HealthPartners Institute

ModelHealth™: Obesity

Microsimulation model technical documentation

Updated: May 2015

Contents

Executive Summary

HealthPartners Institute and ModelHealth™: Obesity-related behavior is a Markov microsimulation model employing annual cycles. It is designed to examine the cost effectiveness of policies targeting obesity through promotion of behavioral change. Specifically, ModelHealth: Obesity presents a unified framework for determining the impact of physical activity and/or dietary interventions on both obesity prevalence and its subsequent cardiovascular disease burden. ModelHealth: Obesity was originally developed with the goal of estimating the health impact and cost-effectiveness of both community interventions recommended in the *Guide to Community Preventive Services* [1] and the clinical obesity screening recommendation of the United States Preventive Services Task Force (USPSTF). Its use and functionality have been expanded to evaluate regional and policy interventions targeting youth and adolescent obesity, community-based physical activity programs, regulation of sugar-sweetened beverages, and food advertising.

ModelHealth: Obesity employs a flexible framework in which the effect of the intervention under analysis is evaluated at the individual level. These individual effects are aggregated to the population or community level. This document presents a description of the model, an overview of its modeling framework, the development of its inputs, a detailed discussion of the modeling framework and embedded algorithms.

The work underlying ModelHealth: Obesity consists of two parts: data and model. The model was originally developed in TreeAge PRO 2009 and was ported to TreeAge 2012 with subsequent updates corresponding to TreeAge 2013, 2014, and 2015. The data underlying this model are the result of an extensive literature search, abstraction, and adjudication process. An exhaustive discussion of that process is beyond the scope of this document. Here, we discuss key results in terms of their parameterization of the ModelHealth: Obesity model. This approach determined key model items such as: disease risk, costs, and intervention effectiveness. Where adequate published estimated were unavailable, primary data analysis was performed using large public-use datasets. These analyses determined initial levels of physical activity, dietary intake, and body mass indexes (BMI) as well as how those three factors evolve over a lifetime.

ModelHealth: Obesity tracks the physical activity, diet, and BMI of an individual over his/her lifetime. For the analysis of many policies, an "energy balance" approach is used to accommodate the heterogeneous reporting of published interventions and effect sizes. Policy effects are modeled as changes in caloric consumption and/or metabolic equivalents. Changes in macronutrient mix
(saturated fat, carbohydrates, etc.) are examined only if clear estimates are available.

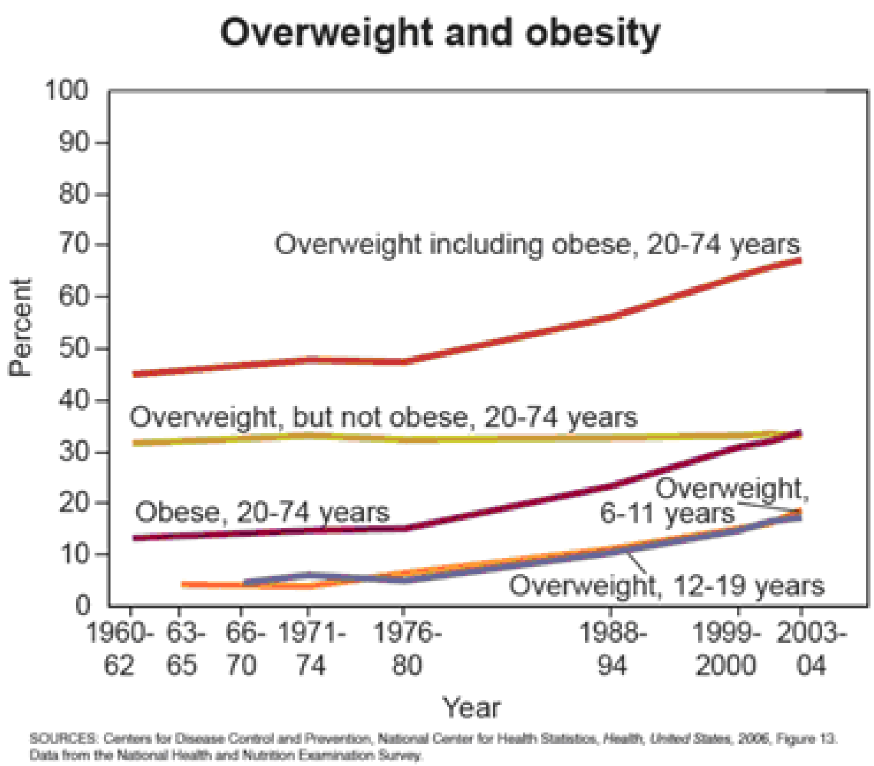
Because policy impacts are assessed at the micro, or individual, level, ModelHealth: Obesity is able to examine impacts on both birth and cross-sectional cohorts.

