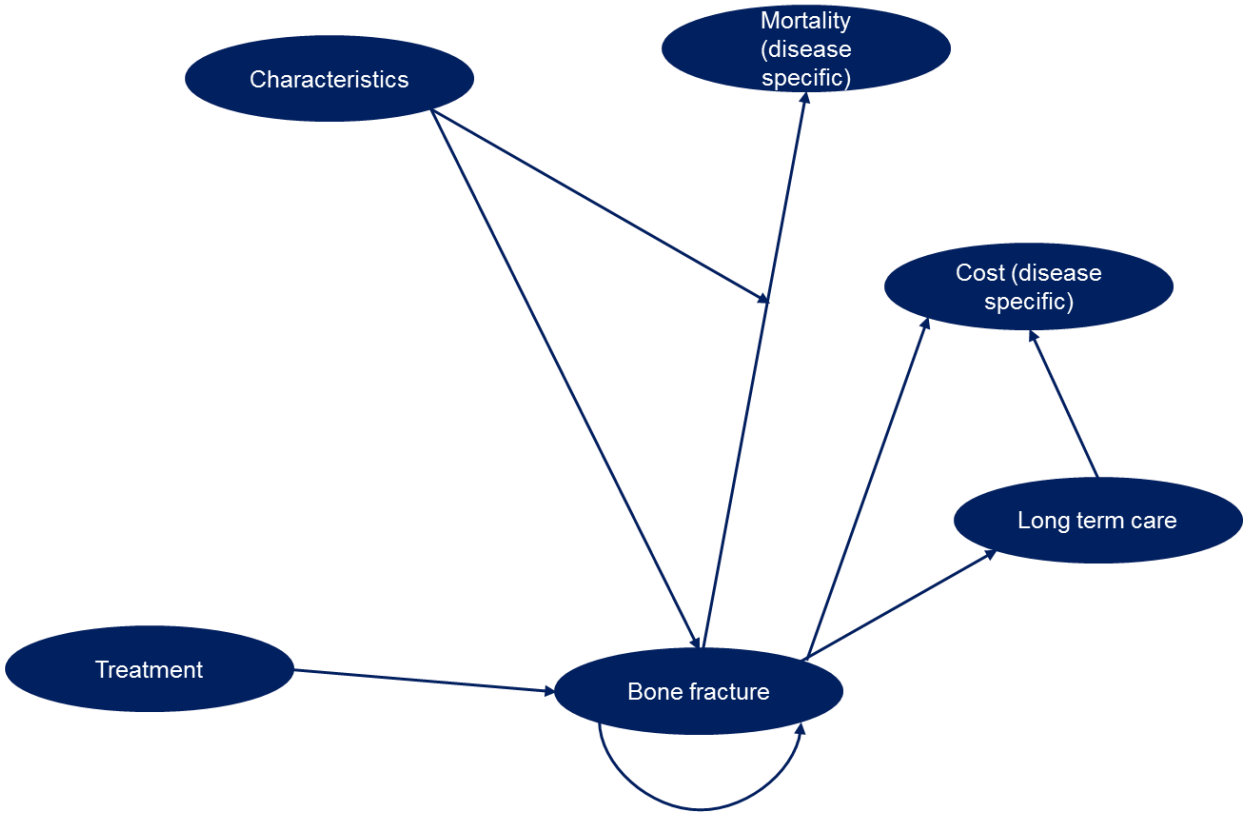
	Both Sexes		Women		Men	
Snoring						
Age	n	%	n	% All Women	n	% All Men
All	234	37	89	25	165	50
	N	%	n	% in Age Group	n	% in Age Group
≤50	144	34	39	9	105	25
>50	110	43	50	20	60	23

Because the equation only applies to people with suspected OSA, applying it to the general population in the model would overestimate the risk. Newman et al reported average 5-year incidence of OSA to be 11.9% in males and 4.9% in females. [107] The equation was then calibrated using a similar population to the one conveyed in this study.

Osteoporosis

As depicted in Exhibit 100, the modeling of osteoporosis is centered on the occurrence of bone fractures and the resulting medical resource use and mortality.

Exhibit 100. Influence Diagram for Osteoporosis



Initial prevalence of fracture history: The number of nursing home residents is 1.4 million (2013)[91], 10% of which had a history of hip fracture[92]. Given 10% of hip fracture could lead to institutionalized care, [93] the number of people with hip fracture history is about 1.4 million, leading to a prevalence of 0.4% with hip fracture history. Another source reported that out of 9 million bone fractures 1.6 million were at the hip.[98] This means the total number of people with bone fracture history in 2013 is 1.4* (9/1.6)=7.9 million.

The risk of bone fracture is significant only for those aged 50 and above.[137] in 2013, the number of people aged 50 and above is 33.65% * 316.5 million = 106.5 million.[94]

In summary, the prevalence of fracture history for people aged 49 and below is 0. For people aged 50 and above, the prevalence is 7.9 mil/106.5 mil=7.4%

[91] CDC, Nursing home care, http://www.cdc.gov/nchs/fastats/nursing-home-care.htm, April 26, 2015, accessed November 13, 2015
[92] Abbasi, AA, et al., Observations on nursing home residents with a history of hip fracture, Am J Med Sci, 1995
[93] Beringer TR, Clarke J, Elliott JR, Marsh DR, Heyburn G, Steele IC. Outcome following proximal femoral fracture in Northern Ireland. Ulster Med J 2006;75:200-6.
[94] US Census, US and World Population Clock, http://www.census.gov/popclock/,

Occurrence of bone fracture: 10-year probability of bone fracture will be projected by age, gender, race, BMI, and the number of clinical risk factors (CRFs) with FRAX® tool.[95] FRAX® is a widely used tool developed by WHO and validated on millions of patients worldwide. A sample risk table is shown below for illustration purpose. Complete risk tables are in the appendix.

Exhibit 101. 10-Year Probability (%) of a Major Osteoporotic Fracture for a 50 Year Asian Female

Age = 50 years

Number of CRFs	BMI (kg/m²)						
	15	20	25	30	35	40	45
0	2.2	2.0	1.9	1.7	1.5	1.3	1.1
1	3.5 (2.4-5.1)	3.2 (2.2-4.5)	3.0 (2.0-4.1)	2.6 (1.8-3.6)	2.3 (1.6-3.2)	2.0 (1.4-2.8)	1.8 (1.2-2.4)
2	5.6 (3.0-9.2)	5.0 (2.6-8.5)	4.7 (2.5-8.1)	4.1 (2.1-7.1)	3.6 (1.9-6.2)	3.1 (1.6-5.5)	2.7 (1.4-4.8)
3	8.8 (4.3-15)	7.7 (3.7-14)	7.1 (3.4-13)	6.2 (2.9-11)	5.4 (2.5-10)	4.7 (2.2-8.8)	4.1 (1.9-7.8)
4	14 (7.6-21)	12 (6.3-19)	11 (5.6-17)	9.2 (4.9-15)	8.1 (4.2-13)	7.1 (3.7-12)	6.2 (3.2-10)
5	21 (13-26)	17 (11-23)	16 (11-21)	14 (9.3-18)	12 (8.1-16)	10 (7.1-14)	9.1 (6.2-12)
6	31	25	22	20	17	15	13

All risk tables are presented in 5-year age bands. To be conservative, it is assumed that patients who are between 50 and 55 years old have the same bone fracture risk as the 50 years old, and those who have BMI between 20-25 has the same risk as BMI 20. The number of clinical risk factors is determined through the following table:[96]

[95] Kanis, JA, et al., FRAX and the assessment of fracture probability in men and women from the UK, Osteoporosis Int, 2008
[96] World Health Organization Collaborating Centre for Metabolic Bone Diseases, FRAX calculation tool (USA), http://www.shef.ac.uk/FRAX/tool.aspx?country=9, accessed November 13, 2015, University of Sheffield, UK

Exhibit 102. Clinical Risk Factors for Determining Bone Fracture Risk

CRF	Description	Modeling approach
Previous fracture	A previous fracture denotes more accurately a previous fracture in adult life occurring spontaneously, or a fracture arising from trauma which, in a healthy individual, would not have resulted in a fracture. Enter yes or no (see also notes on risk factors).	If the individual had a previous facture event, enter yes for this variable
Parent fractured hip	This enquires for a history of hip fracture in the patient's mother or father. Enter yes or no.	Lifetime probability of having a hip fracture was estimated to be 12.1% (CI: 12.1%-12.2%) for women and 4.6% (CI:4.5%-4.7%) for men in 2007.[97] This variable should be generated only once and remain unchanged for the individual for the remainder of the simulation
Current smoking	Enter yes or no depending on whether the patient currently smokes tobacco (see also notes on risk factors).	This will be determined via the modeling of smoking status
Glucocorticoids	Enter yes if the patient is currently exposed to oral glucocorticoids or has been exposed to oral glucocorticoids for more than 3 months at a dose of prednisolone of 5mg daily or more (or equivalent doses of other glucocorticoids) (see also notes on risk factors).	The prevalence of glucocorticoid uptake is not available. To be conservative, the value of this risk factor is always set to 0 “No”.
Rheumatoid arthritis	Enter yes where the patient has a confirmed diagnosis of rheumatoid arthritis. Otherwise enter no (see also notes on risk factors).	The overall prevalence of RA in adult American is about 0.6% estimated from data between 1995 and 2005. This variable should be generated only once and remain unchanged for the individual for the remainder of the simulation
Secondary osteoporosis	Enter yes if the patient has a disorder strongly associated with osteoporosis. These include type I (insulin dependent) diabetes,	If the individual has chronic liver disease (i.e. NAFLD), assign value “yes”.

[97] Hopkins, RB, et al., Estimation of the lifetime risk of hip fracture for women and men in Canada, Osteoporos Int, 2012

CRF	Description	Modeling approach
	osteogenesis imperfecta in adults, untreated long-standing hyperthyroidism, hypogonadism or premature menopause (<45 years), chronic malnutrition, or malabsorption and chronic liver disease	This will be an underestimation as the DPMM doesn't have other factors
Alcohol 3 or more units/day	Enter yes if the patient takes 3 or more units of alcohol daily. A unit of alcohol varies slightly in different countries from 8-10g of alcohol. This is equivalent to a standard glass of beer (285ml), a single measure of spirits (30ml), a medium-sized glass of wine (120ml), or 1 measure of an aperitif (60ml) (see also notes on risk factors).	This will be determined via the modeling of alcohol misuse

Once the 10-year probability is determined, it can be converted to annual probability using the following equation:

$$\text{Annual probability} = 1 - \exp(\text{LN}(1 - \text{10year risk})/10)$$

For instance, if 10-year probability is 31%, annual probability is 1-exp(LN(1-31%)/10)=3.64%

Location of bone fracture: FRAX tool can predict the probability of a general bone fracture or a hip fracture. General bone fracture could happen at 1) hip, 2) clinical spine, or 3) other locations (wrist/forearm or proximal humerus/upper arm fracture). The prediction of fracture location will follow a 2-step approach.

Step 1: Determine whether it's hip fracture. Generate one random number between 0-1 and compare against the annual probability of hip and general bone fracture. For example, if the probability of hip and bone fracture is 0.2 and 0.4 respectively, then one of the following 3 scenarios will happen:

1) Random number ≤0.2: the individual will have a hip fracture
2) 0.2< Random number ≤ 0.4: the individual will have a non-hip facture
3) Random number > 0.4: no fracture

Step 2: For those with non-hip fracture, determine whether it's clinical spine or other fracture: in 2000, among 9 million new osteoporotic fractures, 1.6 mil were at the hip, 1.4 at the spine.[98] The proportion of clinical spine fracture among non-hip fractures is thus 1.4/(9-1.6)= 19%. 81% of the non-hip fractures are at "other locations".

[98] Johnell O and Kanis JA, An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int, 2006, 17:1726.

Treatment effect: Treatment effect will be expressed as a relative risk in bone fracture probabilities.

Mortality: Many studies have reported the increased mortality due to hip or clinical spine fracture is very similar. [99,100,101,102] Based on the results of a meta-analysis, the probability of death is 3.14 (SD:4.03)times higher in the first year following the fracture, and 1.78 (SD:1.69) times higher in subsequent years.[103]

Cost: The direct medical cost of osteoporotic fracture consists of treatment cost for acute cases and long-term care cost for hip fracture patients.

Exhibit 103. Direct Medical Cost of Osteoporotic Fracture

	Direct medical cost (2001 inflated to 2015 USD)
Acute hip fracture	$26,268[104]
1st year long term care after hip fracture	$14,524[105]
Subsequent year long term care after hip fracture	$10,261[105]
Acute clinical spine fracture	$10,924[104]
Acute other fracture	$9,064[104]

Indirect cost due to absenteeism was estimated by Meerding et al. as follows. The cost is only applicable to people who are employed and below age 65.

Exhibit 104. Mean Work Days Lost Due to Osteoporotic Fracture

	Absenteeism (mean work days lost)
Hip fracture	90.7
Clinical spine fracture	38.8
Other fracture	24.7

[99]Johnell O, Kanis JA, Oden A, Sernbo I, Redlund-Johnell I, Petterson C, et al. Mortality after osteoporotic fractures. Osteoporos Int 2004;15:38-42.
[100] Kanis JA, Oden A, Johnell O, De Laet C, Jonsson B. Excess mortality after hospitalisation forvertebral fracture. Osteoporos Int 2004;15:108-12.
[101] Kanis JA, Oden A, Johnell O, De Laet C, Jonsson B, Oglesby AK. The components of excess mortality after hip fracture. Bone 2003;32:468-73.
[102] Cauley JA, Thompson DE, Ensrud KC, Scott JC, Black D. Risk of mortality following clinical fractures. Osteoporos Int 2000;11:556-61.
[103] Haentjens P, Magaziner J, Colon-Emeric CS, Vanderschueren D, Milisen K, Velkeniers B, et al. Meta-analysis: excess mortality after hip fracture among older women and men. Ann Intern Med 2010;152:380-90.
[104] Gabriel SE, Tosteson AN, Leibson CL, Crowson CS, Pond GR, Hammond CS, et al. Direct medical costs attributable to osteoporotic fractures. Osteoporos Int. 2002;13:323-30
[105] Leibson CL, Tosteson AN, Gabriel SE, Ransom JE, Melton LJ. Mortality, disability, and nursing home use for persons with and without hip fracture: a population-based study. J Am Geriatr Soc. 2002;50

Key assumptions:

- Risk of having an osteoporosis related bone fracture before age 50 is 0
- Conservative assumptions were made when calculating the risk of bone fracture. It is assumed that patients falling between age or BMI bands have the same risk as the lower bound of that band.
- No usage data is available on glucocorticoids. To be conservative, it is assumed this risk factor will not increase the risk of bone fracture.

Other Obesity Comorbidities

Other obesity comorbidities in the model included chronic back pain, gallstone disease, gastroesophageal reflux disease, non-alcoholic fatty liver disease, osteoarthritis, pneumonia, and pulmonary embolism.

Similar to the cancers included in the model, our literature search did not yield any usable risk prediction equations for these remaining conditions. As such, we used an approach similar to that used for modeling cancer incidence. Incidence rates and relative risks associated with BMI category were identified in the literature for each condition (the sources are detailed in Exhibit 10). While incidence rates were not always available at as detailed a demographic level as the SEER database for modeling cancers, the most granular data available was used for both incidence rates and relative risks. The process for calculating reference group incidence rate and relative risk function was identical to that used for the cancers for each population group.

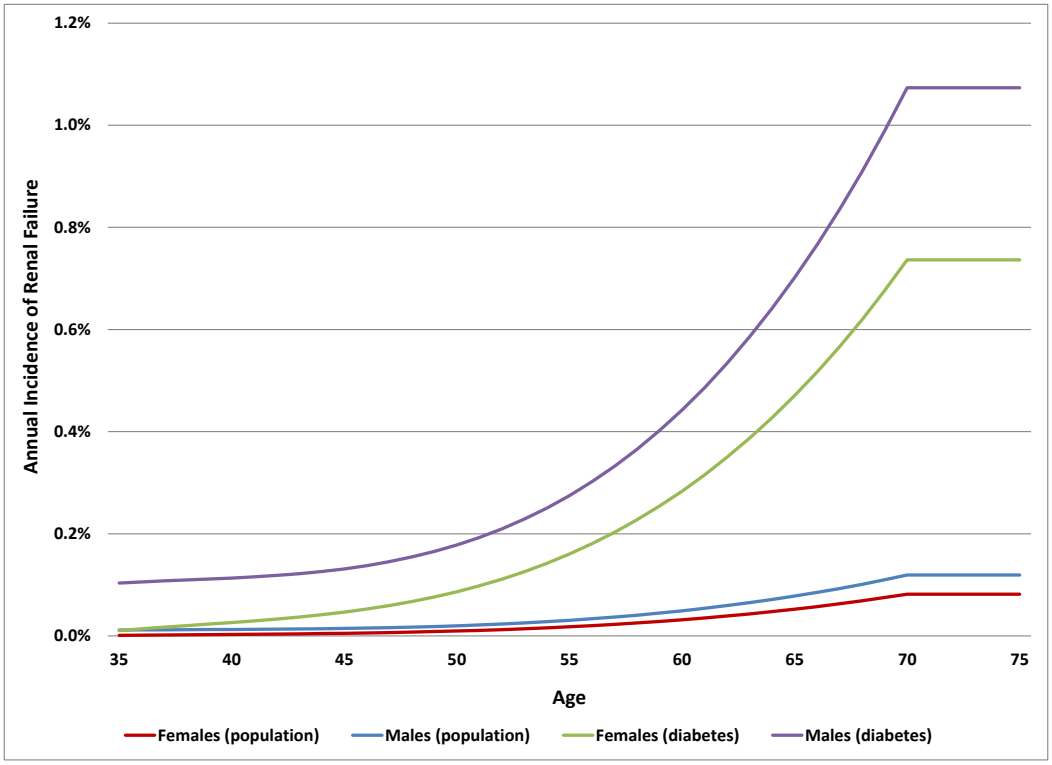
Renal Failure

A literature search to find predictive equations for the incidence of renal failure yielded no usable results. Initial estimates of renal failure incidence from the UKPDS Outcomes Model produced estimates that were lower for the diabetes population than estimated rates for the prediabetes population. As such, incidence rates from Hippisley-Cox and Coupland's analysis of over 1.57 million individuals from 368 QResearch general practices served as the basis for the prediction.[21] That publication provided incidence rates of renal failure in 5 year age bands from age 35-74 for both men and women. A quadratic trend line was fit to each sex to generate annual probability of incident renal failure for the population (Exhibit 105). Incidence rates for the diabetes and non-diabetes populations were calculated using estimates that the relative risk for the population with diabetes was about 9.0 time greater than the risk for the population without diabetes (as calculated from the Multiple Risk Factor Intervention Trial).[22]

Retinopathy

Similar to amputations, retinopathy attributable to diabetes was modeled but all other retinopathy was assumed to be independent of body weight. For people with diabetes, the UKPDS Outcomes Model's equation for retinopathy attributable to diabetes was uses to model incidence.[33] Risk factors modeled are age at diabetes diagnosis, time since diabetes diagnosis, sex, HbA1c, SBP, and cholesterol ratio. Summary statistics for simulated annual incidence among people with diabetes is 1.9%.

Exhibit 105. Renal Failure Incidence Rate by Age

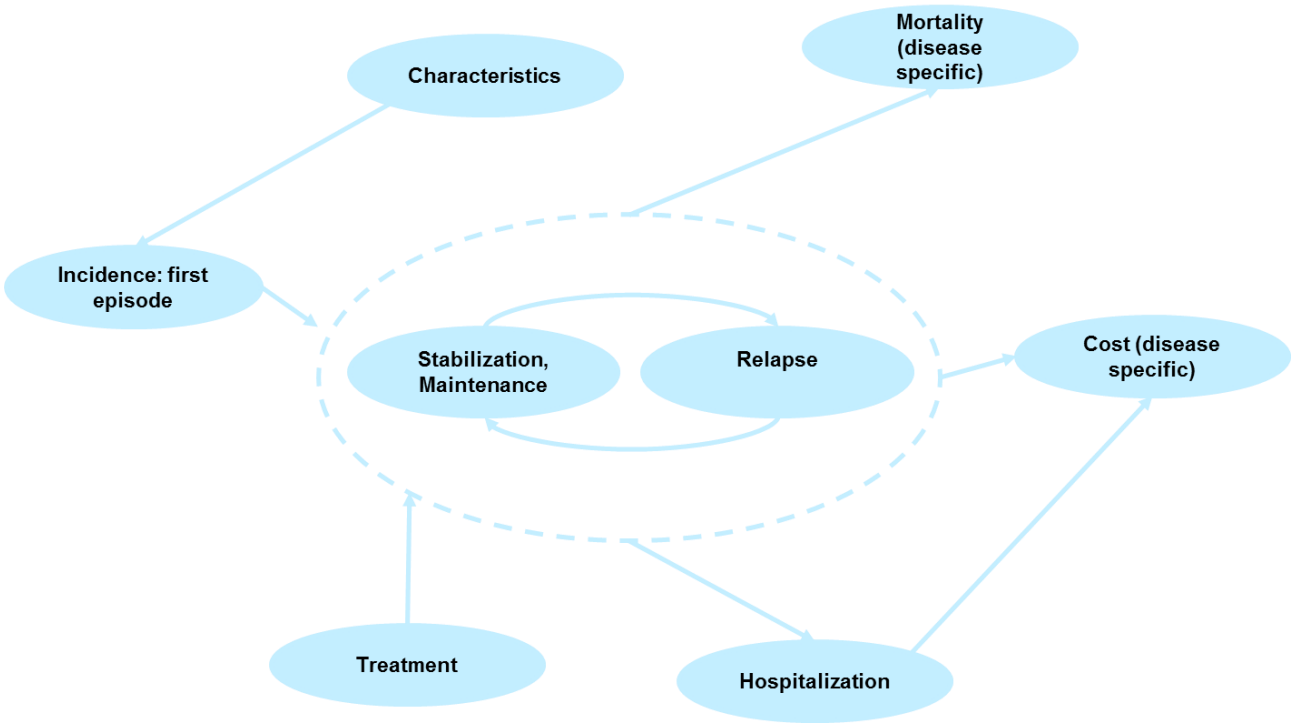


Schizophrenia

Schizophrenia is a life-long chronic illness and will be modeled as such in the DPMM. Treatment goals for the condition center on the following: 1) reduce or eliminate symptoms, 2) maximize quality of life and adaptive functioning, and 3) enable recovery by assisting patients in attaining persona life goals.106 The simulation of schizophrenia in the model will be done as shown in Exhibit 106.