Introduction

Increased rates of obesity, coupled with patterns of reduced physical activity and poor dietary habits, have become major health concerns in the United States. While other industrialized, and industrializing, nations have experienced increased rates of obesity and also deal with undesirable levels of obesity-related behaviors (poor diet and low physical activity), the United States has the highest reported obesity rate in the world.

In addition, U.S. obesity rates have been increasing. In 1962, the reported U.S. obesity rate was 13%. This has steadily increased and appears to be accelerating. Since 1962, obesity rates have increased to 19.4% in 1997; 24.5% in 2004; 26.6% in 2007; and 33.8% in 2008 [2-4] (Figure 1).

Figure 1: Overweight and Obesity



Considerable differences exist in obesity trends and prevalence according to race, sex, education, and income. Among non-Hispanic black and Mexican-American men, obesity prevalence is negatively associated with income: 28.5% of lower-income (130% of poverty level or below) non-Hispanic black men are obese, compared with 44.5% of those with income at or above 350% of poverty level. Similarly, among higher-income Mexican-American men, obesity rates were 40.8% compared to 29.9% among those with lower income.

This relationship is reversed for women. Twenty-nine percent of women in high-income household are obese; 42% in low-income households are obese. Finally, although there appears to be no relationship between education and obesity prevalence among black and Hispanic men, among non-Hispanic white men, obesity is significantly lower among those with a college education. Similarly, a college degree is associated with lower obesity among many groups of women, specifically non-Hispanic white, non-Hispanic black and Mexican-American women.

Historically obesity primarity afflicted adults; however, this has changed during the past 20 years. Currently, nearly one in three youth are either overweight or obese, compared with 1970s rates. The causes are multifactorial. Overweight youth and adolescents consume between 700-1,000 excess calories a day. They spend, on average, more than six hours a day in sedentary activity (watching TV, playing video games, and using other media). Only 2.1% of high schools, 7.9% of middle schools, and 3.8% of elementary schools provide sufficient physical education. In addition, similar to adults, obesity prevalence exhibits significant racial and ethnic differences.

Obesity is associated with signficant health risks. Obese youth have a 52-60% increased risk of asthma. Further, obesity has been linked to elevated cardiovascular disease risk factors, including gluose intolerance, high cholesterol, and high blood pressure. ModelHealth: Obesity was developed to evaluate the health impact and cost-effectiveness of policy reccomendations attempting to address obesity through behavioral change.

USPSTF recommendations

- The USPSTF recommends that clinicians screen children ages 6 and older for obesity and offer them or refer them to comprehensive, intensive behavioral interventions to promote improvement in weight status. (*B Recommendation)* The USPSTF concludes there is moderate certainty that the net benefit is moderate for screening for obesity in children aged 6 years and older and for offering or referring children to moderate- to high-intensity interventions to improve weight status.[5]
- The USPSTF recommends screening all adults for obesity. Clinicians should offer or refer patients with a body mass index (BMI) of 30 kg/m^2 or higher to intensive, multicomponent behavioral interventions. (*B Recommendation)* The USPSTF found that the most effective interventions were comprehensive and of high intensity (12-26 sessions in a year). Although the USPSTF could not determine the effectiveness of other specific intervention components, most of the higher-intensity behavioral interventions included multiple behavioral management activities, such as group sessions, individual sessions, setting weight-loss goals, improving diet or nutrition, physical activity sessions, addressing barriers, active use of self-monitoring, and strategizing how to maintain changes.[6]
- The USPSTF concludes the evidence is insufficient to recommend for or against behavioral counseling in primary care to promote physical activity. (*I Recommendation)*. It found insufficient evidence to determine whether counseling patients in primary care to promote physical activity leads to sustained increases in activity among adults. Controlled trials of physical activity counseling in adult primary care patients were of variable quality and had mixed results. There were no completed trials with children or adolescents that compared counseling with usual care. Data on the feasibility and potential harms of routine physical activity counseling in primary care settings are limited. [7]
- The USPSTF concludes that the evidence is insufficient to recommend for or against routine behavioral counseling to promote a healthy diet in unselected patients in primary care settings. *(I Recommendation)* The USPSTF found fair evidence that brief, low to medium-intensity