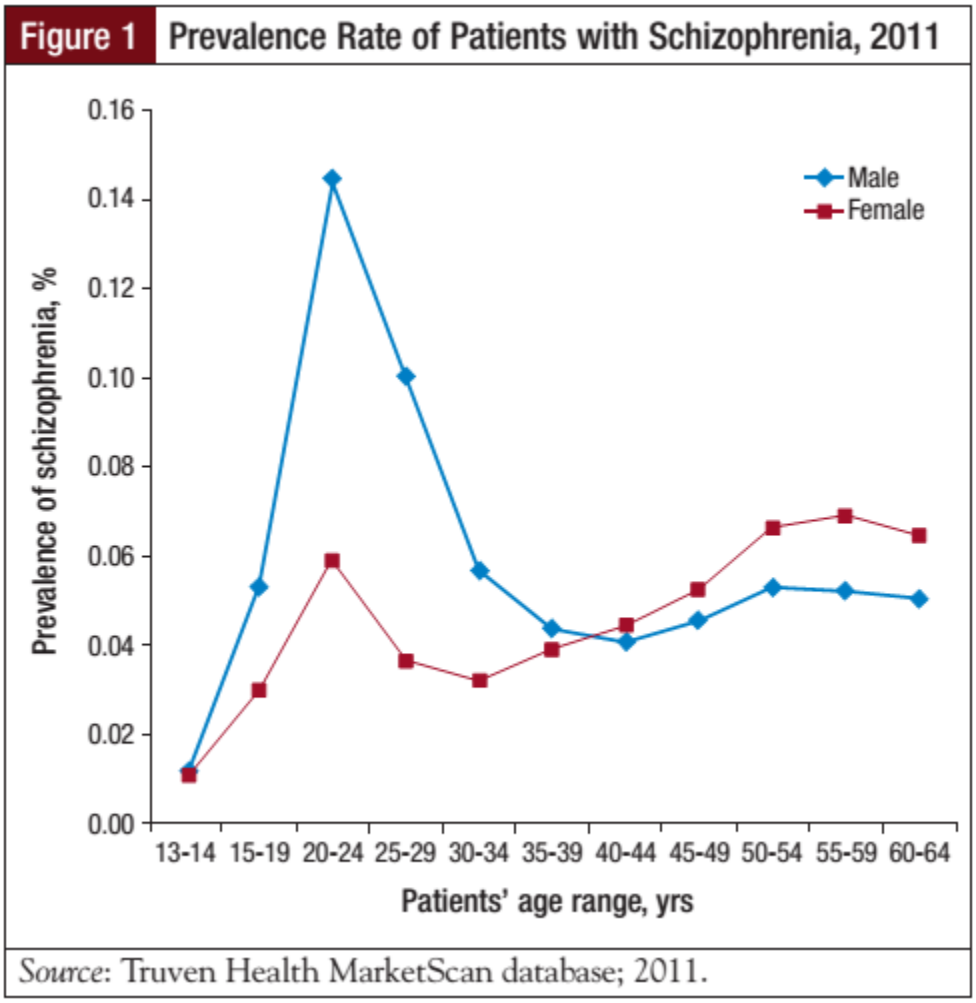
[106] Lehman AF et al. Practice Guideline for the Treatment of Patients With Schizophrenia Second Edition. Work Group on Schizophrenia 2005

Exhibit 106. Influence Diagram for Schizophrenia



Prevalence: In 1993, the National Institute of Mental Health (NIMH) reported that the prevalence of schizophrenia in the USA was 1.1% of the adult population.[107] In their 2006 paper Wu et al. report that the estimated *lifetime* prevalence of schizophrenia and schizophreniform disorder in community epidemiological surveys using fully structured lay-administered diagnostic interviews have been in the range 0.3–1.6%.[108] Fitch et al. analyzed MarketScan claims database and reported prevalence rate as depicted in Exhibit 107.[109]

Exhibit 107. Prevalence Rate of Schizophrenia



The final prevalence rates used in the model are converted from Exhibit 107 above (keeping 2 decimal places).

[107] http://www.nimh.nih.gov/health/statistics/prevalence/schizophrenia.shtml
[108] Wu EQ et al. Annual prevalence of diagnosed schizophrenia in the USA: a claims data analysis approach. Psychological Medicine, 2006, 36, 1535–1540
[109] Fitch, K, Iwasaki, K, Villa, K, Resource utilization and cost in a commercially insured population with schizophrenia, Am Health Drug Benefits, 2014, 7(1):18-26

Exhibit 108. Prevalence Rates for Schizophrenia for Model Inputs (%)

Age group	Prevalence rate (Male)	Prevalence rate (Female)
20-24	0.15%	0.06%
25-29	0.10%	0.04%
30-34	0.06%	0.04%
35-39	0.04%	0.04%
40-44	0.04%	0.04%
45-49	0.04%	0.05%
50-54	0.05%	0.06%
55-59	0.05%	0.06%
60-64	0.05%	0.06%
65+	0.05%	0.06%

Incidence: A 2014 systematic review, by Van Der Werf et al. sought to recalculate the incident rates from published studies by age and sex, hoping to update previous estimates by the inclusion of new studies that were more recently published.[110] However, of the papers included in the meta-analysis, the only paper set in the USA was a 1967 paper by Malzberg et al. Cowan et al. examined the incidence of adult onset schizophrenic disorders in the US military.[111] While the obvious caveat is that the demographics of the military population are not an equivalent match to that of the US civilian population, the military is drawn from all socioeconomic and educational sectors of the US, as well as all states and territories. The study population includes both sexes, and members of all racial subgroups, and the age range 17 to over 60 years.

The most detailed data come from Fitch et al., which provided incidence rates by gender and age. (Exhibit 109) Final model inputs can then be calculated in Exhibit 110 (rounded to the nearest 0.005%)

[110] Van der Werf et al. Systematic review and collaborative recalculation of 133693 incident cases of schizophrenia. Pysch Med 2014

[111] Cowan DV et al. Incidence of adult onset schizophrenic disorders in the US Military: Patterns by sex, race and age. Schizophrenia Research. 2011

Exhibit 109. Incidence Rates of Schizophrenia

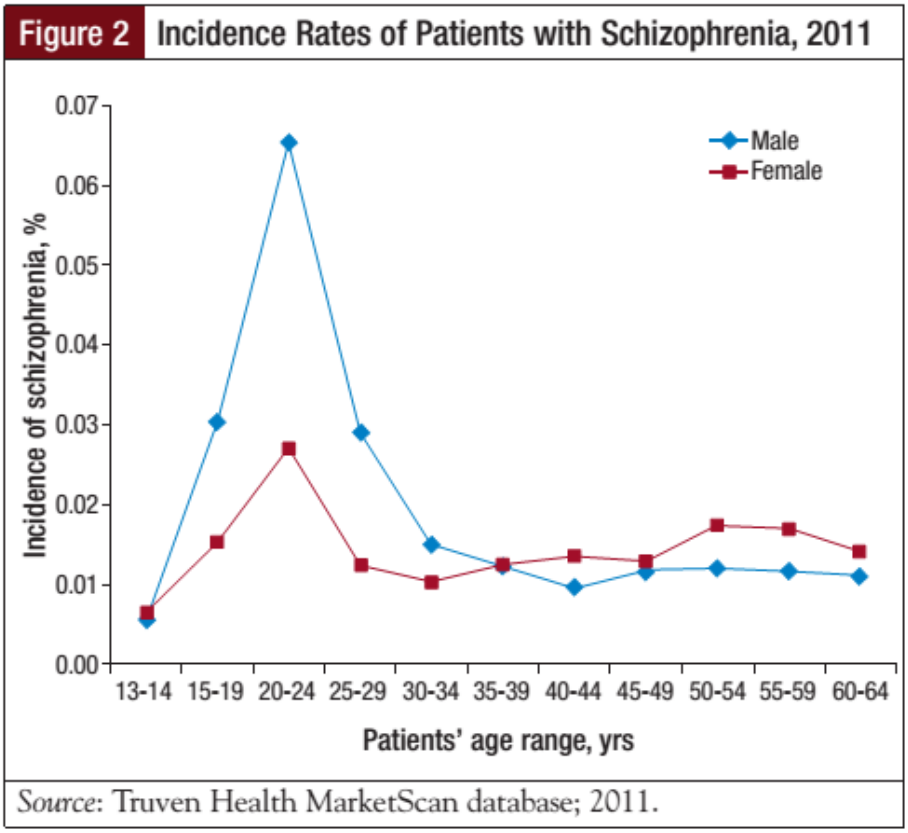


Exhibit 110. Incidence Rate of First Schizophrenic Hospitalization by Sex and Age (%)

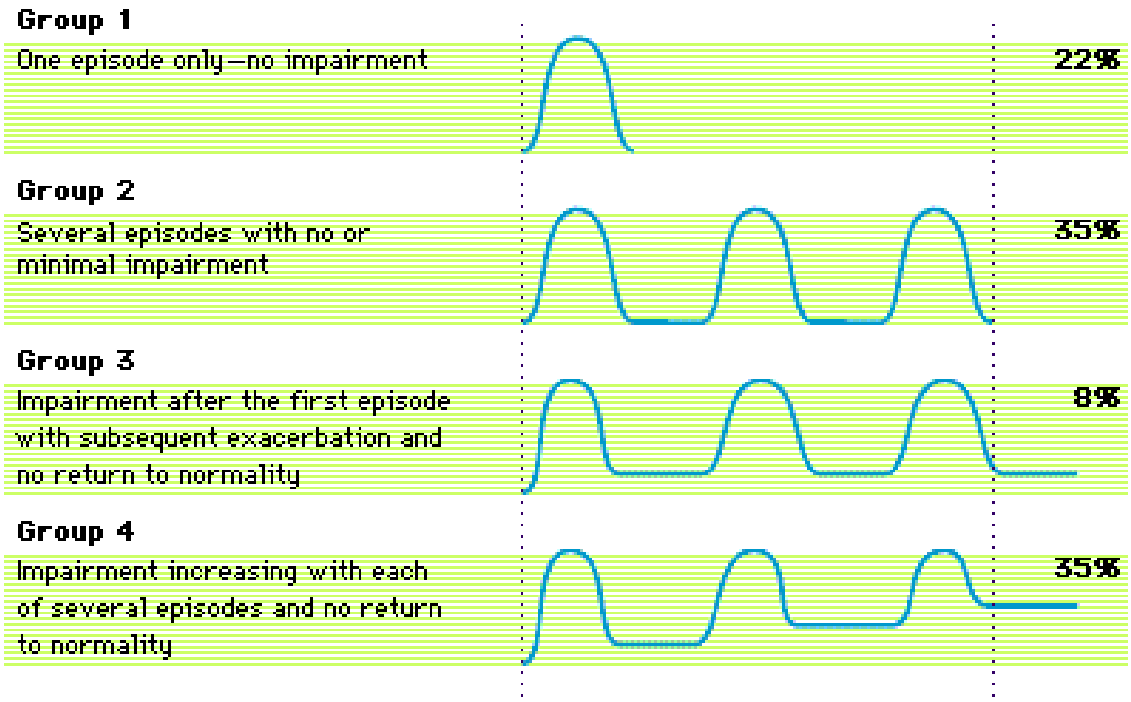
Age group	Prevalence rate (Male)	Prevalence rate (Female)
20-24	0.065%	0.030%
25-29	0.030%	0.010%
30-34	0.015%	0.010%
35-39	0.010%	0.010%
40-44	0.010%	0.010%
45-49	0.010%	0.010%
50-54	0.010%	0.015%
55-59	0.010%	0.015%
60-64	0.010%	0.015%
65+	0.010%	0.015%

Data in Exhibit 110 can be verified independently by Cowan et al.[111] For instance, males usually experience disease onset earlier than females, echoed in the findings of Cowan et al. It has also been noted that women

appear to have two peaks in the age of onset of disease, the fist after menarche and the second, after age 40, which the Cowan data reflects to an extent, with the incidence rate in females being higher than the males in the age 35+ category.[112]

Course of Disease: Schizophrenia is characterized by multiple relapses in most patients who have been diagnosed with the condition. It is noted that there is variation in patient experience, with some suffering only one episode and no permanent impairment, while at the other end of the spectrum, others may suffer multiple episodes and increasing impairment after each. While approximately 20% of patients who have a psychotic break will not have another, the majority of patients experience multiple relapses as seen in Exhibit 111.[113]

Exhibit 111. Schizophrenia Disease Course Variance



Csernansky et al. reported that the risk of monthly relapse in schizophrenia is 3.5% in patients treated with depot antipsychotic drugs, resulting in an annual relapse rate of 42%. Non-compliance amongst patients was estimated to be 7.6% per month, and in these patients, relapse rates increased to 11% per month, making an annual relapse almost a certainty. [114]

[112] Ochoa S et al. Gender Differences in Schizophrenia and First-Episode Psychosis: A Comprehensive Literature Review. Schizophrenia Research and Treatment Volume 2012, Article ID 916198
[113] http://www.schizophrenia.com/szfacts.htm
[114] Csernansky et al. Relapse and Rehospitalisation Rates in Patients with Schizophrenia Effects of Second Generation Antipsychotics. CNS Drugs. 2002

The diagnosis of schizophrenia is often through the first hospitalization. To model the natural course of disease, we will use the relapse and re-admittance rates of patients on placebo as reported by Leucht et al in Exhibit 112.[115] Leucht et al note a lack of evidence pointing to differences in the efficacy of available antipsychotic drugs, and therefore assume any treatment has a similar effect in terms of the outcome of preventing relapses.

Exhibit 112. Probability of Annual Relapse And Re-Admittance for Those Treated with Drugs vs. Placebo

Outcomes	Drug group	Placebo group	Risk ratio
% of patients who relapsed	27%	64%	0.40
% of patients readmitted (% of the total patient population)	10%	26%	0.38

Mortality: Kor et al. report that patients with schizophrenia are known to die earlier than expected, with up to 40% of excess premature mortality attributable to suicide and unnatural death. It is also reported that the lifetime suicide risk for those with schizophrenia is 4.9%.[116] Compared with the general population, schizophrenia patients have a 8.5 fold greater risk of suicide.[117] For the purpose of modeling, we'll use the 40% excess premature mortality as the basis of calculation.

Kasckow et al. report that clozapine, a second generation agent, reduced suicides rates by 88% two years after the start of treatment. In another study conducted over 1 year, current clozapine users had an 83% reduction in death by suicide compared to those who were using the drug but then stopped. National statistics report that 60% of schizophrenic patients get treated, so this population will have their lifetime suicide risk reduced by an average of 86%. For the 40% who are untreated, the suicide rate (x) has been calculated using the following equation:

$$x*40\% + x*(1-86\%)*60\% = 40\%$$

$$x = 82.6\%$$

This means untreated Schizophrenia patients have *82.6%* higher chance of dying due to unnatural causes. Treated cases have 82.6%*(1-86%)= *11.6%* higher chance of dying.

Treatment effect: As our model simulates relapse and hospitalizations as the primary outcomes of schizophrenia, we will use the reduction in % relapse and % re-admittance as measures of treatment effect. (Exhibit 112) Treatment effect will also be modeled via reduced mortality.

[115] Leucht et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. Lancet. 2012
[116] Hor K et al. Suicide and schizophrenia: a systematic review of rates and risk factors. J Psychopharmacol 2010.
[117] Kasckow J et al. Managing Suicide Risk in patients with Schizophrenia. CNS Drugs. 2011

Cost: It is estimated that 40% of individuals with schizophrenia are untreated in any given year.[118] The condition is considered the most debilitating of all mental illnesses, and is estimated to cost approximately USD$63 billion a year (direct, societal and family costs) with 30% attributed to direct treatment. [119]

In Fitch's analysis of a commercially insured (treated) population, the cost of newly diagnosed schizophrenia cases in the first 2 years is reported below.[109]

Exhibit 113. First-Month Cost of Newly Diagnosed Schizophrenia Cases (2011 USD)

Table Snapshot Analysis: Mean PPPM and PMPM Costs and Resource Utilization for Patients with Schizophrenia Compared with Demographically Adjusted Total Commercially Insured Population

Variable	Patients with schizophrenia	Matched population with similar demographics, without schizophrenia
Total PPPM/PMPM cost, $	1806	419
Inpatient PPPM/PMPM cost, $	762	97
Outpatient PPPM/PMPM cost, $	592	239
Prescription drug PPPM/PMPM cost, $	452	83
Annual inpatient admissions per 1000 patients, N	636	48
Schizophrenia-related inpatient cases, N	322	—
Psychiatric/nonschizophrenia-related inpatient cases, N	155	—
Nonschizophrenia/nonpsychiatric-related inpatient cases, N	158	—
Annual emergency department visits per 1000 patients, N	2270	158
Schizophrenia-related emergency department visits, N	242	—
Psychiatric/nonschizophrenia-related emergency department visits, N	513	—
Nonschizophrenia/nonpsychiatric-related emergency department visits, N	1516	—

PMPM indicates per member per month; PPPM, per patient per month.
Source: Truven Health MarketScan database; 2011.

The cost attributable to schizophrenia can be calculated as the difference in cost between patients with and without the condition. (Exhibit 114) Cost attributable to schizophrenia accounts for 77% ($1387/$1806) of the total direct medical cost of patients with schizophrenia.

Exhibit 114. First-Month Cost of Treated Schizophrenia Cases (2011 USD)

Setting	Cost attributable to Schizophrenia	% of total cost
Inpatient	$665	48%
Outpatient	$353	25%
Rx	$369	27%
Total	**$1,387**	**100%**

[118] http://www.treatmentadvocacycenter.org/problem/consequences-of-non-treatment/schizophrenia
[119] http://www.schizophrenia.com/szfacts.htm

The same study reported the average total cost in the first and second year to be $23,512 and $15,252, respectively. Assuming 77% of these costs are directly related to schizophrenia, and that the distribution of inpatient, outpatient, and Rx cost remains the same as in the first month, we get the following table (inflated to 2015 USD).

Exhibit 115. Cost Attributable to Schizophrenia for Treated Patients (2015 USD)

Setting	First year	Subsequent years	% of total cost
Inpatient	$9,629	$6,246	48%
Outpatient	$5,015	$3,253	25%
Rx	$5,416	$3,514	27%
Total	**$20,060**	**$13,013**	**100%**

According to Exhibit 112, untreated patients have 2.37 (64%/27%) times more relapses and 2.6 (26%/10%) times more hospitalizations. Assuming outpatient visits for untreated patients has linear correlation with relapses, the cost for untreated patients can be calculated based on Exhibit 115.

Exhibit 116. Cost Attributable to Schizophrenia (2015 USD) For Untreated Patients

Setting	First year	Subsequent years
Inpatient	$25,035	$16,240
Outpatient	$11,886	$7,710
Rx	0	0
Total	$36,921	$23,950

Kazuhiro et al. reported the indirect cost of schizophrenia in the US is about the same as direct cost.[120] **Please note the indirect cost here includes absenteeism, presenteeism, and caregiver cost, and thus does not need to be converted to presenteeism like for other conditions.**

Exhibit 117. Total Indirect Cost of Schizophrenia (2015 USD)

	First year	Subsequent years
Treated	$20,060	$13,013
Untreated	$36,921	$23,950

Key assumptions:

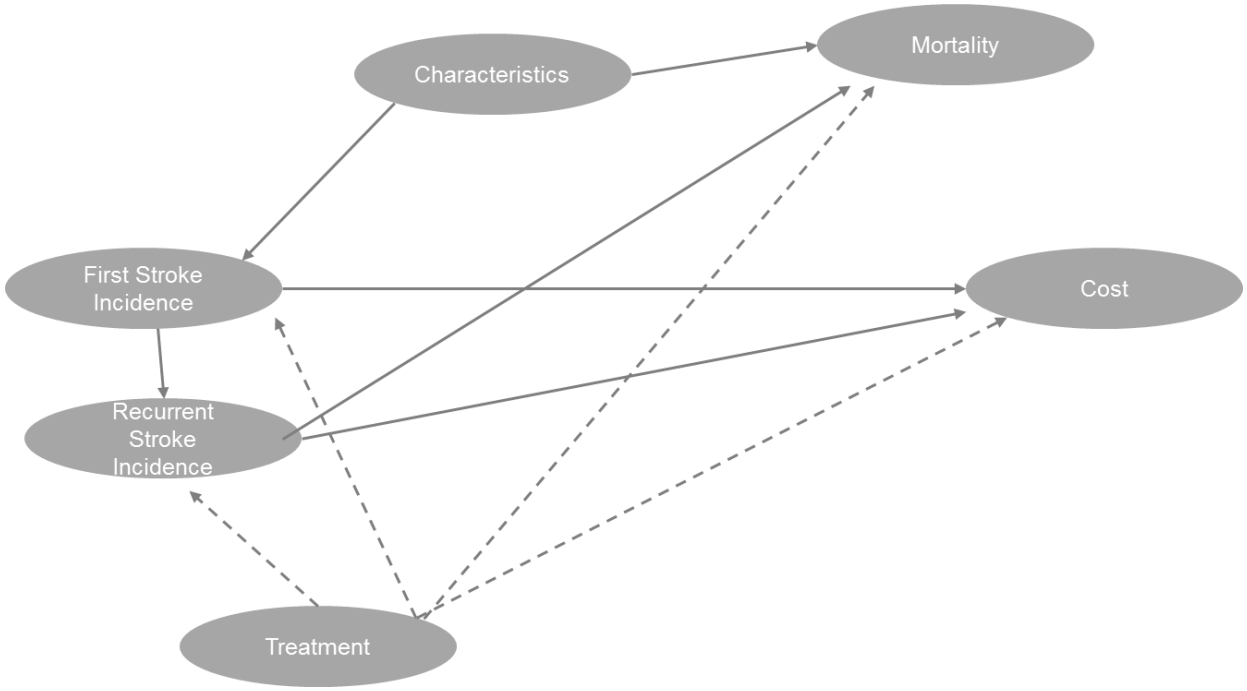
- Assuming the distribution of inpatient, outpatient, and Rx cost remains the same as in the first month

[120] Kazuhiro, TP, et al., Understanding the direct and indirect costs of patients with schizophrenia, F1000Res, Jul 6, 2015

Stroke

Stroke is included in the DPMM as an acute condition. As depicted in Exhibit 118, the modeling of stroke is based on disease occurrence and the resulting medical resource use and mortality. Risk of subsequent (recurrent) stroke is modeled separately from first stroke. Excess mortality risk and costs from recurrent strokes are modeled as well.

Exhibit 118. Influence Diagram for Stroke



First Stroke Incidence: Incidence of first stroke for both the diabetes and non-diabetes populations was predicted using risk functions from the Framingham Heart Study.[121] Separate equations for men and women model stroke risk as a function of age, SBP, diabetes status, smoking, history of cardiovascular disease, atrial fibrillation, LVH, and whether or not a person is on anti-hypertensive medication. The equations predict 10-year stroke risk, which we converted to annual risk by assuming equal probability in each of the 10 years. These equations are displayed in Exhibit 119 below.

[121] D'Agostino RB, Wolf PA, Belanger AJ, Kannel WB. Stroke risk profile: adjustment for antihypertensive medication. The Framingham Study. Stroke 1994; 25(1):40-43.

Exhibit 119. Risk Calculator for First Stroke

Women: Probability of Stroke Within 10 Years

	Points					
	0	+1	+2	+3	+4	+5
Age,y	54-56	57-59	60-62	63-64	65-67	68-70
Untreated SBP, mmHg		95-106	107-118	119-130	131-143	144-155
Treated SBP, mmHg		95-106	107-113	114-119	120-125	126-131
Diabetes	No			Yes		
Cigs	No			Yes		
CVD	No		Yes			
AF	No					
LVH	No				Yes	Yes

	Points				
	+6	+7	+8	+9	+10
Age,y	71-73	74-76	77-78	79-81	82-84
Untreated SBP, mmHg	156-167	168-180	181-192	193-204	205-216
Treated SBP, mmHg	132-139	140-148	149-160	161-204	205-216
Diabetes					
Cigs					
CVD					
AF	Yes				
LVH					

Points	10-Year Probability, %	Points	10-Year Probability, %	Points	10-Year Probability, %
1	1	11	8	21	43
2	1	12	9	22	50
3	2	13	11	23	57
4	2	14	13	24	64
5	2	15	16	25	71
6	3	16	19	26	78
7	4	17	23	27	84
8	4	18	27		
9	5	19	32		
10	6	20	37		

Men: Probability of Stroke Within 10 Years

	Points					
	0	+1	+2	+3	+4	+5
Age,y	54-56	57-59	60-62	63-65	66-68	69-72
Untreated SBP, mmHg	97-105	106-115	116-125	126-135	136-145	146-155
Treated SBP, mmHg	97-105	106-112	113-117	118-123	124-129	130-135
Diabetes	No		Yes			
Cigs	No			Yes		
CVD	No				Yes	
AF	No				Yes	
LVH	No					Yes

	Points				
	+6	+7	+8	+9	+10
Age,y	73-75	76-78	79-81	82-84	85
Untreated SBP, mmHg	156-165	166-175	176-185	186-195	196-205
Treated SBP, mmHg	136-142	143-150	151-161	162-176	177-205
Diabetes					
Cigs					
CVD					
AF					
LVH					

Points	10-Year Probability, %	Points	10-Year Probability, %	Points	10-Year Probability, %
1	3	11	11	21	42
2	3	12	13	22	47
3	4	13	15	23	52
4	4	14	17	24	57
5	5	15	20	25	63
6	5	16	22	26	68
7	6	17	26	27	74
8	7	18	29	28	79
9	8	19	33	29	84
10	10	20	37	30	88

Recurrent Stroke: Incidence of recurrent stroke is modelled based on two sources. For the first year after first stroke, recurrent stroke risk was estimated based on data from the Reduction of Atherothrombosis for Continued Health (REACH) Registry using the Essen Stroke Risk Score (ESRS) as outlined by Weimar, et al.[122] The ESRS takes into account factors such as age, smoking status, and presence of comorbidities. Exhibit 120 summarizes the annual risk of recurrent stroke based on ESRS from the REACH study.

Exhibit 120. Annual Risk of Recurrent Stroke

	REACH		
ESRS Point Sum	N	Event Rate, %	95% CI
0	222	1.82	0.00–3.94
1	1317	2.87	1.81–3.92
2	3017	3.31	2.48–4.14
3	4208	3.92	3.13–4.71
4	3559	4.37	3.48–5.25
5	2054	4.81	3.63–5.98
6	893	4.71	3.05–6.34
>6	335	6.84	3.65–9.92
Total	15 605	4.01	3.46–4.56

In subsequent years the risk of recurrent stroke comes from an older study based on data from the Oxfordshire stroke project.[123] These rates are summarized in Exhibit 121 below.

Exhibit 121. Annual Risk of Recurrent Stroke (2)

	0-6 Months	6-12 Months	2 Years	3 Years	4 Years	5 Years
% Risk, 95% confidence interval	8.6 6.5-10.7	4.6 2.6-6.6	6.7 2.7-7.3	5.0 1.0-5.6	3.3 0.0-3.0	1.3 3.0-15.0
% Cumulative risk, 95% confidence interval	8.6 6.5-10.7	13.2 10.0-16.4	19.9 15.3-23.7	24.9 19.2-30.4	28.2 21.3-34.9	29.5 19.8-39.0
No. at risk	675	463	420	339	233	167

Treatment effect: Treatment effect(s) will be expressed as a relative risk adjustment to incidence of recurrent stroke event and/or mortality.

[122] Weimar, C., H.-C. Diener, M. J. Alberts, P. G. Steg, D. L. Bhatt, P. W.f. Wilson, J.-L. Mas, and J. Rother. "The Essen Stroke Risk Score Predicts Recurrent Cardiovascular Events: A Validation Within the REduction of Atherothrombosis for Continued Health (REACH) Registry." Stroke 40.2 (2008): 350-54. Web.

[123] Burn, J., M. Dennis, J. Bamford, P. Sandercock, D. Wade, and C. Warlow. "Long-term Risk of Recurrent Stroke after a First-ever Stroke. The Oxfordshire Community Stroke Project [published Erratum Appears in Stroke 1994 Sep;25(9):1887]." Stroke 25.2 (1994): 333-37. Web.

Mortality: For first stroke age and sex specific mortality probabilities reflect 1-year mortality rates witnessed in the Arcadia Stroke Registry.[124] Rates were available for ages 18-54, 55-64, 65-74, 75-84, and 85+. These rates are summarized in Exhibit 122 below.

Exhibit 122. 1-Year Mortality Probability From First Stroke

	All events	
Age (y)	*D/S*	*%*
Men:		
18–54	3/18	16.7
55–64	7/36	19.4
65–74	18/74	24.3
75–84	40/106	37.7
≥85	38/75	50.7
Total	106/309	34.3
Women:		
18–54	2/9	22.2
55–64	7/29	24.1
65–74	18/55	32.7
75 84	40/99	40.4
≥85	31/53	58.5
Total	98/245	40.0

For recurrent strokes, Aarnio, et al. conducted a multi-variate Cox proportional hazards analysis of Finnish patients with recurrent strokes and found that recurrent stroke was associated with a mortality hazard ratio of 16.68.[125] This hazard ratio is used to adjust the data from first stroke to account for the excess mortality risk of recurrent stroke.

Cost: The direct medical costs of stroke come from regression analysis with the 2006-2010 Medical Expenditure Panel Survey (MEPS) Full Year Consolidated Data File and Medical Conditions File. We used a generalized linear model (GLM) with gamma distribution and log link to reflect the skewed distribution of annual medical expenditures. The dependent variable was total annual medical expenditures, while age, sex, insurance type, and modeled diseases were the independent variables. All medical cost estimates were converted to 2013 dollars using the medical component of the consumer price index. The American Heart Association's 2015 Update of its Heart Disease and Stroke Statistics publication notes that recurrent stroke patients incur 38% greater medical expenditures than first stroke patients, and this scalar is, thus, applied to the cost estimates for first stroke.

[124] Vemmos KN 1205 , Bots ML, Tsibouris PK, Zis VP, Takis CE, Grobbee DE et al. Prognosis of stroke in the south of Greece: 1 year mortality, functional outcome and its determinants: the Arcadia Stroke Registry. Journal of Neurology, Neurosurgery & Psychiatry 2000; 69(5):595-600.

[125] Aarnio, K., E. Haapaniemi, S. Melkas, M. Kaste, T. Tatlisumak, and J. Putaala. "Long-Term Mortality After First-Ever and Recurrent Stroke in Young Adults." Stroke 45.9 (2014): 2670-676. Web.

Absenteeism due to CHF, MI, or stroke is accounted for in one regression equation. The equation will be able to separately predict the number of missed workdays due to each of the 3 conditions.

Absenteeism will be modeled via a regression analysis on the MEPS. The dependent variable in the Poisson regression will be the number of missed workdays in a year the individual experienced and independent variables will include demographics (i.e. age, sex, race, etc.), socioeconomic characteristics (i.e. insurance status/type, annual income, etc.), biometrics (i.e. BMI, SBP, cholesterol ratio, etc.), and disease status (dummy variables for all modeled conditions). For each individual who has been simulated to be employed in a given cycle this equation will be used with his or her characteristics to predict the number of missed workdays in that year.

Treatment effects can be applied in this approach in two ways, both broadly working within a risk-reduction framework.

- If it is assumed that a treatment will decrease the number of missed workdays by a set amount, that amount can be applied to the total predicted count of missed workdays in the cycle.
- If, however, the assumption is that the treatment reduces missed workdays due to a certain condition by a percentage, the risk reduction can be applied directly to the coefficient in the estimation of the total number of missed workdays. For example, a treatment may reduce the total number of missed workdays of patients by 2 (modeled by subtracting 2 from the total predicted days missed for each patient) or it may half the risk attributable to the disease (modeled by halving the exponentiated coefficient of that disease).

The relative reduction in absenteeism due to treatment will be synchronized with the relative reduction in MI events, to reflect the correlation between reduction in MI event and decrease in absenteeism.

Long term care: According to Kapral et al., 10% of women and 5% of men are admitted to long-term care after a stroke.[126] Annual direct cost of long-term care is assumed to be the same as those for Alzheimer's patients, which is $61,436 per year (2015 cost)

Key assumptions:

- 1/10 of the ten year risk of stroke is a valid approximation of annual risk
- The older Oxfordshire stroke project rates represent current stroke recurrence rates in the US

Health Conditions Associated with Obesity

[126] Kapral, MK, et al., Sex difference in stroke care and outcomes results from the registry of the Canadian stroke network, Stroke, 2005

One main objective of the model was to predict the economic and societal impact of obesity-related comorbidities. We conducted a review of published literature and identified many health conditions that are associated with obesity (Exhibit 123).

Exhibit 123. Summary of Obesity Comorbidities

Comorbidity	Risk Ratio (obese vs normal weight)	Prevalence* or Incidence	Notes
Metabolic syndrome			
Hypertension	F: 2.42 M: 1.84 [35]	67 M * [a]	Risk factor for other conditions
Dyslipidemia		71 M * [a]	Risk factor for other conditions
High blood sugar		105 M * [a]	Model HbA1c; risk factor for other conditions
Diabetes (type 2)	F: 12.41 M: 6.74 [35]	1.9 M [b]	
Cardiovascular diseases			BMI affects cardiovascular disease/event risk both directly and indirectly through increased risk for diabetes and metabolic syndrome
Congestive heart failure	F: 1.78 M: 1.79 [35]	5.1 M * [a]	
Ischemic heart disease	F: 3.10 M: 1.72 [35]		
Myocardial infarction		715,000 [b]	
Left ventricular hypertrophy			Included as a risk factor for myocardial infarction, stroke, and ischemic heart disease
Atrial fibrillation		2.66 M * [a]	Modeled as a risk factor for other conditions; 1 point increase in BMI = 4% increase in AF risk [108]; obesity increased AF risk by 49% with risk escalated in parallel with increased BMI [109]
Cerebrovascular diseases			
Stroke	F: 1.49 M: 1.51 [35]	795,000 [b]	1 point increase in BMI = 6% increase in relative risk for stroke [110]
Cancers			
Breast	RR by BMI Ranges[41] Premenopausal: <22.5=0.96, 25-27.4=0.93, 27.5-29.9=0.99, ≥30 = 0.79 Postmenopausal: <22.5=0.85, 25-27.4=1.1, 27.5-29.9=1.21, ≥30 = 1.29	235,670 c See incidence table for rates	Most studies used BMI; few used WC [35]
Cervical	RR by BMI Ranges [41] F: <22.5=0.96, 25-27.4=0.94, 27.5-29.9=0.79, ≥30 = 1.02	12,360 [c] See Incidence Tables	Few studies have investigated link between incidence risk and body weight; BMI associated with increased

			cancer mortality risk; incidence is low
Colorectal	RR by BMI Ranges[42] F: 25-30:1.03, ≥30 =1.07 M: 25-30:1.16, ≥30 =1.4	96,830 [c] See Incidence Tables	Incidence rates from SEER database
Endometrial	RR by BMI Ranges [43] F: 25-29.9:1.47, 30 – 34.9=2.66, 35-39.9=4.39, ≥40 =7.98	65,630 [c] See Incidence Tables	Most studies used BMI; few used WC [35] Study among European countries estimated 26% to 47% of endometrial cancer cases attributed to overweight and obesity.[111]
Esophageal	OR BY BMI Ranges [44] F: 22-24=0.8, 24-27=2.1, >27=2.6 M: 23-25:1.5, 25-27=2.0, >27=3.0	18,170 [c] See Incidence Tables	Few studies have looked at the link to BMI, the risk ratio is relatively small, and incidence is low [35]
Gallbladder	RR by BMI Ranges [45] F: ≥30 = 1.88 M: ≥30 = 1.35	10,650 [c] See Incidence Tables	Low incidence and only most link to excess body weight
Kidney	RR by BMI Ranges [112] F: <18.5=1.7, 22.5-<25=1.11, 25-<27.5=1.57, 27.5-<30=1.66, 30-<35=2.16, ≥35 = 2.59 M: <18.5=1.32, 22.5- <25=1.15, 25-<27.5=1.43, 27.5-<30=1.64, 30-<35=1.87, ≥35 = 2.47	63,920 [c] See Incidence Tables	Studies used BMI [35]
Leukemia	RR by BMI Ranges[46] B: 25 - <30 = 1.14, ≥ 30 = 1.39	52,380 [c] See Incidence Tables	
Liver	RR by BMI Ranges[113] B: 25-30 = 1.17, ≥30=1.89	33,190 [c] See Incidence Tables	
Multiple myeloma	RR by BMI Ranges [48] F: 22-<2.5=1.1, 25-<30=1.6, ≥30 =1.2 M: 22-<2.5=1.3, ≥30 =2.4	24,050 [c] See Incidence Tables	Link with excess body weight isn't particularly strong; incidence is low
Non-Hodgkin's lymphoma	RR by BMI Ranges [49] B: 25-29.9=1.05, 30-34.9=1.07, ≥35 = 1.29	70,800 [c] See Incidence Tables	
Ovarian	RR by BMI Ranges[41]	21,980 [c]	Link with excess body weight isn't particularly strong;

	<22.5=0.98, 25-27.4=0.99, 27.5=29.9=1.13, ≥30=1.12	See Incidence Tables	incidence is low
Pancreatic	RR by BMI Ranges[50] F:<21 =1.15, 23-24.9=1.08, 25-29.9=1.29, ≥30 = 1.46 M:<21 =1.19, 23-24.9=1.07, 25-29.9=1.09, ≥30 = 1.5	46,420 [c] See Incidence Tables	Studies used BMI [35]
Prostate	RR by BMI Ranges [51] 30 - <35= 0.97, 35 - <40 = 0.84, ≥40 = 0.65	233,000 [c] See Incidence Tables	
Stomach	RR by BMI Ranges [114] F: 25 – 30 = 0.99, ≥30 =1.04 M: 25 – 30 = 1.01, ≥30 =1.12	22,220 [c] See Incidence Tables	
Thyroid	RR by BMI Ranges[52] F: 25-29.99 = 1.11, ≥30 = 1.11 M: 25-29.99 = 1.53, ≥30 = 1.89	60,220 [f]	The study showed a stronger association between BMI and thyroid cancer in men vs. women
Musculoskeletal			
Osteoarthritis	BMI Ranges F:25 –29.99 = 1.8, ≥30 = 1.96 M: 25 –29.99= 2.8, ≥30 =4.20 [35]	26.9 M * [a] Total Annual Incidence: 0.0043 [36]	Most people with osteoarthritis have low annual medical costs associated with their condition, but many receive surgery or other treatment which can be costly. The prevalence is high, and the relationship with body weight is strong.
Chronic back pain	BMI Ranges F:25 –29.99 = 1.6, ≥30 = 2.8 M: 25 –29.99= 1.6, ≥30 =2.8 [35]	Age - Annual Incidence 18-24 - 0.0344 25-44 - 0.0433 45-64 - 0.0519 65+ - 0.0455	Estimated 149 million work days lost due to chronic back pain; total costs $100-$200 billion annually, 2/3 of cost due to reduced productivity
Sleep apnea/respiratory problems			
Pneumonia	Hazard ratios for BMI Ranges [37]	1.1 M hospitalizations [b]	Findings indicate that associations between obesity and pneumonia are stronger in males than in females, where

	F: <22.5 = 1.2, 25.0 – 29.9 = 0.8, 30.0-34.9 = 0.7, ≥35 = 0.8 M: <22.5 = 1.4, ≥35 = 1.2		occurrence of comorbid chronic diseases weakens the relationship [37]
Pulmonary embolism	Relative risks for BMI ≥ 30 by age [38] <40 =5.19 40-49 = 1.94 50-59 = 1.25 60-69 = 1.42 70-79 = 2.07 80+ = 3.15	900,000 [115] See incidence rate table.	Study indicates that obesity is a risk factor for pulmonary embolism in both genders, but is stronger in women, and in both genders under age 40.[38]
Diseases of the gall bladder, kidney, liver and gastrointestinal systems			
Renal failure			Body weight has indirect effect through hypertension and diabetes
Chronic Kidney Disease			Used data from the Atherosclerosis Risk in Communities Study and the Cardiovascular Health Study.[17]
Gallstones and gallbladder disease	Odds ratios for BMI Ranges [23] **F**: 22.0 – 24.9 = 1.4, 25.0-29.9 = 2.3, 30.0-34.9=3.2, ≥35 = 3.7 **M**: 22.0 – 24.9 = 1.2, 25.0-29.9 = 1.6, 30.0-34.9=2.5, ≥35 = 3.3	See Incidence Table	The risk of developing gallstones increased with increasing overweight status [23]
Gastroesophageal reflux disease	Odds ratios for BMI Ranges [25] **F:** 25-30 = 2.0, >30-35 = 3.9, ≥ 35= 6.3 **M:** 25-30 = 2.2, >30-35 = 3.1, ≥ 35= 3.3	See incidence table	

Nonalcoholic fatty liver disease	Odds ratios for BMI Ranges [116] **F:** <18.5= 1.16, 25-29.9 = 2.1, >30-34.9 = 2.7, 35-39.9 = 3.92 , ≥ 40= 5.32 **M:** <18.5= 0.81, 25-29.9 = 2.1, >30-34.9 = 3.7, 35-39.9 = 5.0 , ≥ 40= 6.7	Incidence Rate = 0.02 [34]	
Other problems			
Amputation		1 M * [d]	Body weight has indirect effect through diabetes
Blindness (diabetic retinopathy)		4.1 M * [e]	

Note: The risk ratios summarized in this table reflect risk for people who are obese relative to risk for people who are normal weight. These ratios are based on BMI measures. The microsimulation model will use the underlying equations from the published literature to predict annual probability of disease onset—taking into account multiple risk factors including the presence of other disease.. [a] CDC prevalence estimate (CDC Fact Sheet). [b] CDC estimate of annual incidence (CDC Fact Sheet) http://www.cdc.gov/nchs/fastats/pneumonia.htm [c] Estimated 2014 cancer incidence http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-041770.pdf. [d] Estimated prevalence caused by diabetes and peripheral arterial disease. http://www.amputee-coalition.org/limb-loss-resource-center/resources-by-topic/limb-loss-statistics/limb-loss-statistics/index.html. [e] Estimated prevalence among adults age 40+ with diabetes. http://www.nei.nih.gov/eyedata/pbd_tables.asp. [f] New cases per year: http://www.cancer.org/cancer/thyroidcancer/detailedguide/thyroid-cancer-key-statistics; Costs were substantial and totaled $60,196 per patient during the first year and $35,189 during the second year of follow-up. http://www.jcancer.org/v02p0193.htm. [g] http://www.sleepdisordersguide.com/sleepapnea/sleep-apnea-statistics.html. [h] http://www.cdc.gov/nchs/data/series/sr_10/sr10_249.pdf . [i] The Massachusetts study data suggest there will be approximately 17,781 new cases of ED in Massachusetts and 617,715 in the United States annually. http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/endocrinology/erectile-dysfunction/; Of 18 to 64-year-old males with employer provided insurance average annual expenditures were $4,813 for those treated for ED compared with $3,706 for similar men not treated for the condition: http://www.udaonline.net/publications/file/view/2007/ED_Wessells2007.pdf. [j] http://www.cdc.gov/nchs/data/nhsr/nhsr067.pdf.

In the initial literature search we identified numerous health conditions associated with excessive body weight, though not all those health conditions were included in the final list (shown above). We determined which comorbidities to model according to the following criteria:

There is a general consensus that body weight is a direct risk factor for certain comorbidities, and indirect risk factor for other comorbidities. Indications of this consensus include government and national associations (e.g., American Heart Association, American Diabetes Association) listing the disease as a comorbidity of obesity. All the comorbidities of obesity covered in the model are listed by CDC or other respected organizations as established comorbidities of obesity.

The expected magnitude of the economic impact of reducing body weight is sufficiently large to warrant inclusion in the model. Expected economic impact is a function of three factors (where magnitude is a qualitative assessment of whether a factor will have a sufficiently large impact on medical costs):

The link between body weight and the comorbidity (as measured by rate ratios or some other metric) is large and statistically significant (at the $P<0.05$ level).

The incidence of the comorbidity is large.

The comorbidity is expensive to treat or has a large detrimental effect on mortality, workforce participation, or quality of life.

At least one of these three factors needed to be sufficiently large to include in the model. Still, some high-cost cancers in the model have low incidence so their impact on obesity-attributed costs are modest.

There is sufficient data to model the comorbidity. This criterion is closely associated with the previous two criteria—there is more published literature on the impact of body weight for comorbidities where there is general consensus that body weight is a risk factor, and the comorbidity is prevalent and/or costly.

Assessment by team's clinical expert, Dr. Leigh Perreault. Dr. Perreault is a physician and clinical researcher at the University of Colorado Denver School of Medicine, Aurora, CO. Her overarching research interest lies in understanding the different pathways by which people develop type 2 diabetes, and developing tailored strategies for diabetes prevention.

Based on the above criteria, recognized obesity-related comorbidities that have strong links to body weight, have high incidence and high cost for treatment, and for which data are available for modeling take priority over comorbidities that meet only some or none of these conditions. While some research suggests a link between obesity and prostate cancer, looking across multiple studies the relationship appears to be inconclusive. For example, one study found that obesity lowers risk for localized prostate cancer, but raises risk for advanced prostate cancer and mortality associated with the cancer.[117] Additional studies, however, note that there is no consensus on the link between body weight and prostate cancer.[118;119]

Another example of an obesity-related comorbidity that we added to the OPEM is breast cancer. The relationship between BMI and breast cancer risk differs according to pre or post-menopausal status, with studies reporting an inverse relationship between increasing BMI and breast cancer risk in premenopausal women, and a corresponding rise in risk with increasing BMI in postmenopausal women with no previous use of hormone replacement therapy.[41] For post-menopausal women, it has been shown that each 1kg/m^2 increase in BMI results in an estimated 3% increase in breast cancer risk.[120] Both the annual incidence (232,670 cases per year in 2014 the US) and the cost of breast cancer are relatively high.[127121] There is a wealth of information on the relationship between patient risk factors and cancer incidence, underlying risk for cancer, and medical expenditures. Therefore, adding this condition was a high priority. Most studies of the relationship between excess body weight and breast cancer risk use BMI in the prediction equation, though several studies have used weight circumference (WC). The OPEM used BMI (rather than WC) because BMI is a more common input to published disease prediction models despite the limitations of using BMI.

To model conditions not directly defined by levels of modeled biometrics, we conducted an extensive review of the literature to identify predictive equations of disease onset based on clinical trials or meta-analyses. For the diabetes population, the UKPDS Outcomes Model was a common source for these predictive equations.

[127] SEER Stat Fact Sheets: Breast Cancer http://seer.cancer.gov/statfacts/html/breast.html

Modeling Medical and Indirect Costs

Direct medical cost

The Medical Expenditure Panel Survey (MEPS) data was analyzed to generate annual direct cost expenditure equations by setting and disease type based on primary diagnosis codes. The regression analysis was conducted with the 2009-2013 pooled MEPS data files. The Full Year Consolidated Data File was merged with the individual visit level files including ambulatory (combined office and outpatient), emergency, inpatient and pharmaceutical to develop cost equations for the various settings. Setting-level cost equations were disaggregated by disease type based on the primary ICD-9 diagnosis codes. Exhibit 124 shows the disease categories modeled at the setting level and the corresponding ICD-9 diagnosis code definitions.

Exhibit 124. Disease and ICD-9 Diagnosis Code Mapping

Disease Type	ICD-9 Diagnosis Code
Infectious	001-139
Endocrine	240-279
Hematology	280-289
Mental	290-319; V40
Neuro	320-389
Circulatory	390-459
Resp	460-519
Digestive	520-579
Nephro	580-629
Women's Health	630-679; V22
Musculoskeletal	680-739; 800-848
Congenital	740-759
Perinatal	760-779
Other	Greater than 780
Neoplasms	140-239

In the final analysis, for each of the above settings cost expenditure equations was derived for 11 broad disease categories.[128] We used separate generalized linear models with gamma distribution and log link to model total annual medical expenditures for people age less than 65 and those ages 65 and over. The dependent variable was total

[128] Due to small sample size issue some of the disease categories such as Neonatal/Perinatal, Congenital Conditions, Other rare conditions and Hematology were grouped as a single "All Other" combined disease category in the final regression model. Neoplasm/Cancer cost was not modeled separately in the final model, but based on external published literature estimates.

annual medical expenditures. Explanatory variables include demographics (age group, sex, race/ethnicity); presence of diabetes, hypertension, congestive heart failure, ischemic heart disease, retinopathy, and end-stage renal disease; history of myocardial infarction, stroke, and various cancers; smoking status; and body weight. Separate regressions were estimated for the obese population (which included BMI as a continuous variable) and for the remaining population that was overweight or normal weight (Exhibit 125). In addition to this baseline cost equation, medical expenditures for many chronic disease conditions were derived from published literature, as detailed in the corresponding disease sections above.