behavioral dietary counseling in the primary care setting can produce small to medium changes in average daily intake of core components of an overall healthy diet (especially saturated fat and fruit and vegetables) in unselected patients[7]. The strength of this evidence, however, is limited by reliance on self-reported diet outcomes, limited use of measures corroborating reported changes in diet, limited follow-up data beyond 6 to 12 months, and enrollment of study participants who may not be fully representative of primary care patients. In addition, there is limited evidence to assess possible harms (see "Clinical Considerations"). As a result, the USPSTF concluded there is insufficient evidence to determine the significance and magnitude of the benefit of routine counseling to promote a healthy diet in adults.

- Although the correlation among healthful diet, physical activity, and the incidence of cardiovascular disease is strong, existing evidence indicates the health benefit of initiating behavioral counseling in primary care to promote a healthful diet and physical activity is small. Clinicians may choose to selectively counsel patients rather than incorporate counseling into the care of all adults in the general population. *(BRecommendation)* Cardiovascular disease (CVD) is the leading cause of death in the United States. Adults who adhere to national guidelines for a healthful diet (1) and physical activity (2) have lower cardiovascular morbidity and mortality than those who do not. All persons, regardless of risk status for CVD, can benefit from improved nutrition, healthy eating behaviors, and increased physical activity.[8]
-

Community Guide recommendations

The *Guide to Community Preventive Services* (*Community Guide*)[1] performed systematic evidence reviews in the following areas: Campaigns and informational approaches; behavioral and social approaches; and environmental and policy approaches. The *Community Guide* reports mixed evidence, which is summarized below by topic:

- Campaigns and informational approaches:
    - *Strong evidence:* Community-wide campaigns are effective in increasing physical activity and improving physical fitness among adults and children[9].
    - *Insufficient evidence:* Stand-alone mass media campaigns increase physical activity at the population level
    - *Insufficient evidence:* Classroom-based health education focused on providing information increase physical activity levels and physical fitness.
- In the area of behavioral and social approaches, the *Community Guide* found:
    - *Strong evidence:* Individually adapted health behavior change programs increase physical activity and improve fitness among adults and children
    - *Strong evidence:* Social support interventions in community settings increase physical activity and improve fitness among adults

    - *Strong evidence:* School-based physical education classes improve both physical activity levels and physical fitness among school-aged children and adolescents
    - *Insufficient evidence:* Family-based social support interventions increase physical activity or improve fitness.
    - *Insufficient evidence:* College-based physical education and health education increased physical activity and fitness.
    - *Insufficient evidence:* Health education classes focused on reduced television viewing and video game playing increase physical activity.
- In the area of environmental and policy approaches, the *Community Guide* found[1]:
    - *Strong evidence:* Enhanced access to places for physical activity increase physical activity and improve fitness.
    - *Strong evidence:* Point-of-decision prompts increase the percentage of people choosing to take stairs rather than an elevator or escalator.
    - *Sufficient evidence:* Design and land use policies and practices facilitate physical activity.
    - *Sufficient evidence:* Urban design and land use policy and practice support physical activity.
    - *Insufficient evidence:* Transportation and travel policy and practice support increased physical activity and improved fitness.

Summary of current evidence and modeling approach

This section contains describes policies currently incorporated into ModelHealth: Obesity. For each policy, a brief literature summary is provided. That summary is then followed by a discussion of how the policy is modeled in ModelHealth: Obesity.

Enhanced School-Based Physical Education Curricula

Brief summary of school-based interventions:

The intervention under consideration corresponds to a broader group of school-based interventions cited by the *Community Guide* in support of its recommendations. As noted by the *Community Guide*, this topic covers a wide variety of interventions and modalities. This heterogeneity of design leads to significant modeling challenges.