Exhibit 125. Regression Results for Estimating Total Annual Medical Expenditures

	Age less than 65				Age 65 and over			
	Non obese		Obese		Non obese		Obese	
Parameter	GLM Coefficient	Pr > ChiSq	GLM Coefficient	Pr > ChiSq	GLM Coefficient	Pr > ChiSq	GLM Coefficient	Pr > ChiSq
Intercept	6.2076	<.0001	5.9581	<.0001	7.4378	<.0001	6.749	<.0001
Male	-0.2912	<.0001	-0.4563	<.0001	-0.0748	0.0014	-0.2124	<.0001
Age								
35 to 44	0.2707	<.0001	0.1502	<.0001				
45 to 64	0.4911	<.0001	0.4141	<.0001				
65 to 74					-0.2208	<.0001	-0.0149	0.6662
Race/ethnicity								
Black	-0.4475	<.0001	-0.4456	<.0001	-0.0806	0.1235	-0.0681	0.2933
Other race	-0.3503	<.0001	-0.3629	<.0001	-0.0488	0.4001	0.1718	0.1441
Hispanic	-0.4121	<.0001	-0.3543	<.0001	-0.2202	<.0001	-0.1235	0.0256
Insured	0.6703	<.0001	0.7603	<.0001	1.4213	<.0001	1.6855	<.0001
Insured through Medicaid	-0.1544	<.0001	0.3201	<.0001				
Body weight								
Overweight	-0.0169	<.0001			-0.086	0.0002		
Continuous BMI (for Obese adults)			0.0118	<.0001			0.0158	<.0001
Disease presence								
Hypertension	0.3437	<.0001	0.4009	<.0001	0.2468	<.0001	0.0414	0.3281
Diabetes	0.8477	<.0001	0.7068	<.0001	0.3887	<.0001	0.353	<.0001
Congestive heart failure	1.8533	<.0001	1.2012	<.0001	1.0504	<.0001	0.7898	<.0001
Ischemic heart disease	1.1771	<.0001	1.1101	<.0001	0.6487	<.0001	0.5207	<.0001
History of stroke	0.8110	<.0001	0.3661	<.0001	0.4246	<.0001	0.3629	<.0001
History of myocardial infarction	1.1858	<.0001	1.1731	<.0001	0.7013	<.0001	0.6061	<.0001
Retinopathy	0.7041	<.0001	0.6623	0.0006	0.6043	<.0001	0.2479	0.1458
Renal failure	2.7283	<.0001	2.1572	<.0001	1.4409	<.0001	1.6285	<.0001
Interactions with diabetes								
Congestive heart failure	-0.5183	0.1842	-0.3229	0.2001	-0.1852	0.3232	0.1807	0.7166

	Age less than 65				Age 65 and over			
	Non obese		Obese		Non obese		Obese	
Parameter	**GLM Coefficient**	**Pr > ChiSq**	**GLM Coefficient**	**Pr > ChiSq**	**GLM Coefficient**	**Pr > ChiSq**	**GLM Coefficient**	**Pr > ChiSq**
Ischemic heart disease	-0.3496	0.0361	-0.1973	0.1109	-0.1444	0.0515	0.0853	0.5399
History of myocardial infarction	-0.2685	0.2147	-0.3629	0.0261	-0.1292	0.1874	0.1111	0.7992
History of stroke	-0.1942	0.2303	0.0910	0.5154	0.1341	0.0863	0.0912	0.1816
Renal failure	-1.0415	0.0107	-0.8523	0.04	0.4487	0.142	0.3442	0.2593
Fit statistics:								
N	93,411		24,826		10,570		4,374	
Deviance/DF:	43,894.00		37,316		17,264		13,363	
Scaled deviance/DF:	1.297		1.285		1.26		1.23	
Pearson Chi-square/DF:	3,343,301		510,174		34,431		18,761	
Scaled Pearson Chi-square/DF:	15.4		17.6		2.5		1.7	

We used a zero-inflated log-ratio regression to model the allocation of total medical expenditures across cost categories. The regression used data on all adults in MEPS with the same explanatory variables used to model total annual medical expenditures. Following the approach of Faes et al.,[129] separate regressions were estimated for ambulatory, inpatient, emergency, and prescription drug categories. The dependent variable for each regression reflected the log of the ratio of category expenditures to expenditures in the "all other" category (e.g. log of ratio of inpatient expenditures to all other expenditures). There are two components from the regression outcome. One component involved estimating logistic regressions to model the probability an individual incurred expenditures in that category. The second component analyzed the log transformed ratio using ordinary least squares model with the same explanatory variables. Combining the information from 2 components allowed us to calculate the proportion of each person's total annual medical expenditures across the five cost categories. Summary of regression coefficients is shown in Exhibit 127. Similar approach was employed to derive disease-level cost equations within each setting, regression coefficients were summarized in Exhibit 128-Exhibit 131.

Exhibit 126 shows similar health expenditure distribution at medical setting-level from raw MEPS data and predicted cost allocation from log-ratio regression. This speaks to the validity of model predictions.

Exhibit 126 Comparison of MEPS health cost and predicted allocation at medical setting-level

	MEPS							Log-ratio regression						
Settings	Minimum	25th Pctl	Median	Mean	75th Pctl	Maximum	Mean %	Minimum	25th Pctl	Median	Mean	75th Pctl	Maximum	Mean %
ER	$0	$0	$0	$188	$0	$163,576	4%	$0	$3	$46	$221	$175	$81,956	5%
Ambulatory	$0	$0	$182	$1,342	$945	$346,548	31%	$0	$8	$176	$1,324	$1,002	$305,546	31%
Inpatient	$0	$0	$0	$1,307	$0	$557,107	30%	$0	$19	$255	$1,228	$956	$1,678,429	28%
Rx	$0	$0	$33	$999	$609	$2,313,850	23%	$0	$2	$55	$1,094	$503	$189,137	25%
All other	$0	$1	$1	$497	$250	$231,367	11%	$0	$2	$66	$465	$346	$165,344	11%
Total	$0	$49	$757	$4,333	$3,530	$2,313,850	100%	$0	$49	$757	$4,333	$3,530	$2,313,850	100%

[129] 1. Faes C, Molenberghs G, Hens N, Muller A, Goossens H, Coenen S (2011) Analysing the composition of outpatient antibiotic use: a tutorial on compositional data analysis. J Antimicrob Chemother 66 Suppl 6: vi89-vi94. dkr461 [pii];10.1093/jac/dkr461 [doi].

Exhibit 127. Regression Results for Allocation of Medical Expenditures at Medical Settings

	Model parameter estimate							
	Emergency		Inpatient		Ambulatory		Prescription drug	
Coefficients	Log-ratio	Logistic	log-ratio	Logistic	Log-ratio	Logistic	Log-ratio	Logistic
Intercept	1.6257**	-2.7869**	4.0254**	-3.0416**	1.9275**	0.1947**	1.2800**	-0.0984*
Age 18 to 34	1.3526**	0.1956**	1.6755**	-0.1508**	-0.0719	-0.7639**	-0.8794**	-0.9023**
Age 35 to 44	1.5597**	-0.0148**	1.1773**	-0.5261**	-0.1038	-0.6117**	-0.6817**	-0.7868**
Age 45 to 64	1.1497**	-0.3034**	0.8203**	-0.7007**	-0.4016**	-0.4485**	-0.7479**	-0.6136**
Age 65 to 74	0.3993	-0.2114	0.2763	-0.2041**	0.0015	-0.2038**	-0.0627	-0.2011**
Hispanic	0.5988*	0.2485**	0.5374	0.1646**	0.7314**	-0.1382**	0.5004**	-0.0321
Black	0.5629*	0.5302**	0.4120	0.2757**	0.4856**	-0.2011**	0.3389**	0.0070
White	-0.3465	0.4275**	-0.5426	0.3487**	-0.0639	0.5090**	-0.0607	0.6714**
Normal weight	-0.4712**	-0.0373	-0.1855	-0.0011	-0.2907**	-0.0788**	-0.2543**	-0.1248**
Obese	0.1289	0.1420**	-0.0238	0.0669*	0.1379*	0.0651**	0.1919**	0.1345**
Male	0.4178**	-0.2985**	0.2779	-0.6011**	0.0278	-0.8146**	0.1546**	-0.7661**
Smoker	0.9508**	0.4164**	0.9071**	0.0410	0.8251**	-0.2010**	0.8521**	0.0679**
Presence or history of								
Hypertension	-0.1248	0.3125**	-0.2492	0.3404**	0.4362**	0.7276**	0.8444**	1.2579**
Cardiovascular disease	-0.5035**	0.5142**	0.1009	0.6220**	0.0553	0.5937**	0.0622	0.7397**
Heart attack	0.1725	0.1941**	-0.2486	0.5041**	0.3384**	0.0653	0.6114**	0.2733**
Stroke	-0.7694**	0.5802**	-1.0009**	0.6067**	-0.3254**	0.2007**	-0.1359	0.5735**
Diabetes	-0.1345	0.2442**	-0.0440	0.3774**	0.3522**	0.8885**	1.1373	1.7220**
Arthritis	-0.6162**	0.3974**	-0.7568**	0.3753**	0.1743**	0.8168**	0.1664**	0.9026**
Insured	-1.0678**	0.0162	-0.9866**	0.2749**	-1.2608**	1.2191**	-1.3034**	0.8902**
Has Medicaid	-0.2507	0.7074**	0.2682	0.8279**	1.0365**	0.1247**	1.4200**	0.2885**
Fit statistics (N=121,360)								
-2 Log Likelihood	186,711		122,139		654,374		620,176	
Akaike information criterion (AIC)	186,793		122,221		654,456		620,258	
Bayesian information criterion (BIC)	187,191		122,619		654,854		620,656	
Pearson Statistic	120,099		115,370		120,557		123,202	

** significant at 0.05 level

Exhibit 128. Regression Results for Allocation of Disease-level Medical Expenditures from Emergency Care

Emergency	Logistic										
Coefficients	**Circulatory**	**Neuro**	**Resp**	**Infectious**	**Endocrine**	**Digestive**	**Nephro**	**Mental**	**Musculo**	**WomensHealth**	**Neoplasms**
Intercept	-5.68	-6.4282	-4.7798	-6.1154	-6.1606	-4.9789	-5.5127	-6.4343	-4.2968	-10.4787	-6.8068
Age 18 to 34	-0.6789	1.1123	-0.00301	0.9492	-0.1894	0.07273	0.5919	0.4471	0.2665	6.4656	-1.7194
Age 35 to 44	-0.2119	0.9400	-0.2338	0.4436	-0.2779	-0.02335	0.3775	0.5593	0.1162	5.4568	-0.7648
Age 45 to 64	-0.0975	0.4345	-0.4798	0.1304	-0.4263	-0.3431	-0.3378	-0.1048	-0.1297	0.4373	-0.5478
Age 65 to 74	-0.07275	0.3748	-0.18	0.05011	-0.2991	0.01653	-0.1248	-0.09885	-0.1428	-4.6902	-0.3181
Hispanic	0.1776	0.1853	0.1166	0.00769	-0.03242	0.4360	0.4680	0.4181	0.1584	0.2729	0.7248
Black	0.5359	0.2873	0.5453	0.5378	0.7963	0.4532	0.4359	0.2104	0.5057	0.3295	1.0271
White	0.1643	0.4034	0.3561	0.4469	0.3950	0.4950	0.7366	0.5027	0.4370	0.05765	0.6824
Normal weight	0.01432	-0.171	0.1035	0.06051	0.06682	-0.09578	-0.1258	0.02113	-0.07174	-0.2715	0.1345
Obese	0.1092	0.2908	0.4135	0.2590	0.1018	0.07271	0.1241	-0.03407	0.1046	-0.01579	0.1640
Male	-0.07925	-0.6582	-0.4277	-0.3472	-0.07953	-0.3219	-0.7792	-0.6012	-0.1612	0	-0.2351
Smoker	0.1221	0.5049	0.5663	0.2723	0.1870	0.4942	0.3635	0.8496	0.5343	0.0615	-0.06535
Presence or history of											
Hypertension	1.2260	0.3792	0.3476	0.1612	0.4274	0.2290	0.2101	0.2334	0.1399	0.02216	0.1385
Cardiovascular disease	1.2531	0.3349	0.5062	0.2074	0.4612	0.3513	0.3613	0.2384	0.2244	0.2464	0.6761
Heart attack	0.7190	-0.3111	0.1484	-0.1731	-0.3006	-0.03428	0.06504	0.1155	-0.06849	0.008369	0.06056
Stroke	1.1242	0.5268	0.2391	0.1037	0.2007	0.2132	0.2578	0.2646	0.3239	0.1229	0.6180
Diabetes	-0.0268	-0.246	-0.05397	0.1716	2.1502	0.1708	0.3957	-0.0161	0.1383	-0.7547	0.1573
Arthritis	0.06651	0.4002	0.3117	0.3123	0.1410	0.4372	0.2182	0.5844	0.6910	-0.1109	-0.1945
Insured	-0.1095	-0.09985	-0.3121	-0.09551	-0.5026	-0.2427	0.004079	-0.5168	-0.2095	-0.5638	-0.2321
Has Medicaid	0.5399	0.6820	0.8923	0.7510	0.6207	0.7666	0.4892	1.1710	0.6599	1.4677	0.5540
	Log-ratio										
Coefficients	**Circulatory**	**Neuro**	**Resp**	**Infectious**	**Endocrine**	**Digestive**	**Nephro**	**Mental**	**Musculo**	**WomensHealth**	**Neoplasms**
Intercept	5.9595	5.4846	5.4799	5.4863	5.1293	4.6206	6.1207	4.9946	5.1433	5.2927	5.8803
Age 18 to 34	0.06299	0.7405	0.3789	0.7676	0.2391	0.9621	0.3107	0.4215	0.4779	0.8485	1.2155
Age 35 to 44	0.1398	0.6411	0.7055	0.9497	0.008982	1.2587	0.2520	0.8946	0.6236	1.0865	1.3878
Age 45 to 64	0.3722	0.9993	0.7509	0.9998	0.4056	1.2879	0.6650	0.8792	0.7369	2.6027	1.1018
Age 65 to 74	-0.03603	-0.2316	0.2196	0.3300	-0.697	0.6167	-0.2239	0.4597	0.3301	-6.1498	0.2690
Hispanic	0.1014	-0.1402	-0.1501	-0.4339	0.5522	0.4270	-0.08934	-0.2686	0.03365	-0.7114	-1.5322
Black	-0.374	-0.09401	-0.4438	-0.4999	0.5306	0.4815	-0.4273	-0.2439	-0.08164	-0.3295	-1.417
White	-0.2576	-0.2887	-0.5187	-0.3731	0.09456	0.4453	-0.3898	-0.5456	-0.04031	-0.2118	-1.1154
Normal weight	-0.1765	-0.2134	-0.1049	-0.1675	-0.2614	-0.00563	-0.3796	-0.1664	-0.147	0.1551	1.4662
Obese	-0.2041	-0.07102	-0.1566	-0.2734	-0.3116	-0.06135	0.1460	0.2317	-0.2273	-0.1842	1.4174
Male	-0.0373	0.1008	0.2404	-0.04592	0.2912	0.2435	0.5576	0.4113	0.2051	0	-0.6805
Smoker	0.03758	-0.1073	0.04713	-0.6973	0.0987	-0.294	0.1312	0.1583	-0.1619	-0.2448	-0.3141
Presence or history of											
Hypertension	-0.06087	-0.22	-0.00365	-0.105	-0.3046	-0.1454	-0.4458	-0.9991	-0.204	-0.3961	0.1475
Cardiovascular disease	0.0825	-0.1891	-0.1576	-0.03029	-0.6096	-0.4564	-0.4713	0.1354	-0.1451	-0.4093	0.09111
Heart attack	0.1566	-1	0.01844	0.3081	0.8900	0.2938	-0.134	0.9430	-0.1283	0.9958	-0.9284
Stroke	0.07579	-0.6916	-0.1229	-0.3678	-0.4347	-0.6114	-0.1745	-0.7656	-0.04656	-0.3876	-1.6398
Diabetes	-0.08461	0.0043	-0.07738	-0.08676	-0.06519	0.1705	-0.1166	0.1575	-0.08396	-0.4741	0.3783
Arthritis	-0.09506	0.1657	-0.31	-0.136	-0.3093	-0.3613	0.007589	-0.2407	0.002664	-0.7189	0.4511
Insured	0.2865	0.4155	0.3546	0.5828	1.1151	0.5415	0.3829	1.0180	0.6108	0.7186	-0.3617
Has Medicaid	-0.6546	-1.161	-0.7007	-0.7427	-1.1213	-1.3553	-1.0756	-1.2251	-0.9836	-1.1346	-0.5379

Exhibit 129. Regression Results for Allocation of Disease-level Medical Expenditures from Ambulatory Care

Ambulatory	Logistic										
Coefficients	Circulatory	Neuro	Resp	Infectious	Endocrine	Digestive	Nephro	Mental	Musculo	pmensHea	Neoplasms
Intercept	-3.1522	-3.6153	-3.4103	-3.8172	-15.3879	-2.3064	-2.8548	-3.5078	-4.8377	-2.6052	-3.6153
Age 18 to 34	-1.8602	-2.2462	-0.3425	-0.5115	13.2045	-1.3031	-1.2977	0.3332	1.0209	-0.2406	-2.2462
Age 35 to 44	-1.1573	-1.4517	-0.1023	-0.07999	12.163	-1.1669	-0.401	0.4372	1.1569	0.1371	-1.4517
Age 45 to 64	-0.6477	-0.8185	-0.17	-0.0026	7.1373	-0.8833	-0.03275	0.3786	0.8576	0.2834	-0.8185
Age 65 to 74	-0.2885	-0.1247	-0.02492	0.07138	-6.8749	-0.3689	0.05368	0.1520	0.3441	0.1411	-0.1247
Hispanic	-0.05342	0.1636	0.2047	0.1975	-0.0712	-0.1463	-0.09312	-0.0326	-0.04649	-0.1149	0.1636
Black	-0.01269	0.3957	-0.0328	-0.1881	-0.3032	-0.2196	-0.4174	-0.2487	-0.2618	-0.2025	0.3957
White	-0.00748	1.1742	0.4598	0.3676	0.1064	0.4875	0.1757	0.5020	0.6766	0.4487	1.1742
Normal weight	-0.1276	0.09689	-0.0057	-0.03941	-0.3532	-0.03849	-0.1874	-0.04166	0.000405	-0.05447	0.09689
Obese	0.1345	-0.1243	-0.04167	-0.03679	-0.03239	0.03223	0.1433	0.1698	0.1071	0.08466	-0.1243
Male	-0.01621	-0.08715	-0.8592	-0.3003	-7.3467	-0.3632	-0.2787	-0.468	-0.4017	0	-0.08715
Smoker	-0.0548	-0.2499	-0.1249	-0.1169	-0.4003	-0.1876	-0.1661	-0.00298	0.5317	-0.07077	-0.2499
Presence or history of											
Hypertension	2.5497	0.08266	0.1498	0.1141	-0.2016	0.1031	-0.2091	0.05436	0.1782	0.02773	0.08266
Cardiovascular disease	0.9989	0.2692	0.4251	0.3486	-0.03063	0.1645	0.1365	0.3252	0.3019	0.2140	0.2692
Heart attack	0.5337	-0.1516	0.04571	-0.2281	-0.4809	-0.1934	-0.1079	-0.1296	-0.3491	-0.2721	-0.1516
Stroke	0.4572	-0.01333	0.04333	-0.02514	-0.3224	0.1661	0.0403	0.02143	0.1964	-0.1152	-0.01333
Diabetes	0.3238	0.02428	0.2008	-0.03019	-0.2738	0.1355	2.4307	-0.2477	-0.02237	-0.2484	0.02428
Arthritis	0.1928	0.2790	0.2911	0.5029	-0.2846	0.4229	0.3199	0.4117	0.5241	1.2410	0.2790
Insured	0.3685	0.8502	0.6058	0.6255	0.3912	0.7367	0.6886	0.9576	0.2682	0.8001	0.8502
Has Medicaid	0.3336	0.2119	0.1044	0.2350	0.8270	0.1435	0.07065	-0.1796	1.0109	0.07396	0.2119
	Log-ratio										
Coefficients	Circulatory	Neuro	Resp	Infectious	Endocrine	Digestive	Nephro	Mental	Musculo	pmensHea	Neoplasms
Intercept	0.7732	1.7920	1.0920	1.3381	2.4263	1.4820	1.3089	0.7918	2.3333	0.7996	2.1920
Age 18 to 34	0.5624	0.9227	1.4216	1.4684	1.3161	0.7543	1.2986	1.2576	1.3288	0.8476	0.9227
Age 35 to 44	0.7341	0.7948	1.2148	1.4421	1.1201	0.5484	0.7628	1.0307	1.1672	0.6918	0.7948
Age 45 to 64	0.2045	0.5645	0.5554	0.7287	-0.07255	0.1079	0.3659	0.5922	0.4836	0.3404	0.5645
Age 65 to 74	0.01659	0.1355	-0.00129	0.1679	-2.8762	-0.328	-0.05345	0.0371	0.1922	-0.03227	0.1355
Hispanic	0.2871	0.3341	0.1987	0.07832	0.01038	0.4636	0.3793	0.2862	0.02322	0.1311	0.3341
Black	-0.05289	0.2545	0.6648	0.0471	0.001504	0.3410	-0.06276	0.1212	0.08689	0.0622	0.2545
White	-0.6082	-0.7685	-0.4912	-0.1664	-0.2411	-0.1522	-0.4287	-0.3032	-0.3552	-0.3541	-0.7685
Normal weight	0.1112	0.1317	-0.00864	-0.1238	-0.06745	-0.08535	-0.00405	0.09153	0.04828	-0.02999	0.1317
Obese	0.05824	0.01564	-0.1662	-0.00408	-0.187	-0.05223	-0.0528	-0.02392	-0.04382	-0.1118	0.01564
Male	0.6384	0.7032	0.7190	1.1288	0.5576	0.8928	0.7098	1.0834	0.9188	0	0.7032
Smoker	0.4907	0.7515	0.5754	0.5783	-0.1539	0.5078	0.3711	0.6748	0.6575	0.6176	0.7515
Presence or history of											
Hypertension	0.1952	-0.07888	0.1238	-0.1569	-0.2991	0.06907	-0.2423	0.005021	-0.09242	-0.09874	-0.07888
Cardiovascular disease	-0.00978	-0.2546	-0.1103	-0.4993	-0.567	-0.4535	-0.3118	-0.3391	-0.266	-0.3174	-0.2546
Heart attack	0.2064	-0.05059	-0.1095	0.2040	1.4221	-0.0655	0.2159	0.05179	0.03553	0.01069	-0.05059
Stroke	0.3884	0.2073	0.2014	-0.4303	-0.9571	0.1964	0.08381	0.1535	0.3372	-0.09027	0.2073
Diabetes	0.2200	0.2445	0.4155	-0.04586	0.03562	0.3015	0.9214	-0.095	-0.1488	0.04291	0.2445
Arthritis	-0.2523	-0.3494	-0.2802	-0.3511	-0.4471	-0.2589	-0.3731	-0.2211	-0.3304	0.1243	-0.3494
Insured	-0.9195	-0.8944	-0.8197	-0.9886	-0.446	-1.0363	-0.8852	-0.9055	-0.9512	-1.0328	-0.8944
Has Medicaid	0.3642	0.3351	0.3658	-0.03359	0.4907	-0.00385	-0.01992	-0.04917	0.6416	0.3754	0.3351

Exhibit 130. Regression Results of Disease-level Allocation of Medical Expenditures from Inpatient Care

Inpatient	Logistic										
Coefficients	**Circulatory**	**Neuro**	**Resp**	**Infectious**	**Endocrine**	**Digestive**	**Nephro**	**Mental**	**Musculo**	**WomensHealth**	**Neoplasms**
Intercept	-5.3214	-6.6257	-5.9361	-7.0428	-7.6436	-5.0783	-6.4516	-7.1816	-5.6835	-17.5623	-5.9822
Age 18 to 34	-1.7903	-1.2707	-1.6696	-0.5182	-0.1372	-0.9075	-0.6184	-0.3626	-1.2674	14.8285	-1.8811
Age 35 to 44	-0.9368	-1.1302	-1.4219	-0.4212	-0.01725	-0.5587	-0.1305	-0.1306	-0.8163	13.7389	-0.5996
Age 45 to 64	-0.5566	-1.1068	-1.1499	-0.71	-0.3517	-0.5739	-0.5226	-0.6523	-0.6469	8.6432	-0.2788
Age 65 to 74	-0.1395	-0.517	-0.2979	-0.1288	-0.1786	-0.09529	-0.1152	-0.6047	-0.2562	-4.8163	0.3029
Hispanic	-0.1214	-0.07491	0.1226	0.1635	0.6310	0.1993	0.3240	0.5522	-0.03704	0.03975	-0.06672
Black	0.3959	-0.08046	0.4109	0.3408	1.2390	0.0542	0.2945	0.6511	0.1030	-0.05751	0.4159
White	0.1165	0.1123	0.5404	0.7039	0.9445	0.2268	0.5153	0.8531	0.4736	0.1373	0.2880
Normal weight	-0.0276	-0.3478	0.4395	0.04843	0.1704	-0.0712	0.08458	0.1556	0.1573	-0.2162	0.1526
Obese	0.1425	-0.1364	0.3547	0.1566	-0.1343	0.03498	0.1703	0.02494	0.1472	-0.07743	0.04445
Male	0.04838	0.1879	-0.126	-0.01919	-0.2163	-0.2489	-0.7754	-0.2117	0.1216	0	-0.1451
Smoker	0.08936	0.2675	0.3950	0.1843	0.01987	0.1540	0.2033	1.1450	-0.00994	-0.3479	-0.08258
Presence or history of											
Hypertension	0.8789	0.1830	0.4474	0.4515	0.5182	0.2221	0.3597	0.5252	0.2621	-0.2396	0.2510
Cardiovascular disease	1.5605	-0.1386	0.6062	0.5197	0.5756	0.2654	0.4163	0.2020	0.1549	-0.05271	0.005331
Heart attack	0.7585	0.3180	0.1833	0.2625	-0.4434	-0.00431	0.05374	0.07975	-0.0129	-0.8545	0.1936
Stroke	1.0566	0.5776	0.3744	0.2583	0.3027	0.1362	0.5817	0.6619	0.2525	-0.2294	0.2685
Diabetes	0.1978	0.4986	0.4112	0.3527	2.2773	0.1892	0.4717	-0.1167	0.3069	-0.2413	0.3453
Arthritis	0.03365	0.5229	0.4215	0.4766	0.1957	0.4706	0.1755	0.4647	1.2823	-0.1308	0.1542
Insured	-0.0068	0.3469	0.1114	-0.2188	-0.05699	0.1996	0.5692	-0.3209	0.2953	0.1508	0.4218
Has Medicaid	0.6126	1.0749	1.2197	0.8978	0.5995	0.6275	0.5800	1.4559	0.6204	0.8920	0.4130
	Log-ratio										
Coefficients	**Circulatory**	**Neuro**	**Resp**	**Infectious**	**Endocrine**	**Digestive**	**Nephro**	**Mental**	**Musculo**	**WomensHealth**	**Neoplasms**
Intercept	9.2748	8.4662	8.5063	8.4911	7.4307	7.8350	7.3513	7.4387	8.6324	5.4968	8.8789
Age 18 to 34	0.1074	0.7749	-0.1698	0.1885	0.6670	0.5356	1.2583	0.2338	0.4162	2.6210	0.9263
Age 35 to 44	0.3881	0.1067	0.1580	0.2178	1.1889	0.6649	1.4155	0.2584	0.3376	2.7059	0.9359
Age 45 to 64	0.2025	0.6470	0.2402	0.4636	1.1026	0.4903	1.1694	0.2992	0.4967	1.3638	0.7174
Age 65 to 74	-0.1427	0.2227	0.3115	0.1100	-0.1589	0.4803	0.7663	1.4201	-0.546	-7.7689	0.9057
Hispanic	-0.4504	1.7739	-0.2283	-0.5651	-0.9244	0.004678	-0.3261	-0.2631	-0.2989	-0.04105	-0.9835
Black	-0.2622	0.3327	-0.06911	-0.4677	-1.4388	-0.3442	-0.03884	0.6950	-0.5897	-0.02513	-0.923
White	-0.2948	0.9253	-0.07034	-0.7657	-1.0336	-0.1669	-0.1218	0.7123	-0.2778	-0.0257	-0.4558
Normal weight	-0.1795	-0.1292	0.004393	-1.1994	0.4303	-0.09212	-0.3697	0.4564	-0.1085	0.001451	0.6169
Obese	-0.2183	0.4249	0.3049	-0.5534	0.04949	-0.1297	-0.01967	0.1266	-0.1778	0.007334	0.1297
Male	0.4837	-0.5815	0.05887	1.0599	0.03306	0.01439	0.2655	0.05227	0.1437	0	-0.06369
Smoker	0.1184	0.2336	-0.5203	-0.4045	0.2245	0.2516	0.3451	-0.4786	-0.05499	-0.05059	-0.0299
Presence or history of											
Hypertension	-0.1716	-0.5158	-0.7062	-0.01872	0.4341	-0.06317	0.2480	-0.3986	0.1625	-0.01226	-0.213
Cardiovascular disease	0.1864	-1.2982	-0.4418	-0.3336	-0.2735	0.2506	-0.3275	0.1316	0.1722	-0.1371	-0.00609
Heart attack	0.2293	1.6675	0.2667	0.2234	0.7985	-1.1722	-0.324	0.2369	-0.7613	0.3548	-0.00958
Stroke	0.04884	-1.6376	-0.4237	-2.8567	-0.8816	-0.3081	0.7367	-0.4072	-0.2803	-0.4411	-1.3418
Diabetes	-0.1514	0.01611	0.2370	-0.01132	0.001387	-0.2628	-0.2286	0.7625	-0.3639	-0.4936	0.03955
Arthritis	-0.3438	0.5935	-0.2506	-0.09916	-0.3321	0.08011	0.2945	-0.1598	0.09765	-0.08284	-0.05903
Insured	0.1577	-0.4465	0.7361	0.4425	1.2912	0.5743	0.2815	-0.08961	0.5892	0.4718	0.06133
Has Medicaid	-0.5628	-1.6791	-0.5443	-0.3139	-0.5406	-0.07288	-0.9855	0.3095	-0.5463	-0.4483	0.1191

Exhibit 131. Regression Results of Disease-level Allocation of Medical Expenditures from Prescription Drug

RX	Logistic										
Coefficients	**Circulatory**	**Neuro**	**Resp**	**Infectious**	**Endocrine**	**Digestive**	**Nephro**	**Mental**	**Musculo**	**WomensHealth**	**Neoplasms**
Intercept	-2.5897	-2.7908	-3.1132	-4.5026	-1.5692	-2.7755	-3.0489	-3.7562	-2.363	-18.3917	-4.5835
Age 18 to 34	-2.8287	-1.0329	0.2545	0.7301	-2.3411	-0.8575	-0.7143	0.1394	-0.5313	15.5302	-2.2925
Age 35 to 44	-1.9346	-0.7918	0.4132	0.7160	-1.3748	-0.4515	-0.6572	0.3973	-0.2499	14.6087	-1.3439
Age 45 to 64	-1.0985	-0.6355	0.3790	0.5220	-0.6048	-0.2546	-0.4381	0.3357	-0.102	9.6788	-0.6255
Age 65 to 74	-0.3863	-0.2902	0.2561	0.3302	-0.01733	-0.02077	-0.1119	0.2130	0.01419	-7.4989	0.005244
Hispanic	-0.2709	0.005368	-0.09314	0.1704	-0.3294	0.2175	0.2577	0.06483	-0.0131	-0.1761	0.3634
Black	0.1600	0.07457	-0.05135	0.2133	-0.5263	0.1080	0.1489	-0.1936	-0.04484	-0.2333	0.5525
White	0.1356	0.5764	0.4407	0.5518	0.1836	0.4765	0.6260	1.0566	0.2593	0.005474	0.8762
Normal weight	-0.2372	-0.02271	-0.1128	-0.01357	-0.3491	-0.1408	-0.02164	-0.07377	-0.01567	-0.3078	0.07514
Obese	0.2236	0.08006	0.2082	0.05327	0.1725	0.1306	-0.02788	0.1503	0.1598	0.07484	-0.02906
Male	-0.1428	-0.4876	-0.4673	-0.4772	-0.2407	-0.3306	-0.5692	-0.6065	-0.4023	0	-0.1544
Smoker	-0.1974	-0.00801	0.0806	0.05588	-0.1808	0.1213	-0.07548	0.5770	0.2369	-0.2781	-0.1452
Presence or history of											
Hypertension	3.6403	0.1847	0.2145	0.03467	0.8991	0.3682	0.1146	0.3866	0.2889	-0.1504	0.07168
Cardiovascular disease	1.1593	0.1423	0.3802	0.2492	0.4911	0.4108	0.3313	0.3486	0.1780	0.1769	0.1292
Heart attack	0.7005	-0.1356	-0.01204	-0.1649	0.5348	0.07575	0.09788	-0.07595	-0.1597	-0.6104	-0.00375
Stroke	0.5798	0.3348	0.1173	0.07496	0.3062	0.04589	0.04361	0.4219	0.07124	-0.5572	0.1199
Diabetes	0.6323	0.2034	-0.0503	0.0374	3.1813	0.1909	0.2827	0.2175	0.03241	-0.2655	0.1508
Arthritis	0.1931	0.4953	0.5267	0.3797	0.3665	0.6766	0.3542	0.6348	1.3927	-0.1806	0.2778
Insured	0.5732	0.5889	0.7957	0.6418	0.7341	0.3631	0.4603	0.3607	0.4623	0.3163	0.6469
Has Medicaid	0.09344	0.4341	0.1260	0.1258	0.05839	0.5637	0.2524	0.8782	0.5066	0.9136	0.4433
	Log-ratio										
Coefficients	**Circulatory**	**Neuro**	**Resp**	**Infectious**	**Endocrine**	**Digestive**	**Nephro**	**Mental**	**Musculo**	**WomensHealth**	**Neoplasms**
Intercept	1.3277	2.8738	2.2597	1.4955	1.4728	1.7074	2.2999	2.2279	1.8366	2.3469	2.7458
Age 18 to 34	0.1072	-0.494	0.3566	0.3745	0.5737	0.4083	0.08354	1.0938	0.7315	0.3841	0.4691
Age 35 to 44	0.5697	0.1168	0.7632	0.9586	0.9158	0.9803	0.4558	1.1349	0.8475	0.4418	0.7778
Age 45 to 64	0.7173	0.0909	0.7812	1.0341	1.0130	0.9492	0.6979	0.8128	0.7582	0.8442	0.9106
Age 65 to 74	0.2513	-0.00026	0.3968	0.5030	0.4954	0.4421	0.5321	0.1478	0.4435	-2.0876	0.5382
Hispanic	-0.3563	-0.07272	-0.1822	0.04173	-0.09374	0.04924	-0.1528	0.01294	-0.3634	-0.3407	-0.2095
Black	0.05636	0.2534	0.07911	0.2244	0.03085	0.1806	-0.01506	0.1128	-0.3353	-0.2784	-0.3206
White	-0.349	-0.1025	-0.1234	-0.3569	-0.09972	0.1894	-0.3214	0.1055	-0.2107	-0.6561	-0.4662
Normal weight	-0.2901	0.07692	-0.09015	-0.04602	-0.2317	-0.2038	-0.1164	-0.0928	0.1004	-0.1383	-0.1966
Obese	-0.02017	-0.1804	-0.01263	-0.182	-0.1519	-0.04371	-0.2414	0.08958	-0.04054	-0.1327	-0.2248
Male	0.2369	0.4940	0.6121	1.0098	0.4421	0.3022	0.8562	0.4997	0.2055	0	0.1046
Smoker	-0.2834	-0.2382	-0.04794	0.2502	-0.2533	-0.1548	-0.2208	0.0937	-0.04828	-0.03944	-0.1404
Presence or history of											
Hypertension	0.7257	-0.2775	-0.1685	-0.3777	-0.07055	-0.2222	-0.2672	-0.4043	-0.2526	-0.2813	-0.4603
Cardiovascular disease	-0.1224	-0.6193	-0.1957	-0.4298	-0.3968	-0.4145	-0.4787	-0.4929	-0.415	-0.1229	-0.5509
Heart attack	0.3526	0.04959	-0.358	-0.00349	0.1279	-0.3097	-0.06651	-0.2808	-0.3582	0.9269	-0.5184
Stroke	-0.3718	-0.5989	-0.5955	-0.5645	-0.6923	-0.4802	-0.9463	-0.5285	-0.7531	-0.1508	-0.7264
Diabetes	-0.3745	-0.64	-0.6291	-0.5776	1.0112	-0.502	-0.5526	-0.4651	-0.5685	-0.6324	-0.4243
Arthritis	-0.4725	-0.3957	-0.2302	-0.6421	-0.4469	-0.191	-0.4023	-0.4743	0.2446	0.04782	-0.4742
Insured	0.0973	-0.1041	-0.1374	-0.353	0.2550	0.2274	-0.3188	-0.1231	-0.2797	0.01496	0.1636
Has Medicaid	-1.2154	-0.4265	-0.4896	-0.2163	-1.1133	-0.9849	-0.5865	-0.2219	-0.5184	-0.2865	-0.925

Indirect Cost

Studies have established a relationship between presence of chronic disease and lower productivity in the form of reduced employment, reduced earnings, absenteeism (missed work days), and presentism (reduced productivity while at work). To develop prediction equations for the microsimulation model, we conducted regression analysis with MEPS data on years of employment (logistic regression), household and personal income (OLS regression), missed work days (negative binomial regression), and receipt of Supplemental Security Insurance for disability (logistic regression). As shown in Exhibit 132, presence of obesity-related comorbidities contributed to lower employment, lower earnings, increased absenteeism, and increased likelihood of receiving disability payments.

The recent revisions to the indirect cost regression include updating the missed work days cost equation. The missed work days cost equation was updated to incorporate the recent MEPS data from 2009-2013. The regression analysis was revised to include several new explanatory variables including disease conditions COPD, Osteoporosis, Bipolar Disorder, Alzheimer's disease, Depression, Cancer, Asthma and smoker as a health risk indicator.

Exhibit 132. Regression Results for Indirect Costs

Variable	Employment Probability		Household Income ($)		Missed Work Days		SSI Disability Probability	
	Odds Ratio	**95% CI**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Odds Ratio**	**95% CI**
Intercept			73,647**	747	0.8378**	0.0327**		
Male	1.681*	1.625-1.739	6,444**	476	-0.0994**	0.0195**	0.792*	0.716-0.876
Age								
35 to 44	1.636*	1.553-1.723	15,958**	650	0.0476	0.0269	0.912	0.766-1.087
45 to 64	1.1*	1.054-1.148	24,726**	593	0.1178**	0.0245	1.277*	1.104-1.478
65 to 74	0.182*	0.172-0.194	23,405**	1,242	0.0108	0.0524	1.013	0.838-1.223
Race/ethnicity								
Black	0.721*	0.69-0.754	(23,175) **	664	0.3488**	0.0342	2.724*	2.421-3.064
Other race	0.844*	0.794-0.898	7,880**	865	-0.0838**	0.045	1.162	0.942-1.433
Hispanic	0.69*	0.661-0.719	(22,931) **	604	0.1655**	0.0238	1.942*	1.704-2.213
Health Risk Factors								
Overweight	1.282*	1.229-1.337	(3,833) **	594	0.0159	0.0244	0.738*	0.646-0.844
Obese	1.159*	1.11-1.21	(7,673) **	629	0.1776**	0.0252	0.895	0.788-1.015
Smoker					0.2221**	0.0262		
Disease presence								
Hypertension	0.741*	0.713-0.771	(5,461) **	584	0.1967**	0.0231	1.627*	1.447-1.829
History of stroke	0.392*	0.357-0.43	(15,077) **	1,932	0.5306**	0.0804	2.109*	1.794-2.48
Diabetes	0.627*	0.595-0.66	(6,766) **	912	0.1829**	0.0367	1.647*	1.458-1.86
IHD	0.517	0.469-0.57	1,679	1,994	0.4436**	0.0771	1.735*	1.448-2.079
History of MI	0.683*	0.6-0.778	(3,559)	2,592	0.3879**	0.1023	1.43*	1.142-1.79
CHF	0.359*	0.286-0.45	(889)	5,107	0.4987*	0.207	1.589*	1.174-2.149
Retinopathy	0.397*	0.312-0.506	(10,231)*	4,747	0.0798	0.1943	3.623*	2.488-5.277
Chronic Renal Disease	0.497	0.209-1.182	(4,337)	17,988	1.5661**	0.4923	0.75	0.165-3.405
Renal Failure	0.247*	0.179-0.343	5,665	7,555	1.254**	0.2987	1.626*	1.075-2.458
Pulmonary Embolism	0.731	0.434-1.232	(6,292)	9,117	1.0909**	0.4464	0.689	0.158-3.002
Osteo-arthritis	0.557*	0.518-0.598	$201	1,362	0.4936**	0.0509	1.962*	1.69-2.279
Back Pain	0.824*	0.779-0.871	(1,930)*	847	0.3415**	0.0334	1.344*	1.157-1.561
Gallstone	0.811*	0.665-0.988	(215)	3,188	1.0038**	0.1231	0.641	0.323-1.274
GERD	0.646*	0.607-0.688	2,573 *	1,067	0.3142**	0.0483	1.739*	1.511-2.002
Liver Disease	0.475*	0.337-0.67	5,472	6,408	0.3489	0.2531	1.239	0.536-2.865
Pneumonia	0.604*	0.533-0.685	(6,158)*	2,183	0.8181**	0.0844	1.943*	1.501-2.516
Cancer					0.616**	0.0405		
Asthma					0.293**	0.0329		
COPD					0.34**	0.0461		

Osteoporosis			0.3567**	0.0812
Bipolar			0.7256**	0.1033
Alzheimer			0.342	0.9495
Depression			0.5065**	0.0357
Schizophrenia			0.5641	0.3743
	n=83,931 % Concordant=72.6% % Discordant=26.9% c-statistic=0.729	n=59,763 F value=185.68 R-squared=0.0853	n=50,728	n=67,907 Concordant=72.2% % Discordant=24.7% c-statistic=0.738

** Statistically significant at the 0.01 level. * Statistically significant at the 0.05 level. CHF=congestive heart failure. MI=myocardial infarction. IHD=ischemic heart disease.

An individual was designated as employed or unemployed in that year by comparing his or her probability of employment to a random number generated from a uniform distribution between 0 and 1. That is, if 100 adults each had a calculated 60 percent probability of employment based on individual characteristics then the simulation model would have classified approximately 60 of them as employed and 40 as unemployed. The same methodology was used to predict whether or not an individual received SSI payments.

For people who were modeled as unemployed, their productivity loss associated with reduced employment was calculated using estimates of median personal income by age and sex.[130] (Probability of unemployment took into consideration health risk factors, but lost earnings associated with being unemployed was assumed independent of health conditions). The value of SSI payments was calculated through an analysis of the March 2011 Current Population Survey. The average disability income received by those receiving SSI payments was calculated by sex and age group (<40, 5 year age bands from 40-69, and >70). This average payment was only assigned to those who were simulated to be receiving SSI payments in the year. The value of missed work days was calculated by predicted personal income ÷ 235 (under the assumption of a 235 day work year).

Modeling Health-Related Quality of Life

Presence of chronic disease and adverse medical events reduces quality of life—measured as quality adjusted life years (QALYs)—which we capture in the model through the use of "disutility" scalars. The EQ-5D instrument is commonly used in health technology assessments as a means of eliciting information on preference weights to calculate quality of life disutilities. As noted by Tengs et al, weights for conditions can vary widely across published sources and trials.[122] The primary source used for this study was from Sullivan et al, who sought to create a "nationally representative catalog of preference-based scores for chronic conditions."[69] U.S. based condition-specific disutilities for each modeled condition, except amputation come from this source. The amputation disutility estimate comes from an analysis by Zhang et al.'s analysis of the Translating Research Into Action For Diabetes study.[70] These were also U.S. based preference weights from an EQ-5D instrument. QALY estimates were calculated by taking the median MEPS general EQ-5D score and subtracting the disutility of each existing medical condition. Age has a disutility of -0.00029 per year, in accordance with the methodology from

[130] http://www.bls.gov/news.release/pdf/wkyeng.pdf , Table 3.

Sullivan, and there are additional utility decrements dependent on the number of chronic conditions that an individual has in a given year.

Exhibit 133. QALY Values

Condition	Utility Decrement
Amputation*	0.108
Blindness*	0.0498
Breast Cancer*	0.0156
Cervical Cancer*	0.014
CHF	0.0635
Chronic Back Pain*	0.0455
CKD*	0.0527
Colorectal Cancer*	0.014
Diabetes*	0.0351
Endometrial Cancer*	0.014
Esophageal Cancer*	0.014
Gallbladder Cancer*	0.014
Gallbladder Disease	0.0288
GERD*	0.0216
Hypertension*	0.025
IHD*	0.0336
Kidney Cancer*	0.014
Leukemia*	0.014
Liver Cancer*	0.014
MI	0.0409
Multiple Myeloma*	0.014
NAFLD*	0.0567
Non-Hodgkin's Lymphoma*	0.014
Obesity*	0.05
Osteoarthritis*	0.0642
Ovarian Cancer*	0.014
Pancreatic Cancer*	0.014
Pneumonia	0.0277
Prostate Cancer*	0.0153
Pulmonary Embolism	0.0198
Renal Failure	0.0603
Stomach Cancer*	0.014
Stroke	0.0524
Thyroid Cancer*	0.014

*Denotes a chronic condition. Note: Baseline utility is 0.844

Model Validation

Overview

Following ISPOR guidelines on validation best practices, there are 5 main types of validations: face validity, verification of internal validity, cross validity, external validity, and predictive validity: [123]

- Review by subject matter experts (face validity)
    - Does the model framework conform to observations about how the system works, and is it consistent with theory?
    - Does the model use the best available inputs and parameters?
    - Are the model outputs consistent with expectations of subject matter experts?
- Internal validation (verification, or technical validity)
    - Review computer code for accuracy
    - Validate parameters in the model against their source
    - Put model through a "stress test" by modeling extreme input values to test whether the model produces expected results
    - Replicate results from published studies used in the model development
    - Replicate incidence, prevalence, and other statistics in the data sources used to create the model (e.g., simulating weight gain associated with aging from NHANES for comparison to the weight by age distribution in NHANES)
- External and predictive validation
    - Replicate findings of studies (e.g., clinical trials) not used in model development
    - Use longitudinal database to compare predicted to actual outcomes
- Between-model validation (cross validation)
    - Compare model outputs for consistency with published results of other models (e.g., compare cancer prediction results to results from the US National Cancer Institute's CISNET program)

The model structure has been reviewed by subject matter experts with clinical backgrounds in obesity, endocrinology, health services research, modeling, and health economics and outcomes research. Both during and after completing model programming, we conducted sensitivity analyses on the model by varying key model inputs to ensure no anomalies.

An important part of the external validation was to use a longitudinal database to simulate future outcomes based on the initial year of data for each person, and then compare predicted to actual outcomes for later years. Ideally, the database would include patient demographics, as well as lab results for the biometrics and indicators

of the diseases listed in Exhibit 10. To the extent that some variables were missing from the longitudinal database, we predicted values based on other information known about each person. Many of the disease indicators were constructed from medical claims data using ICD-9 diagnosis codes.