Examples of the variety of modalities in and across interventions are:

- **Enhanced PE curriculum:** Changing activities taught in physical education classes (e.g. rule modification, different games, selection of activities transferable outside of class)
- **Comprehensive lifestyle education:** Incorporation of healthy lifestyle choices, nutritional education in both the PE and traditional classrooms
- **Lifestyle integration and outreach:** Incorporating parents and additional family activities into the school-based intervention.

Further, current published intervention studies have a variety of settings, durations, targeted/enrolled populations, and reported outcome measures. Examples of the variety in each of the listed categories are:

- **Setting:** Intervention settings ranged from rural elementary to predominantly minority, urban magnet schools.
- **Duration:** Reported follow-up periods ranged from a maximum of 3 years to a minimum of 12 weeks
- **Enrolled population:** Similar to the diversity of settings, the targeted and enrolled populations varied considerably. Some focused on specific sub-groups (adolescent, inner-city, minority girls) while others had more general representation.
- **Reported outcomes:** Reported outcomes ranged from body composition (Body Fat, BMI, Waist-Hip ratio) to aerobic capacity (VO_2 max) to impact on academic performance. Similarly, reported physical activity changes ranged from reporting of intervention design to surveyed 3-day physical activity recall (PAR).

Finally, while many of the studies reported a dietary component that was either a direct change in diet (e.g. working with food service workers in the school) or a healthy-choice education component, quantified dietary change was limited to only a few studies and limited to fat and/or total caloric intake.

Strategy for modeling enhanced school-based physical education curricula:

The abstracted intervention trials meeting the criteria listed are intended to assess the impact of enhanced school-based physical education over a stated time-window and toward a specific set of stated outcomes.

ModelHealth: Obesity incorporates relationships between diet, physical activity, and youth BMI (as BMI z-score) that were empirically estimated using the continuous NHANES data (2001-2010). Specifically, to accommodate individual variation in baseline physical activity levels, we related a percentage change in physical activity to BMI z-score. In our base case, a youth's BMI percentile (z-score) will change only in response to a behavioral change, and the magnitude of that BMI change is determined by the estimated coefficients listed in **Table 2**. This modeling approach has specific implications. To assess potential impact, we require a clear reporting of the *behavior change* that occurred as a result of the intervention. Physiological outcomes, such as BMI and aerobic capacity, are not directly usable in the current model structure.

The impact of interventions targeting youth (ages ≤ 18) is modeled using an energy balance approach. In this context, any quantifiable dietary change is expressed as a net change in energy expenditure, or metabolic equivalents (METs). This assumption is made for two reasons. First, while some studies mention healthy school lunches and/or healthy snacks as part of a multi-component intervention, they did not systematically track total diet and dietary change. For instance, the studies may report how school lunches were changed, but they did not track whether students subsequently altered their diet outside of school as well. Second, many of the studies focused only on physical education class. Using an energy balance approach allowed us to combine results of diverse interventions that reported results differently.

With the transition to adulthood at age 18, we will assume the entire behavioral change is incorporated into an individuals' baseline level of adult physical activity. Thus, the long-term impact of the intervention will be both a reduction in initial adult BMI as well as a higher level of physical activity on entering adulthood.

The approach is best described as simulating a composite trial that is perfectly implemented. With this approach, for those studies where information is clearly reported, the intended change in physical activity and diet has been averaged. An individual's change in METs is calculated in the following way:

$$\%\ Change\ in\ METs = \left[\begin{matrix} Added\ Number\ of\ Opportunities\ (e.g. PE\ classes) \\ X\ Increased\ Length\ of\ Opportunity\ (e.g\ extended\ class\ period) \\ X\ Intensity\ of\ Activity\ (METs) \end{matrix}\right] \Big/ Current\ PA\ level\ (METs)$$

Two items are important to note in this modeling approach. They are:

- The strategy is potentially optimistic, in that it attempts to capture the intended behavior change and assumes additional activity is not offset by reduced activity during private (unobserved) times.
- The strategy is potentially conservative, in that it assumes no synergy between in-school activity and out-of-school activities. That is, there is not carryover during unobserved times.