The primary longitudinal file for external validation was the GE Centricity Electronic Medical Records database. While microsimulation models such as the OPEM will have limited ability to predict outcomes for a specific patient (due to lack of inputs on family history, genetic markers, and substantial variability across individuals in actual health outcomes), the validation efforts have been focused on the degree to which the model simulates correctly for the overall patient population and for subsets of the population. Important subsets were defined by starting year demographics (age group and sex), clinical measures such as body weight status and normal/prediabetes/diabetes status, and other characteristics such as presence of select chronic conditions.

Because simulation logic is the backbone of the model, it was also the focus of the validation. Though we checked the face validity of the cost and QALY calculations, these measures were not validated against external data.

Validation of Diabetes

Diabetes is a major comorbidity of obesity and a major cost driver. Therefore, validation activities for diabetes were more extensive than for other parts of the model. We simulated a prediabetes population with BMI>=24, which is similar to the initial enrollment criteria of Diabetes Prevention Program and Outcomes Study (DPPOS).[124] Even though the DPPOS population was, not completely identical to the simulated population, this scenario serves as a close proxy to external validation of the transition rate from prediabetes to diabetes. DPP reported average, annual diabetes incidence rates over three years of 11.0% and 4.8%, respectively, for the placebo and lifestyle intervention groups.[125] During the follow-up period of DPPOS, annual diabetes incidence was 5.9% for the lifestyle group and 5.6% for the placebo group. The cumulative 10-year incidence rate for DPPOS was 40% for the lifestyle group and 50% for the placebo group.[124] (Exhibit 134)

In the original DPP (average follow up of 3.2 years) a 15 percentage point reduction in the incidence of diabetes was observed in the group undergoing lifestyle intervention compared to the placebo group. Under a scenario that simulated the effects of the DPP intervention, this reduction was 8.6 percentage points.

After 10 years, DPPOS reported that incidence of diabetes in the original lifestyle intervention group was 10 percentage points greater than the placebo group (a 34% reduction[131], with a large 95% confidence interval of 24-42%). However, comparison of 10-year results for the DPPOS lifestyle to placebo arms potentially underestimates the intervention effect. After 3 years masked treatment was discontinued in the DPP when it was shown that lifestyle intervention reduced diabetes incidence by 58% relative to the placebo group.[124] At that time, all participants—including the placebo group—were provided lifestyle sessions similar to the original DPP lifestyle intervention and 57% of the placebo group participated in at least one lifestyle session. After DPPOS enrollment

[131] "x percentage point less" means absolute difference, while y% reduction refers to relative difference. For example, between 40% and 50%, the former is 10 percentage points less than the latter, representing a 20% reduction.

the annual diabetes incidence rates were slightly higher similar for the original DPP lifestyle [5.9%] group relative to the placebo group [5.6%].

Our simulated scenario had a 17.4 percentage point reduction, larger than the 10-year benefit of the DPP. As noted earlier, the DPPOS results for the placebo group became tainted when the placebo group was unblinded after 3 years, patients in the placebo group were offered lifestyle intervention, and 57% of the placebo group participated in at least one lifestyle session. Thus, the DPPOS-reported difference in 10-year incidence between the lifestyle and original placebo groups potentially underestimates the long-term impact of intervention on diabetes incidence, complicating comparisons of the simulated results to those observed.

Exhibit 134. Comparison of Prediabetes-to-Diabetes Incidence Rates

		Cumulative %		Absolute reduction from Lifestyle Intervention	
	Annual %	3-Year	10-Year	3-Year	10-Year
OPEM (simulation) [a]					
Non-intervention					
Total modeled population	9.0	22.0	53.0		
Lifestyle					
Total modeled population	5.1	13.4	35.6	8.6	17.4
DPP/DPPOS (actual).[124;125]					
Non-intervention (placebo) group	11.0 (years 1-3) 5.9 (years 4-10) [b]	33	50[b]		
Lifestyle group	4.8 (years 1-3) 5.6 (years 4-10)	18	40	15	10[b]
Literature (clinical trials)[126]	5-10				
ADDITION Study (Denmark), high risk individuals[127;128]	11.8-17	27-38			

Note: [a] Average annual incidence was calculated by dividing total onset of diabetes over a 10-year period (numerator) by number of people without diabetes (denominator). The 3-year and 10-year incidence rates reflect cumulative incidence (numerator) divided by size of the initial population (denominator). [b] When DPP showed dramatic improvement after 3 years, masked treatment was discontinued in the DPP program and 57% of participants in the placebo group participated in at least one lifestyle session.

Validation with NHANES

An additional validation exercise was to start with the 2003-2004 NHANES sample and simulate the effects of aging on health outcomes over six years for comparison to the population in the 2009-2010 NHANES samples. The NHANES is not a longitudinal file, so the respondents in each NHANES wave are a different group of people—though each wave is weighted to be nationally representative of the U.S. non-institutionalized population. Exhibit 135 compares average BMI in the simulated population to average BMI in the 2009-2010 NHANES sample. For the simulated population, average BMI increases steadily for women in their 30s, 40, and 50s. BMI holds steady in their 60s, and declines in their 70s. Average BMI among women in the 2009-2010 NHANES has a less defined pattern, but in general is higher than simulated rates for women in their 60s and 70s. For men, the average BMI

for the 2009-2010 NHANES sample was relatively unchanged for age groups under 70, whereas for the simulated population average BMI increased from their 30s to their 40s, before starting to decline in their 50s and 60s.

Exhibit 135. Average BMI by Year of Age

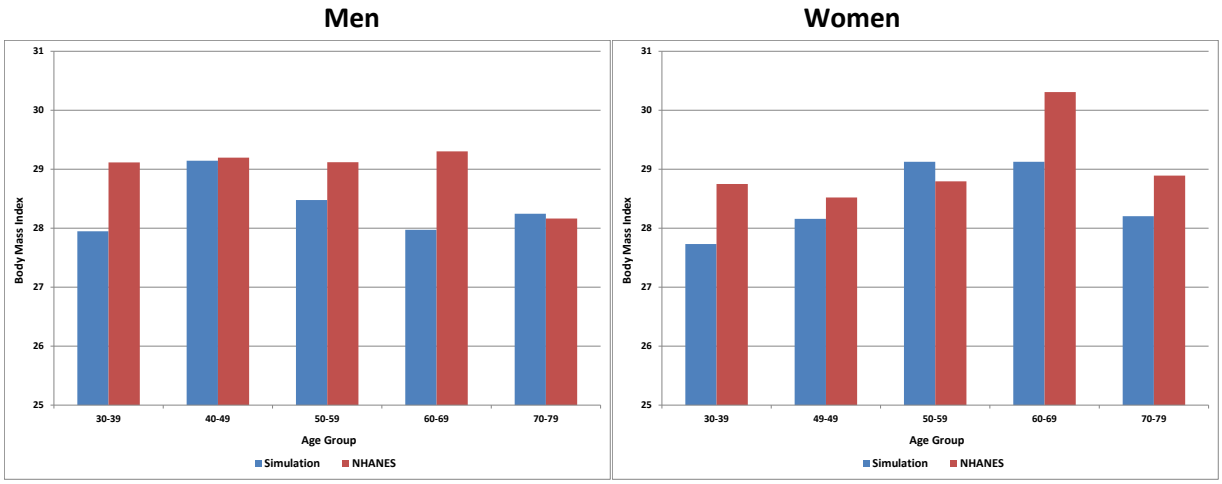


Exhibit 136 shows hypertension prevalence from the 2009-2010 and 2003-2004 NHANES files, as well as prevalence rates from the 2003-2004 population simulated to 2009-2010. Simulated prevalence of hypertension tended to be slightly less than actual rates for each age group.

Exhibit 137 shows similar information on the prevalence of IHD. Prevalence is similar for the two NHANES files and for the simulated population. For the age 65-74 population, IHD prevalence is substantially higher among the 2003-2004 NHANES sample compared to the 2009-2010 sample. This discrepancy could be associated with small sample size. The simulated prevalence falls around the midpoint of the 2003-2004 and 2009-2010 estimates.

Exhibit 136. Prevalence of Hypertension

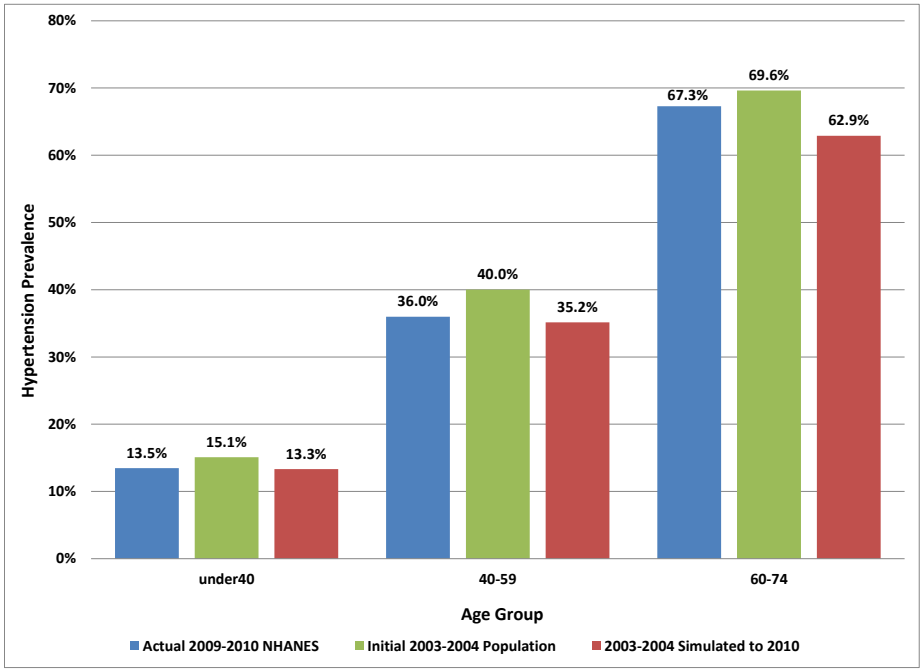
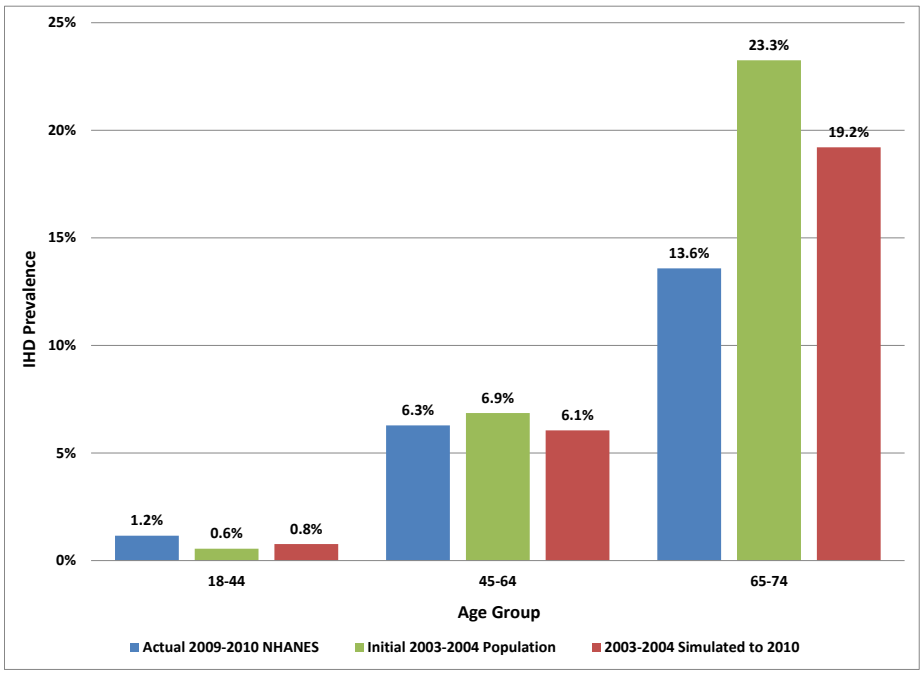


Exhibit 137. Prevalence of IHD



Validation of Cancer Incidence

We drew a population from NHANES that is representative of the general US population, and compared simulated model outcomes of this population against the actual cancer epidemiology data of the US.[129] Average predicted cancer incidence in the first two years was consistent with published estimates. For example, the OEPM predicted 64 cases of breast cancer per 100,000 population, whereas published estimates is 62.3 cases per 100,000.

Exhibit 138. Validation of Cancer Incidence

Type	Predicted incidence /100K	Actual incidence /100K
Breast cancer	64	62.3
Prostate cancer	71	73.9
Colorectal cancer	46	50.0
Non-Hodgkin's Lymphoma	22	19.7
Endometrial cancer	20	24.6
Kidney cancer	15	15.5
Pancreatic cancer	13	12.3
Leukemia	10	13
Liver cancer	5	7.9
Thyroid cancer	12	12.9

Other Validations

Other model validation activities included simulating disease incidence for comparison to published statistics. In Exhibit 139, we compare model simulation incidence for myocardial infarction to estimates from four published studies: Atherosclerosis Risk in Communities Study, Olmstead County Study, Worcester, and Corpus Christi.[84] The simulated incidence rates for men and women are consistent with the range of published estimates.

Exhibit 139. Annual Incidence of Myocardial Infarction

	Atherosclerosis Risk in Communities Study	Olmstead County Study	Worcester Heart Attack Study	Corpus Christi Study	Simulated Year 1 Incidence
Males	0.2%	N/A	N/A	0.2%-0.4%	0.2%
Females	0.4%	N/A	N/A	0.3%-0.5%	0.5%
Both Sexes	N/A	0.2%	0.2%	N/A	0.3%

For validation we simulated CKD incidence for a representative sample of U.S. adults for comparison with rates reported by NIDDK. [132] For the age 20-64 population the simulated rate was slightly lower than the NIDDK-

[132] http://kidney.niddk.nih.gov/kudiseases/pubs/kustats/#3

reported estimates (0.4% versus 0.5%), but the simulated rate was higher for the population age 65 and older (5.3% versus 4.4%) (Exhibit 140).

Simulated incidence of renal failure produced aggregate rates for the age 20-44 population and the age 65-74 population similar to estimates for those age groups reported by the United States Renal Data System (USRDS) (Exhibit 141).[133] For the age 45-64 population the simulated rate was lower than the USRDS rate. For the total U.S. age 75 and older population, the simulated rate was higher than the rate reported by USRDS.

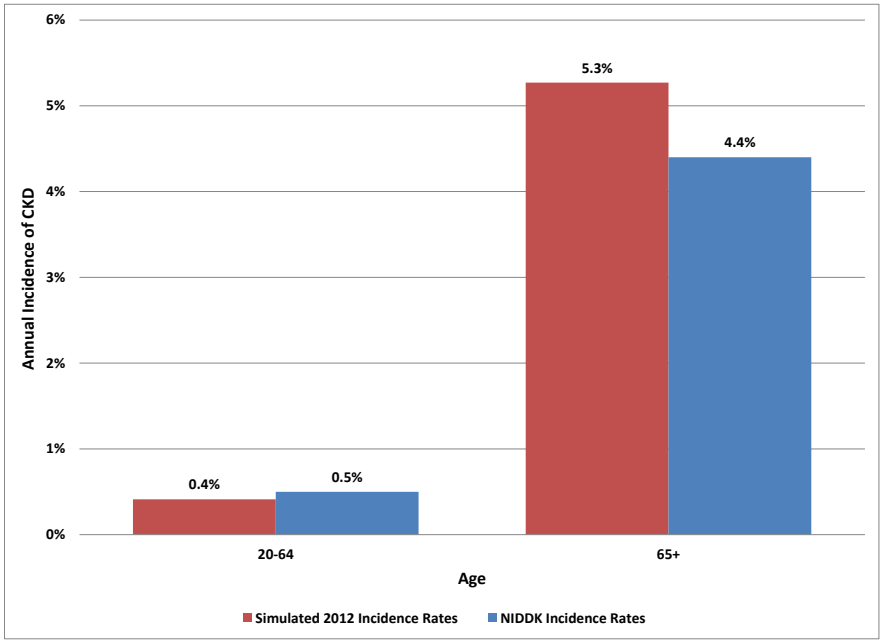
Using the model we simulated over 10 years the prevalence of ischemic heart disease, and compared prevalence estimates at year 10 to current population prevalence estimates reported by CDC.[134] Prevalence estimates were similar for the population age 18-44, slightly lower for the population age 45-64, and slightly higher for the population age 65 and older (Exhibit 142).

Simulated annual probability of stroke was similar to national patterns reported by the American Heart Association (Exhibit 143).[135] For the population age 45 to 64, simulated incidence rate tended to be low for blacks males relative to AHA-reported incidence. Simulated rates tended to be higher than AHA-reported statistics for whites age 45-54 and age 65-74. While the issue of over/under -prediction like this may exist for some demographic groups, it was present among the simulated population for both the intervention and non-intervention scenarios and the effect of such prediction error was further mitigated when estimating the difference in incidence between the two scenarios—which is the primary focus of this analysis.

For diabetes-related amputations we identified few studies that could be used for external validation. Simulated annual probability of diabetes-related amputations were higher than rates reported for a population in Sweden from 1997 through 2006, though the small sample size for that population produced large 95% confidence intervals and simulated rates for males age 65 and older and females age 75-84 were within the reported 95% confidence interval (Exhibit 144). Simulated incidence for the other age groups was higher than the estimates from Sweden. Overall, though, the incidence of amputation is low and sensitivity analysis, discussed later, suggested that amputation incidence had little impact on estimated medical cost savings from the intervention modeled.

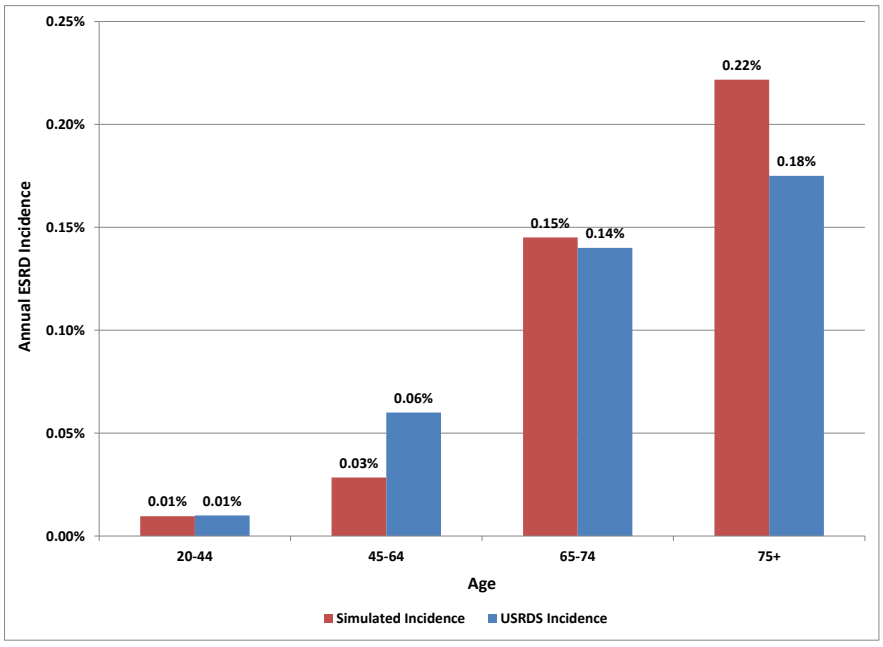
[133] http://www.usrds.org/2012/view/v2_01.aspx
[134] http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6040a1.htm?s_cid=mm6040a1_w
[135] http://www.heart.org/idc/groups/heart-public/@wcm/@sop/@smd/documents/downloadable/ucm_449858.pdf

Exhibit 140. Chronic Kidney Disease Incidence Comparison



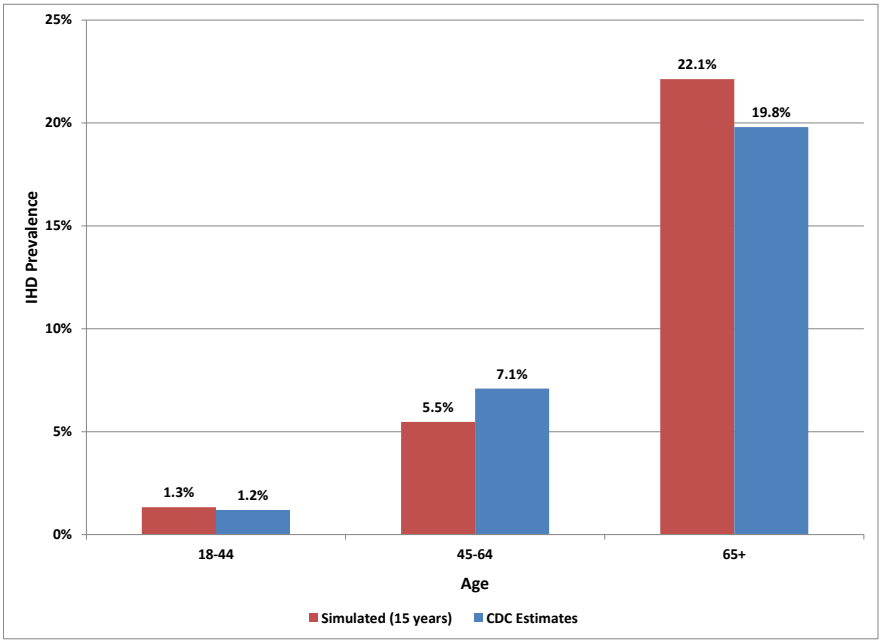
Note: NIDDK= National Institute of Diabetes and Digestive and Kidney Diseases.

Exhibit 141. Renal Failure Disease Incidence Comparison



Note: USRDS=United States Renal Data System.

Exhibit 142. Ischemic Heart Disease Prevalence Comparison



Note: CDC= Centers for Disease Control and Prevention.

Exhibit 143. Stroke Incidence Comparison

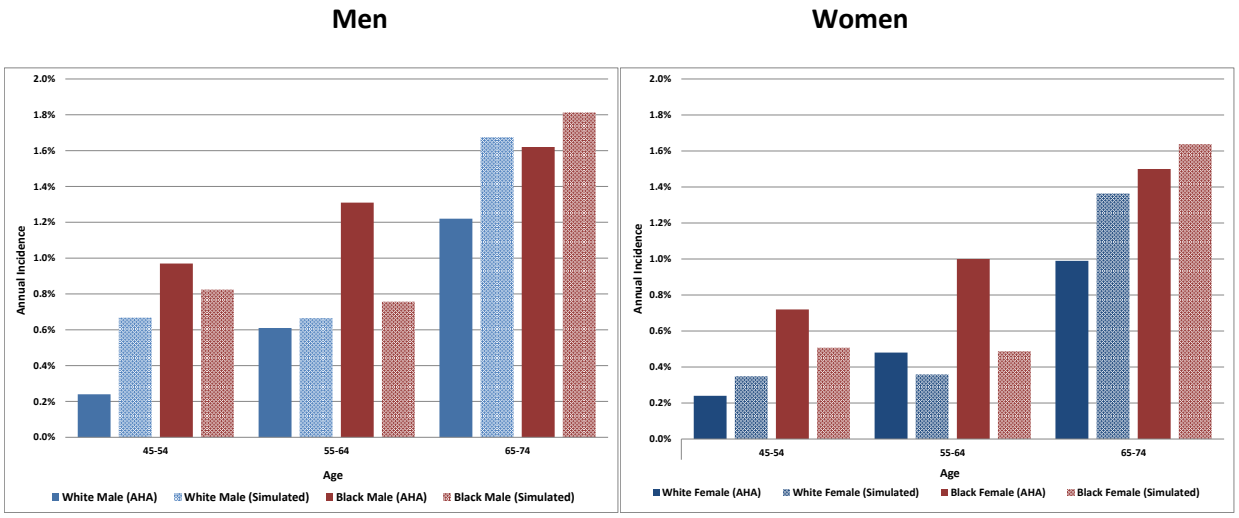
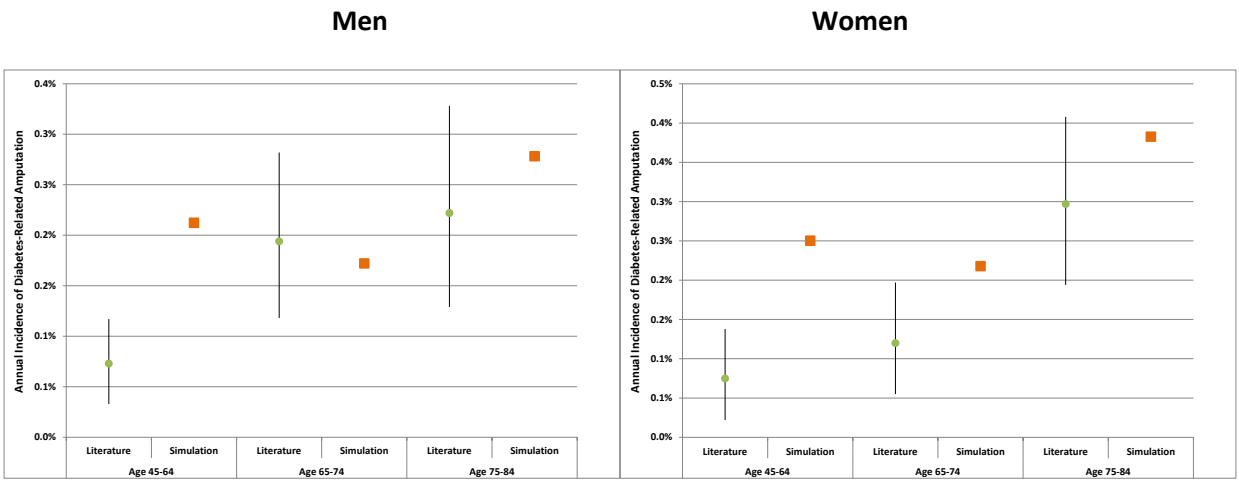


Exhibit 144. Diabetic Amputation Incidence Comparison



Modeling Mortality

An individual's annual probability of mortality was modeled as a function of demographics, clinical characteristics, and presence of disease. The OPEM modeled mortality separately for each medical condition tracked, and then captured the remaining all-cause mortality probability. Simulation of mortality compared an individual's calculated probability of mortality for each potential cause in a given year to a series of random numbers between number 0 and 1 generated from a uniform distribution. If the calculated mortality probability from the prediction equations for a cause exceeded the random number for that cause, then death was assumed. For example, if an individual with CHF has a probability of CHF-related death of 1% from the prediction equation, then a random number ≤ 0.01 would lead to death attributed to CHF, whereas a random number > 0.01 would not.

While there is some evidence that diabetes increases case fatality rate, our review of the literature found insufficient information to quantify any differences in case fatality rate between the populations with and without diabetes with respect to IHD, MI, CHF, stroke, and renal failure.[130;131] Therefore, model results possibly downplay the benefits in reduced mortality associated with preventing or delaying diabetes onset.

We modeled mortality associated with the following:

- **Ischemic Heart Disease**. Estimates of mortality risk from IHD are from the Framingham Heart Study and used a non-proportional hazards Weibull accelerated failure time model to predict the probability of death.[53] If the person has IHD, then risk factors included: sex, log of age, log of SBP, smoking status, log of cholesterol ratio, presence of diabetes, and history of left ventricular hypertrophy.

- **Myocardial Infarction**. Data on mortality within the first 365 days of an incident myocardial infarction came from the Swedish Socialstyrelsen Registry, with rates reported by sex for 5 year age bands.[57]
- **Congestive Heart Failure**. CHF-related mortality came from the National Swedish register on hospital discharges and cause-specific death, where they reported mortality rates by sex and 10-year age bands from age 45-84.[56]
- **Stroke**. Age and sex specific mortality probabilities reflect 1-year mortality rates witnessed in the Arcadia Stroke Registry.[59] Rates were available for ages 18-54, 55-64, 65-74, 75-84, and 85+.
- **Renal failure**. Mortality probability for renal failure came from Lins et al. who examined outcomes after acute renal failure up to one year out.[132] That study found that 51% cases of acute renal failure die in-hospital, and an additional 11% die within a year of event incidence. Thus, a mortality probability of 62% was assumed in cycle for patients experiencing incident renal failure in the model.
- **Chronic Kidney Disease**. A systematic review of CKD and mortality risk by Tonelli et al. suggests that CKD increases mortality risk from a variety of causes (beyond renal failure), with mortality risk relative to the population without CKD ranging from 0.94 to 5.0.[55] We used the midpoint of this range (2.94) and applied a relative risk adjustment to the risk of all-cause mortality for individual's living with CKD.
- **Cancers**. Cancer mortality data were derived from SEER database by National Cancer Institute at http://seer.cancer.gov/data/.
- **All-Cause Mortality**. The final mortality component modeled covers the remaining causes of death. This all-cause mortality was calculated from the CDC WONDER Underlying Cause of Death files for 1999-2010.[133] A top-down approach was used, wherein all deaths by age and sex were the starting point, and deaths attributable to modeled conditions were subsequently subtracted. The ICD-10 codes excluded from the all-cause mortality analysis are listed in Exhibit 146. Following these removals, the updated mortality rates were calculated by dividing remaining deaths by remaining population, to avoid double counting. The all-cause mortality rates used are displayed in Exhibit 145.

Exhibit 145. All-Cause Mortality Rates

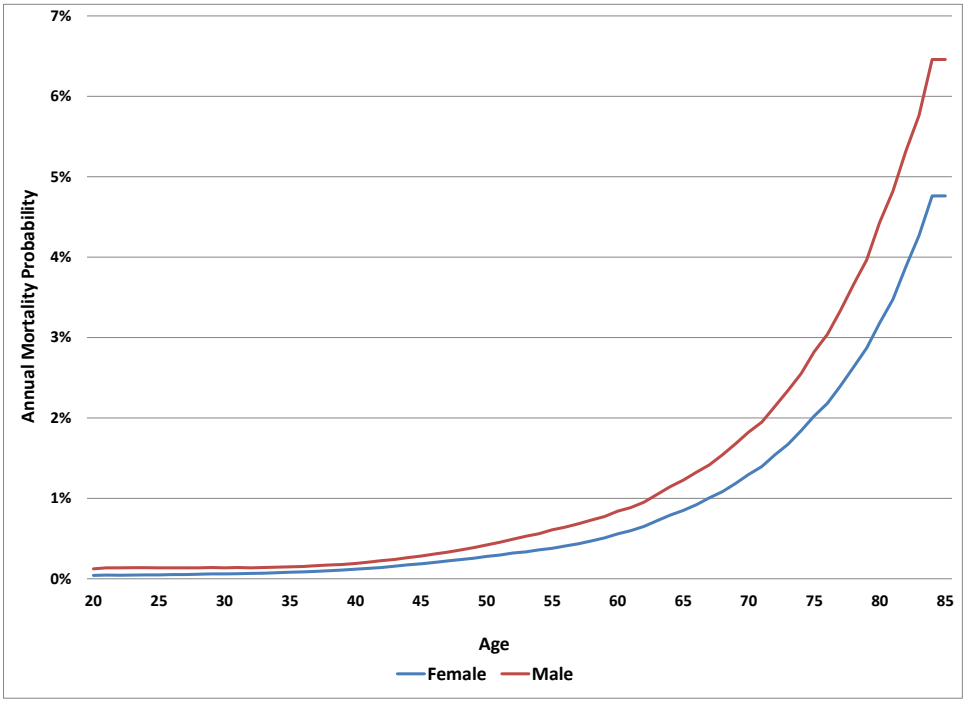


Exhibit 146. ICD-10 Codes Excluded from All-Cause Mortality Analysis

Condition	ICD-10 Codes for All-Cause Mortality
Stroke	I60-169
IHD	I20-I25
MI	I21-I22 (included in IHD)
CHF	I50
Renal failure	N17,N19
CKD	N18
Breast cancer	C50
Cervical cancer	C53
Colorectal cancer	C18,C20
Endometrial cancer	C54,C55
Esophageal cancer	C15
Gallbladder cancer	C23
Kidney cancer	C64
Leukemia	C91-C95
Liver cancer	C22
Multiple myeloma	C90
Non-Hodgkin's lymphoma	C82-C83
Ovarian cancer	C56
Pancreatic cancer	C25
Prostate cancer	C61
Stomach cancer	C16
Thyroid cancer	C73
Pneumonia	J12-J18
Pulmonary embolism	I26

Model validation efforts included running a 10-year simulation for a representative sample of the adult population in the U.S. and comparing mortality rates by cause of death with published estimates. Examples for stroke and CHF morality can be found in Exhibit 148 and Exhibit 147. Published mortality rates associated with CHF came from the Rotterdam Study.[56] Published rates for stroke mortality came from the 2013 American Heart Association Heart Disease and Stroke Statistical Update.[134] The results below indicate good agreement between simulated and actual data.

Exhibit 147. CHF Mortality Rate

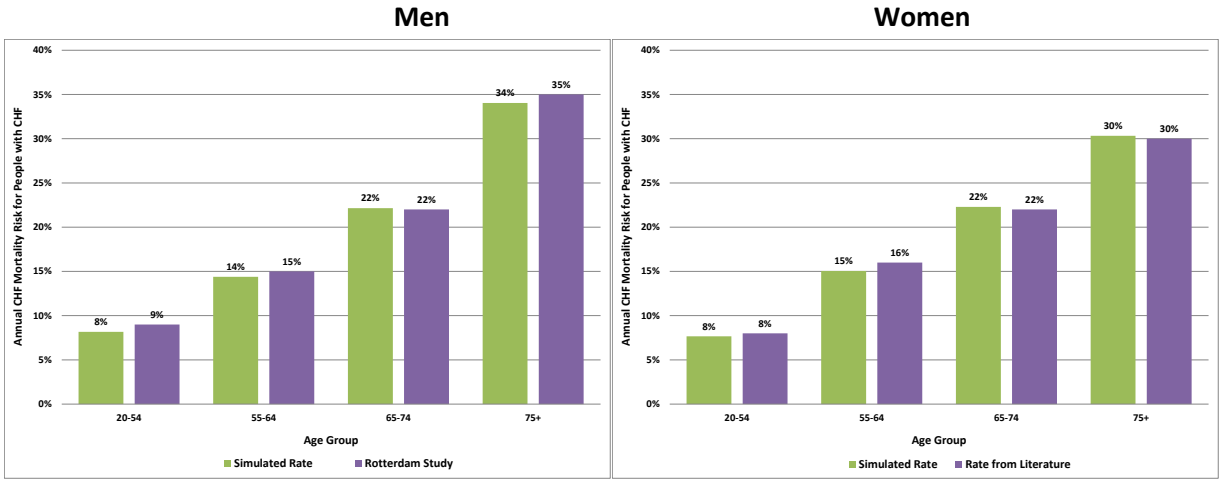
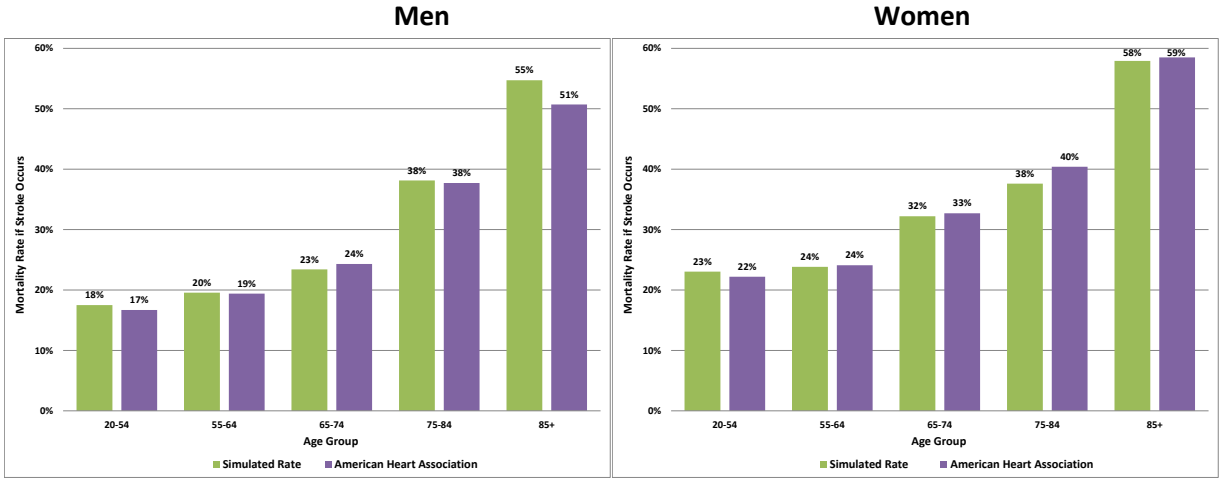


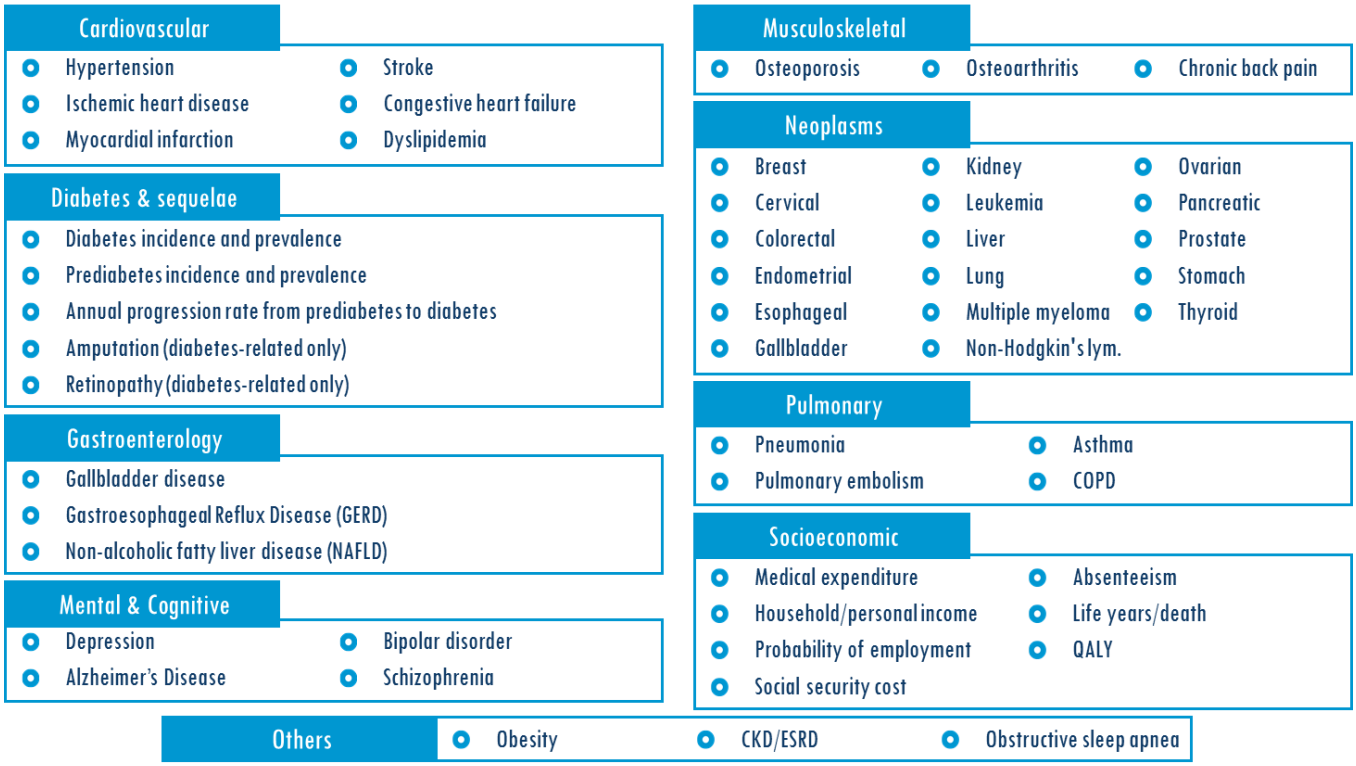
Exhibit 148. Stroke Mortality Rate



Model Outcomes

The DPMM is capable of simulating more than 50 clinical and economic outcomes.(Exhibit 149)

Exhibit 149. DPMM Model Outcomes



Model Limitations

Models are simplified representations of complex systems, and this model covered three complex systems: (1) the human body and the epidemiology of disease, (2) the health care system and the relationship between patient health states and medical expenditures, and (3) the economic system and the relationship between patient health states and economic outcomes. Like all models, the OPEM makes assumptions and uses incomplete or imperfect data to quantify complex relationships. Limitations of the model include the following:

1. The lack of a single longitudinal data source of the U.S. prediabetes population that covers a sufficient time period and is of sufficient size to quantify the relationships between disease onset and patient characteristics meant that the parameters and equations in the simulation model came from multiple sources. The characteristics of the population in these sources varied. Some sources (e.g., Framingham) included representative samples of the population including people with normal blood glucose levels. Other sources (e.g., UKPDS) collected data on a population outside the U.S.
2. To fill gaps in the literature, some model parameters were built on analyses with cross-sectional data from NHANES. Validation activities found, however, that the key relationships used in the model based

on cross-sectional data produced population outcomes consistent with outcomes based on published longitudinal trends.

3. While modeled risk factors and disease incidence/prevalence were generally consistent with published estimates, for some populations the validation activities suggested that simulated growth rate in biometrics and disease onset appeared to be high (or low) when compared to published sources. Sensitivity analysis suggests that over (under) predicting annual change in patient health states had relatively little impact on the estimated program impact, in large part because any over (under) estimation occurred among both the intervention and non-intervention scenarios and program impact was calculated as the difference in outcomes between these two scenarios.
4. Older data sources were sometimes used (e.g., Framingham and UKPDS), and standards of care such as statin use have evolved over time. This may lead to a cohort effect that biases the risk estimation of certain health conditions. For example, data from the Look-AHEAD trial and other studies report that statin use has increased over time and is associated with decreased risk of adverse CVD events, and that after controlling for cholesterol levels the impact of body weight loss on CVD outcomes largely disappears.[135-137]

Reference List

(1) Yang W, Dall TM, Zhang Y et al. Simulation Of Quitting Smoking In The Military Shows Higher Lifetime Medical Spending More Than Offset By Productivity Gains. *Health Affairs* 2012;31:2717-2726.

(2) Dall TM, Storm MV, Semilla AP, Wintfeld N, O'Grady M, Narayan KV. Value of lifestyle intervention to prevent diabetes and sequelae. *American journal of preventive medicine* 2015;48:271-280.

(3) Su W, Huang J, Chen F et al. Modeling the clinical and economic implications of obesity using microsimulation. *J Med Econ* 2015;1-27.

(4) Hsu CY, Bates DW, Kuperman GJ, Curhan GC. Diabetes, hemoglobin A(1c), cholesterol, and the risk of moderate chronic renal insufficiency in an ambulatory population. *Am J Kidney Dis* 2000;36:272-281.

(5) Kshirsagar AV, Bang H, Bomback AS et al. A simple algorithm to predict incident kidney disease. *Arch Intern Med* 2008;168:2466-2473.

(6) Sheehan TJ, DuBrava S, DeChello LM, Fang Z. Rates of weight change for black and white Americans over a twenty year period. *Int J Obes Relat Metab Disord* 2003;27:498-504.

(7) Clarke PM, Gray AM, Briggs A et al. A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68). *Diabetologia* 2004;47:1747-1759.

(8) Neter JE, Stam BE, Kok FJ, Grobbee DE, Geleijnse JM. Influence of weight reduction on blood pressure a meta-analysis of randomized controlled trials. *Hypertension* 2003;42:878-884.

(9) Danaei G, Finucane MM, Lu Y et al. 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* 2011;378:31-40.

(10) Stevens GA, Finucane MM, De-Regil LM et al. Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995GÇô2011: a systematic analysis of population-representative data. *The Lancet Global Health* 2013;1:e16-e25.

(11) Gadde KM, Allison DB, Ryan DH et al. Effects of low-dose, controlled-release, phentermine plus topiramate combination on weight and associated comorbidities in overweight and

obese adults (CONQUER): a randomised, placebo-controlled, phase 3 trial. *The Lancet* 2011;377:1341-1352.

(12) Heianza Y, Arase Y, Fujihara K et al. Longitudinal Trajectories of HbA1c and Fasting Plasma Glucose Levels During the Development of Type 2 Diabetes The Toranomon Hospital Health Management Center Study 7 (TOPICS 7). *Diabetes care* 2012;35:1050-1052.

(13) Heianza Y, Arase Y, Fujihara K et al. Screening for pre-diabetes to predict future diabetes using various cut-off points for HbA1c and impaired fasting glucose: the Toranomon Hospital Health Management Center Study 4 (TOPICS 4). *Diabetic Medicine* 2012;29:e279-e285.

(14) Wilson PW, Anderson KM, Harri T, Kannel WB, Castelli WP. Determinants of change in total cholesterol and HDL-C with age: the Framingham Study. *Journal of gerontology* 1994;49:M252-M257.

(15) Kannel WB, Wolf PA, Benjamin EJ, Levy D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. *The American journal of cardiology* 1998;82:2N-9N.

(16) Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC public health* 2009;9:88.

(17) Kshirsagar AV, Bang H, Bomback AS et al. A simple algorithm to predict incident kidney disease. *Arch Intern Med* 2008;168:2466-2473.

(18) Hsu CY, Bates DW, Kuperman GJ, Curhan GC. Diabetes, hemoglobin A(1c), cholesterol, and the risk of moderate chronic renal insufficiency in an ambulatory population. *Am J Kidney Dis* 2000;36:272-281.

(19) D'Agostino RB, Vasan RS, Pencina MJ et al. General cardiovascular risk profile for use in primary care the Framingham Heart Study. *Circulation* 2008;117:743-753.

(20) World Health Organization. Definition and diagnosis of diabetes. 2010. 1-31-2015.

Ref Type: Online Source

(21) Hippisley-Cox J, Coupland C. Predicting the risk of Chronic Kidney Disease in Men and Women in England and Wales: prospective derivation and external validation of the QKidney-Scores. *BMC family practice* 2010;11:49.

(22) Brancati FL, Whelton PK, Randall BL, Neaton JD, Stamler J, Klag MJ. Risk of end-stage renal disease in diabetes mellitus: a prospective cohort study of men screened for MRFIT. Multiple Risk Factor Intervention Trial. *JAMA* 1997;278:2069-2074.

(23) Field AE, Coakley EH, Must A et al. Impact of overweight on the risk of developing common chronic diseases during a 10-year period. *Archives of internal medicine* 2001;161:1581-1586.

(24) Maram ES, Linos DA, Thistle JL, Linos A, O'Fallon WM, Kurland LT. The incidence of cholelithiasis and cholecystitis in Rochester, Minn, 1950 through 1969. *Archives of internal medicine* 1990;150:1833-1836.

(25) Nilsson M, Johnsen R, Ye W, Hveem K, Lagergren J. Obesity and estrogen as risk factors for gastroesophageal reflux symptoms. *JAMA* 2003;290:66-72.

(26) Ruigomez A, Garcia Rodriguez LA, Wallander M, Johansson S, Graffner H, Dent J. Natural history of gastro-oesophageal reflux disease diagnosed in general practice. *Alimentary pharmacology & therapeutics* 2004;20:751-760.

(27) de Simone G, Devereux RB, Roman MJ, Alderman MH, Laragh JH. Relation of obesity and gender to left ventricular hypertrophy in normotensive and hypertensive adults. *Hypertension* 1994;23:600-606.

(28) Hooi JD, Kester AD, Stoffers HE, Overdijk MM, van Ree JW, Knottnerus JA. Incidence of and risk factors for asymptomatic peripheral arterial occlusive disease: a longitudinal study. *American journal of epidemiology* 2001;153:666-672.

(29) Nichols GA, Reinier K, Chugh SS. Independent contribution of diabetes to increased prevalence and incidence of atrial fibrillation. *Diabetes care* 2009;32:1851-1856.