Encouraging Stair Use

Brief summary of interventions:

The intervention under consideration corresponds to a group of community-based interventions listed in the *Community Guide* as, Point-of-Decision Prompts to Encourage Use of Stairs[10]. As noted by the *Community Guide*, this topic covers a wide variety of interventions, locations, and modalities. The unifying factors are that the signs, or prompts:

- Inform people about health or weight loss benefits from taking the stairs, and/or
- Remind people already predisposed to becoming more active, for health or other reasons, about an opportunity to do so

The intervention literature cited by the *Community Guide* included both stand-alone interventions as well as interventions used in combination with other factors, such as music to encourage stair use.

The literature in this category was heterogeneous in its design, reporting, and modality. Examples of the variety of modalities in and across interventions are:

- Worksite interventions: Signs and prompts placed at strategic locations throughout the workplace to encourage stair use during the workday. These studies monitored impact by counting the number of times stairs were used, surveying within the workplace, and reported number of flights of stairs.
- Shopping mall interventions: Signs places by escalators and elevators to promote use of stairs as shoppers travel from floor to floor. These studies measured impact by observing stair use before and after sign placement.
- Mass transit: Signs, prompts and "exemplars" (i.e. persons modeling the use of stairs) encouraged the use of stairs at parking lots, mass transit stations, and airports. Impact was measured impact by observing the amount of stair use before and after placement of prompts.

Finally, while most of the studies reported a positive impact in terms of the frequency of stair use, they did not report a net change in physical activity as a result of that increased utilization. For instance, while 10 of 11 studies cited by the *Community Guide* reported more people using the stairs, they did not report how many stairs were climbed.

Strategy for modeling stair use:

The abstracted intervention trials meeting the criteria listed are intended to assess the impact of point of decision prompts on the frequency of stair use. They did not examine the impact on overall physical activity as a result of placing prompts.

ModelHealth: Obesity incorporates relationships between diet, physical activity, and BMI using continuous relationships. These relationships estimate the impact of a given increase in physical activity at an individual, or micro, level by examining it relative to each simulated individual's baseline level of physical activity. Thus, the impact of taking the stairs on the lifetime BMI trajectory of a person who is inactive will be much greater than one who is already active. The impact of encouraging stair use was determined by identifying and accounting for three determining factors: potential effect size, exposure to prompts, and fidelity to intervention.

Determination of potential effect size:

To develop more reliable estimates of the potential change in physical activity attributable to this intervention, data from the Commercial Buildings Energy Consumption Survey (CBECS) Public Use Files[11] were used to specify average building and structural height for shopping malls and worksites. Data from the National Transportation Database [12]were used to determine the average number of stairs at different public transportation sites. Frequency of shopping mall and public transportation use by region was informed by data from the American Community Survey[13]. A summary of building heights by workers and number of elevators within each building in provided in Table 1.

Table 1: Summary of workplace building height

Building Height (floors)	Number of Buildings (1000s)	Number of workers (1000s)	Number of Elevators (1000s)	Average Elevators per Building	Number of Possible Workers Exposed
1	440.90	5,317	-	0.00	-
2	234.34	5,314	33	0.14	741
3	98.32	4,192	50	0.51	2,126
4	29.55	2,014	27	0.90	1,815
5	7.23	1,761	17	2.41	1,761
6	2.82	758	8	2.71	758
7	2.15	679	6	2.73	679
8	1.40	932	8	5.61	932
9	0.72	304	4	5.56	304
10	0.86	333	4	4.32	333
11	0.75	446	3	4.50	446
12	0.98	335	5	5.16	335
13	0.18	139	1	6.20	139
14	0.22	166	2	8.71	166
	Summary				
	Number buildings 2+ floors				**383**
	Average number of floors, if 2+				**2.84**
	Average number workers per bldg, if 2+				**60**

floors	
Average number elevators per bldg, if 2+ floors	**0.57**
Potential number of daily workday stairs	**57**

The *potential effect size attributed to worksite stair use* was determined by calculating the number of workers that could be exposed to a decision prompt (at a building with 2 or more floors *and* with an elevator). For a building with 2-4 floors, the average number of elevators was used to weight down the number of possible workers. The final potential number of daily workday stairs was a weighted average of possible workers exposed and building height assuming a floor height of 12 steps.