(30) Selvin E, Erlinger TP. Prevalence of and risk factors for peripheral arterial disease in the United States results from the National Health and Nutrition Examination Survey, 1999-2000. *Circulation* 2004;110:738-743.

(31) Zhang Y, Dall TM, Mann SE et al. The economic costs of undiagnosed diabetes. *Population Health Management* 2009;12:95-101.

(32) Wilson PW, Bozeman SR, Burton TM, Hoaglin DC, Ben-Joseph R, Pashos CL. Prediction of first events of coronary heart disease and stroke with consideration of adiposity. *Circulation* 2008;118:124-130.

(33) Clarke PM, Gray AM, Briggs A et al. A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68). *Diabetologia* 2004;47:1747-1759.

(34) Ratziu V, Bellentani S, Cortez-Pinto H, Day C, Marchesini G. A position statement on NAFLD/NASH based on the EASL 2009 special conference. *Journal of hepatology* 2010;53:372-384.

(35) Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC public health* 2009;9:88.

(36) Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis & Rheumatism* 1995;38:1134-1141.

(37) Kornum JB, N++rgaard M, Dethlefsen C et al. Obesity and risk of subsequent hospitalization with pneumonia. *European Respiratory Journal* 2010.

(38) Stein PD, Beemath A, Olson RE. Obesity as a risk factor in venous thromboembolism. *Am J Med* 2005;118:978-980.

(39) Silverstein MD, Heit JA, Mohr DN, Petterson TM, O'Fallon WM, Melton LJ. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Archives of internal medicine* 1998;158:585-593.

(40) Green LE, Dinh TA, Smith RA. An estrogen model: the relationship between body mass index, menopausal status, estrogen replacement therapy, and breast cancer risk. *Computational and mathematical methods in medicine* 2012;2012.

(41) Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *Bmj* 2007;335:1134.

(42) Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. *Cancer Epidemiology Biomarkers & Prevention* 2007;16:2533-2547.

(43) Nagle CM, Marquart L, Bain CJ et al. Impact of weight change and weight cycling on risk of different subtypes of endometrial cancer. *European Journal of Cancer* 2013;49:2717-2726.

(44) Chow WH, Blot WJ, Vaughan TL et al. Body mass index and risk of adenocarcinomas of the esophagus and gastric cardia. *Journal of the National Cancer Institute* 1998;90:150-155.

(45) Larsson SC, Wolk A. Obesity and the risk of gallbladder cancer: a meta-analysis. *Br J Cancer* 2007;96:1457-1461.

(46) Larsson SC, Wolk A. Overweight and obesity and incidence of leukemia: A meta-analysis of cohort studies. *International journal of cancer* 2008;122:1418-1421.

(47) Larsson SC, Wolk A. Obesity and the risk of gallbladder cancer: a meta-analysis. *British journal of cancer* 2007;96:1457-1461.

(48) Birmann BM, Giovannucci E, Rosner B, Anderson KC, Colditz GA. Body mass index, physical activity, and risk of multiple myeloma. *Cancer Epidemiology Biomarkers & Prevention* 2007;16:1474-1478.

(49) Lim U, Morton LM, Subar AF et al. Alcohol, smoking, and body size in relation to incident Hodgkin's and non-Hodgkin's lymphoma risk. *American journal of epidemiology* 2007;166:697-708.

(50) Genkinger JM, Spiegelman D, Anderson KE et al. A pooled analysis of 14 cohort studies of anthropometric factors and pancreatic cancer risk. *International journal of cancer* 2011;129:1708-1717.

(51) Wright ME, Chang S, Schatzkin A et al. Prospective study of adiposity and weight change in relation to prostate cancer incidence and mortality. *Cancer* 2007;109:675-684.

(52) Leitzmann MF, Brenner A, Moore SC et al. Prospective study of body mass index, physical activity and thyroid cancer. *International journal of cancer* 2010;126:2947-2956.

(53) Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. *American heart journal* 1991;121:293-298.

(54) D'Agostino RB. Stroke risk profile: adjustment for antihypertensive medication. The Framingham Study. *Stroke* 1994;25:40-43.

(55) Tonelli M, Wiebe N, Culleton B et al. Chronic kidney disease and mortality risk: a systematic review. *J Am Soc Nephrol* 2006;17:2034-2047.

(56) Schaufelberger M, Swedberg K, Koster M, Rosen M, Rosengren A. Decreasing one-year mortality and hospitalization rates for heart failure in Sweden Data from the Swedish Hospital Discharge Registry 1988 to 2000. *European heart journal* 2004;25:300-307.

(57) Socialstyrelsen. Swedish Health and Welfare Statistical Databases: AMI Statistics. *Socialstyrelsen* [serial online] 2013.

(58) Carson JL, Kelley MA, Duff A et al. The clinical course of pulmonary embolism. *N Engl J Med* 1992;326:1240-1245.

(59) Vemmos KN, Bots ML, Tsibouris PK et al. Prognosis of stroke in the south of Greece: 1 year mortality, functional outcome and its determinants: the Arcadia Stroke Registry. *Journal of Neurology, Neurosurgery & Psychiatry* 2000;69:595-600.

(60) Yabroff KR, Lamont EB, Mariotto A et al. Cost of care for elderly cancer patients in the United States. *Journal of the National Cancer Institute* 2008;100:630-641.

(61) Glasgow RE, Cho M, Hutter MM, Mulvihill SJ. The spectrum and cost of complicated gallstone disease in California. *Archives of Surgery* 2000;135:1021-1025.

(62) Colice GL, Morley MA, Asche C, Birnbaum HG. Treatment costs of community-acquired pneumonia in an employed population. *CHEST Journal* 2004;125:2140-2145.

(63) Park B, Messina L, Dargon P, Huang W, Ciocca R, Anderson FA. Recent Trends in Clinical Outcomes and Resource Utilization for Pulmonary Embolism in the United States: Findings From the Nationwide Inpatient Sample. *CHEST Journal* 2009;136:983-990.

(64) Kotlarz H, Gunnarsson CL, Fang H, Rizzo JA. Insurer and out-of-pocket costs of osteoarthritis in the US: Evidence from national survey data. *Arthritis & Rheumatism* 2009;60:3546-3553.

(65) Bloom BS, Jayadevappa R, Wahl P, Cacciamanni J. Time trends in cost of caring for people with gastroesophageal reflux disease. *The American journal of gastroenterology* 2001;96:S64-S69.

(66) Crow WT, Willis DR. Estimating cost of care for patients with acute low back pain: a retrospective review of patient records. *JAOA: Journal of the American Osteopathic Association* 2009;109:229-233.

(67) Younossi ZM, Zheng L, Stepanova M, Henry L, Venkatesan C, Mishra A. Trends in Outpatient Resource Utilizations and Outcomes for Medicare Beneficiaries With Nonalcoholic Fatty Liver Disease. *Journal of clinical gastroenterology* 2014.

(68) Riley GF, Lubitz JD. Long-term trends in Medicare payments in the last year of life. *Health Serv Res* 2010;45:565-576.

(69) Sullivan PW, Ghushchyan V. Preference-based EQ-5D index scores for chronic conditions in the United States. *Medical Decision Making* 2006;26:410-420.

(70) Zhang P, Brown MB, Bilik D, Ackermann RT, Li R, Herman WH. Health Utility Scores for People With Type 2 Diabetes in US Managed Care Health Plans Results from Translating Research Into Action for Diabetes (TRIAD). *Diabetes care* 2012;35:2250-2256.

(71) Centers of Disease Control. National Health and Nutrition Examination Survey 2009 - 2010 Data Documentation, Codebook, and Frequencies: Smoking - Cigarette Use (SMQ_F). 9-1-2011.

Ref Type: Online Source

(72) Kiefe CI, Williams OD, Lewis CE, Allison JJ, Sekar P, Wagenknecht LE. Ten-year changes in smoking among young adults: are racial differences explained by socioeconomic factors in the CARDIA study? *American Journal of Public Health* 2001;91:213.

(73) Centers for Disease Control and Prevention. Fact Sheets- Quitting Smoking. 5-21-2015.

Ref Type: Online Source

(74) Garvey AJ, Bliss RE, Hitchcock JL, Heinold JW, Rosner B. Predictors of smoking relapse among self-quitters: a report from the Normative Aging Study. *Addictive behaviors* 1992;17:367-377.

(75) Swan GE, Denk CE, Parker SD, Carmelli D, Furze CT, Rosenman RH. Risk factors for late relapse in male and female ex-smokers. *Addictive behaviors* 1988;13:253-266.

(76) Centers for Disease Control and Prevention. BRFSS 2013 Survey Data and Documentation. 9-23-2014.

Ref Type: Online Source

(77) Centers for Disease Control and Prevention. NHANES 2011- 2012 Data Documentation, Codebook and Frequencies Alcohol Use (ALQ_G). 10-1-2013.

Ref Type: Online Source

(78) National Institutes of Health. The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. 2000. 6-30-2014.

Ref Type: Online Source

(79) Dutton G, Loria C, Carnethon M et al. Abstract PO21: Secular and Aging-Related Weight Gain Slows among U.S. Adults: The CARDIA Study [abstract]Dutton G, Loria C, Carnethon M et al. *Circulation* 2013;127:AP021

(80) Sheehan TJ, DuBrava S, DeChello LM, Fang Z. Rates of weight change for black and white Americans over a twenty year period. *Int J Obes Relat Metab Disord* 2003;27:498-504.

(81) Ford ES, Dietz WH. Trends in energy intake among adults in the United States: findings from NHANES. *Am J Clin Nutr* 2013;97:848-853.

(82) Dwyer-Lindgren L, Freedman G, Engell RE et al. Prevalence of physical activity and obesity in US counties, 2001--2011: a road map for action. *Popul Health Metr* 2013;11:7.

(83) Orpana HM, Tremblay MS, Fines P. Trends in weight change among Canadian adults. *Health Rep* 2007;18:9-16.

(84) Roger VL, Go AS, Lloyd-Jones DM et al. Heart Disease and Stroke Statistics-2012 Update A Report From the American Heart Association. *Circulation* 2012;125:e2-e220.

(85) Vasan R, Larson MG, Leip E, Evans J, O'Donnell C, Kannel W. Impact of high-normal blood pressure on the risk of cardiovascular disease. *N Engl J Med* 2001;345:1291-1297.

(86) Stamler J, Stamler R, Neaton J. Blood pressure, systolic and diastolic, and cardiovascular risks: US population data. *Archives of internal medicine* 1993;153:598.

(87) Kannel W. Blood pressure as a cardiovascular risk factor. *JAMA* 1996;275:1571-1576.

(88) Sagie A, Larson MG, Levy D. The natural history of borderline isolated systolic hypertension. *N Engl J Med* 1993;329:1912-1917.

(89) MacMahon S, Peto R, Collins R, Godwin J, Cutler J, Sorlie P. Blood pressure, stroke, and coronary heart disease: part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 2014;335:765-774.

(90) Neaton J, Wentworth D. Blood Pressure, Cigarette Smoking, and Death From Coronary Heart Disease Overall Findings and Differences by Age for 316099 White Men. *Arch Intern Med* 1992;152:56.

(91) Rodgers A, MacMahon S, Yee T, Clark T, Keung C, Chen Z. Blood pressure, cholesterol, and stroke in eastern Asia. *Lancet* 1998;352:1801-1807.

(92) van den Hoogen P, Feskens E, Nagelkerke N, Menotti A, Nissinen A, Kromhout D. The relation between blood pressure and mortality due to coronary heart disease among men in different parts of the world. *N Engl J Med* 2000;342:1-8.

(93) Wang Y, Beydoun MA, Liang L, Caballero B, Kumanyika SK. Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. *Obesity* 2008;16:2323-2330.

(94) Boundurant S, Wedge R. Combating tobacco use in military and veteran populations. *National Academies Press* 20109.

(95) Bunn W, Stave G, Downs K, Alvir J, Dirani R. Effect of smoking status on productivity loss. *Journal of Occupational and Environmental Medicine* 2006;48:1099-1108.

(96) Goldman D, Zheng Y, Girosi F, Michaud P, Olshansky S, Cutler D. The benefits of risk factor prevention in Americans aged 51 years and older. *American Journal of Public Health* 2009;99.

(97) American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. *Diabetes care* 2010;33:S62-S69.

(98) Burgin NE, Gosden TP. Donepezil- Submission to the National Institute for Health and Clinical Excellence Mutliple Technology Appraisal. 5-3-2010.

Ref Type: Online Source

(99) Bond M, Rogers G, Peters J et al. The effectiveness and cost-effectiveness of donepezil, galantamine, rivastigmine and memantine for the treatment of AlzheimerGÇÖs disease (review of Technology Appraisal No. 111): a systematic review and economic model. 2012.

(100) Fargo K, Bleilar L. 2014 Alzheimer's Disease Facts and Figures, Alzheimer's & Dementia. *Alzheimer's Association* 2014;10.

(101) Alzheimer's Association. What Is Alzheimer's? 2016.

Ref Type: Online Source

(102) Gurland BJ, Wilder DE, Lantigua R et al. Rates of dementia in three ethnoracial groups. *International journal of geriatric psychiatry* 1999;14:481-493.

(103) Hebert LE, Beckett LA, Scherr PA, Evans DA. Annual incidence of Alzheimer disease in the United States projected to the years 2000 through 2050. *Alzheimer Disease & Associated Disorders* 2001;15:169-173.

(104) Bowen JD, Malter AD, Sheppard L et al. Predictors of mortality in patients diagnosed with probable Alzheimer's disease. *Neurology* 1996;47:433-439.

(105) Viner S, Szalai J, Hoffstein V. Are history and physical examination a good screening test for sleep apnea. *Annals of internal medicine* 2011;115.

(106) Philips B. Sleep Apnea: Prevalence of Risk Factors in a General Population. *Southern Medical Journal,* 1989.

(107) Newman A, Foster A, Givelber R, Neito F, Redline S, Young T. Progression and regression of sleep-disordered breathing with changes in weight: the Sleep Heart Health Study. *Arch Intern Med* 2005;165.

(108) Wang TJ, Parise H, Levy D et al. Obesity and the risk of new-onset atrial fibrillation. *JAMA* 2004;292:2471-2477.

(109) Wanahita N, Messerli FH, Bangalore S, Gami AS, Somers VK, Steinberg JS. Atrial fibrillation and obesity--results of a meta-analysis. *Am Heart J* 2008;155:310-315.

(110) Kurth T, Gaziano JM, Berger K et al. Body mass index and the risk of stroke in men. *Arch Intern Med* 2002;162:2557-2562.

(111) Bergstrom A, Pisani P, Tenet V, Wolk A, Adami HO. Overweight as an avoidable cause of cancer in Europe. *Int J Cancer* 2001;91:421-430.

(112) Adams KF, Leitzmann MF, Albanes D et al. Body size and renal cell cancer incidence in a large US cohort study. *American journal of epidemiology* 2008;168:268-277.

(113) Larsson SC, Wolk A. Overweight, obesity and risk of liver cancer: a meta-analysis of cohort studies. *British journal of cancer* 2007;97:1005-1008.

(114) Chen Y, Liu L, Wang X et al. Body mass index and risk of gastric cancer: a meta-analysis of a population with more than ten million from 24 prospective studies. *Cancer Epidemiology Biomarkers & Prevention* 2013;22:1395-1408.

(115) Raskob GE, Silverstein R, Bratzler DW, Heit JA, White RH. Surveillance for deep vein thrombosis and pulmonary embolism: recommendations from a national workshop. *Am J Prev Med* 2010;38:S502-S509.

(116) Clark JM, Brancati FL, Diehl AM. Nonalcoholic fatty liver disease. *Gastroenterology* 2002;122:1649-1657.

(117) Discacciati A, Orsini N, Wolk A. Body mass index and incidence of localized and advanced prostate cancerGÇöa doseGÇôresponse meta-analysis of prospective studies. *Annals of Oncology* 2012;mdr603.

(118) Giovannucci E, Liu Y, Platz EA, Stampfer MJ, Willett WC. Risk factors for prostate cancer incidence and progression in the health professionals followGÇÉup study. *International journal of cancer* 2007;121:1571-1578.

(119) Efstathiou JA, Bae K, Shipley WU et al. Obesity and mortality in men with locally advanced prostate cancer. *Cancer* 2007;110:2691-2699.

(120) Endogenous Hormones Breast Cancer Collaborative Group. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *Journal of the National Cancer Institute* 2003;95:1218-1226.

(121) Campbell JD, Ramsey SD. The costs of treating breast cancer in the US. *Pharmacoeconomics* 2009;27:199-209.

(122) Tengs T, Wallace A. One thousand health-related quality-of-life estimates. *Medical care* 2000;583-637.

(123) Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. *Value Health* 2012;15:843-850.

(124) Knowler WC, Fowler SE, Hamman RF et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. *Lancet* 2009;374:1677-1686.

(125) Knowler WC, Barrett-Connor E, Fowler SE et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393-403.

(126) Gerstein HC, Santaguida P, Raina P et al. Annual incidence and relative risk of diabetes in people with various categories of dysglycemia: a systematic overview and meta-analysis of prospective studies. *Diabetes Res Clin Pract* 2007;78:305-312.

(127) Rasmussen SS, Glumer C, Sandbaek A, Lauritzen T, Borch-Johnsen K. Determinants of progression from impaired fasting glucose and impaired glucose tolerance to diabetes in a high-risk screened population: 3 year follow-up in the ADDITION study, Denmark. *Diabetologia* 2008;51:249-257.

(128) Rasmussen SS, Glumer C, Sandbaek A, Lauritzen T, Borch-Johnsen K. Progression from impaired fasting glucose and impaired glucose tolerance to diabetes in a high-risk screening programme in general practice: the ADDITION Study, Denmark. *Diabetologia* 2007;50:293-297.

(129) National Cancer Institute. Surveillance, Epidemiology, and End Results Program. 2011. 10-9-2014.

Ref Type: Online Source

(130) Haffner SM, Lehto S, Ronnemaa T, Pyrorala K, Laakso M. Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. *N Engl J Med* 1998;339:229-234.

(131) Ho KK, Anderson KM, Kannel WB, Grossman W, Levy D. Survival after the onset of congestive heart failure in Framingham Heart Study subjects. *Circulation* 1993;88:107-115.

(132) Lins RL, Elseviers MM, Daelemans R. Severity scoring and mortality 1 year after acute renal failure. *Nephrol Dial Transplant* 2006;21:1066-1068.

(133) Centers for Disease Control and Prevention NCHS. Underlying Cause of Death 1999-2010 on CDC WONDER Online Database, released 2012. *Centers for Disease Control and Prevention, National Center for Health Statistics* [serial online] 2013; Accessed July 1, 2013.

(134) Go AS, Mozaffarian D, Roger VL et al. Heart disease and stroke statistics--2013 update: a report from the American Heart Association. *Circulation* 2013;127:e6-e245.

(135) Nylen ES, Faselis C, Kheirbek R, Myers J, Panagiotakos D, Kokkinos P. Statins modulate the mortality risk associated with obesity and cardiorespiratory fitness in diabetics. *J Clin Endocrinol Metab* 2013;98:3394-3401.

(136) Wing RR. Long-term effects of a lifestyle intervention on weight and cardiovascular risk factors in individuals with type 2 diabetes mellitus: four-year results of the Look AHEAD trial. *Arch Intern Med* 2010;170:1566-1575.

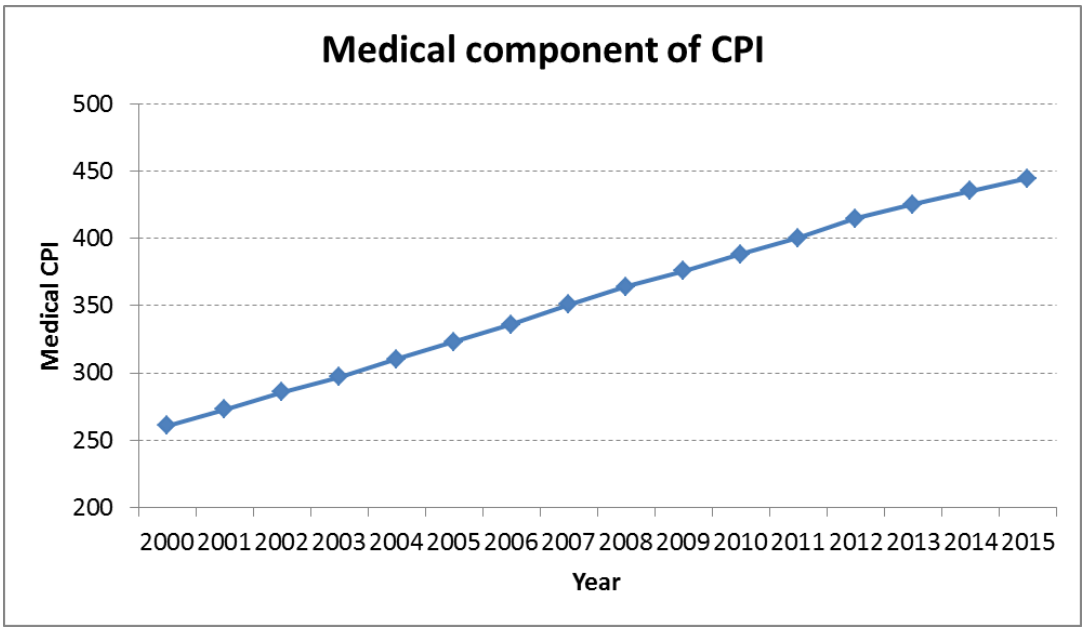
(137) Wing RR, Bolin P, Brancati FL et al. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. *N Engl J Med* 2013;369:145-154.

Appendix

Medical Component of US Consumer Price Index (CPI)[136]

Year	Medical CPI
2015 (first half)	444.65
2014	435.29
2013	425.13
2012	414.92
2011	400.26
2010	388.44
2009	375.61
2008	364.07
2007	351.05
2006	336.2
2005	323.2
2004	310.1
2003	297.1
2002	285.6
2001	272.8
2000	260.8

Note: 1982-84 cost year = 100



[136] http://www.bls.gov/cpi/cpi_dr.htm#2013 , accessed Oct 30 2015

FRAX® 10-year fracture probability charts[137]

Location	Race	Gender	10-year risk table by CRFs, age, and BMI
General bone fracture	Asian	Female	Bone_Asian_F.pdf
General bone fracture	Asian	Male	Bone_Asian_M.pdf
General bone fracture	Black	Female	Bone_Black_F.pdf
General bone fracture	Black	Male	Bone_Black_M.pdf
General bone fracture	Hispanics	Female	Bone_His_F.pdf
General bone fracture	Hispanics	Male	Bone_His_M.pdf
General bone fracture	White	Female	Bone_White_F.pdf
General bone fracture	White	Male	Bone_White_M.pdf
Hip	Asian	Female	Hip_Asian_F.pdf
Hip	Asian	Male	Hip_Asian_M.pdf

[137] World Health Organization Collaborating Centre for Metabolic Bone Diseases, FRAX calculation tool (USA), http://www.shef.ac.uk/FRAX/charts.aspx, accessed November 13, 2015, University of Sheffield, UK

Hip	Black	Female	Hip_Black_F.pdf
Hip	Black	Male	Hip_Black_M.pdf
Hip	Hispanics	Female	Hip_His_F.pdf
Hip	Hispanics	Male	Hip_His_M.pdf
Hip	White	Female	Hip_White_F.pdf
Hip	White	Male	Hip_White_M.pdf

US life tables

Race	Gender	Life Table
Hispanics	Female	Hispanic F.xlsx
Hispanics	Male	Hispanic M.xlsx
Non-Hispanic Black	Female	Non-Hispanic Black F.xlsx

Non-Hispanic Black	Male	Non-Hispanic Black M.xlsx
Non-Hispanic White	Female	Non-Hispanic White F.xlsx
Non-Hispanic White	Male	Non-Hispanic White M.xlsx
Non-Hispanic Other (US national table)	Female	US F.xlsx
Non-Hispanic Other (US national table)	Male	US M.xlsx