The *potential effect size attributable to mass transit use* was determined in a similar manner. The number of workers reporting using mass transit was multiplied by the average number of stairs (including parking ramps) at mass transit stations in their city.

The *potential effect size attributable to shopping malls* was determined by multiplying the average number of trips to shopping malls reported by an assumed number of opportunities to use stairs (3) per trip to the shopping mall.

Determination of final, modeled effect size:

The final effect size incorporated into the model was determined by a two-step process for each of the modalities (Shopping Mall, Worksite, or Mass Transit). First, a random draw determined if the person was adherent to the modality. The likelihood of adherence was abstracted from the intervention literature. For instance, the median stair use in the presence of workplace decision prompts was 20%. Second, given adherence, an individual effect (i.e. number of stairs) was determined by randomly determining the number of stairs possible. For large cities, worksite building height was assumed between 3 and 14 stories. For small cities, it was assumed between 0 and 5 stories. For large cities, mass transit was assumed between 1 and 3 stories. For small cities, it was assumed 1 story (i.e. no stairs). For both small and large cities, shopping mall height was assumed 1 to 3 stories.

Increasing worksite activity

Brief summary of interventions:

This intervention corresponds to a broad group of interventions and programs designed to improve health-related behaviors at the worksite. These interventions combine multiple approaches, including dietary counseling, improving access to healthy foods, educational materials, social support strategies, build environment. They are frequently bundled with larger, more comprehensive, worksite wellness programs targeting multiple health issues such as tobacco use, stress management, and work-family balance.

The literature included in this category was heterogeneous in its design, reporting, and study design. Examples of the variety of designs comprising the evidence base are:

- Randomized trials: The majority of these studies compared intervention groups to control; however, several included an active control arm.
- Self-reported outcomes: Studies in this category based results on changes to previously developed wellness assessment instruments provided as part of a larger worksite wellness program.
- Before-after studies: This group of studies reported changes to a given set of employees at a given work environments coinciding with implementation of a new program, change in build environments, or other initiative.

Similarly, the modality of intervention varied considerably across the evidence base. Examples of these different modalities were:

- Informational strategies: The include providing employees with literature, posting information regarding diet and exercise in strategic locations, and having lectures and seminars.
- Social strategies: These included walking teams with incentives, skill-building activities, and inclusion of social-support networks.
- Environmental changes: These interventions included providing healthier cafeteria options, onsite exercise programs, and subsidizing gym memberships.

Strategy for modeling interventions:

The abstracted intervention trials report impact in varied ways. The three most common outcomes were BMI, weight, and percentage of body fat.   ModelHealth: Obesity incorporates relationships between diet, physical activity, and BMI using continuous relationships. These relationships estimate the impact of a given increase in physical activity at an individual, or micro, level by examining it relative to each simulated individual's baseline level of physical activity. Thus, in order to determine the impact of this intervention in a comparable manner to the others that were modeled, the intervention impact needed to be expressed in terms of a change in energy balance.

To specify the modeled impact, specifics regarding the worksite policies and interventions were abstracted. These were used to specify an average percentage change in energy balance that could be attributable to any combination of changes to physical activity or diet.

Community-Scale Urban Design and Land Use Policies and Practices (Trails and Land Use Policies)

Brief summary of interventions:

All eligible studies are observational, cross-sectional, and typically do not report outcomes in a form that readily demonstrates how this class of interventions affects an individual's body composition (i.e., outcomes are generally minutes of walking, rates of active commuting, etc.). Most of the studies within this class of interventions compared the behavior of residents in automobile-oriented (or suburban) communities with those in traditional (or urban) communities. Overall, the median improvement in some aspect of physical activity (e.g., number of walkers or bicyclists) was 161%. Additional benefits that may have been brought about by these interventions included:

- Improvements in green space
- Increased sense of community and decreased isolation
- Increased consumer choice for places to live
- Reduced crime and stress
- Increased walking and bicycling on urban streets, although beneficial, also pose the risk of increased injury to pedestrian or cyclist, because of increased exposure to motor vehicles.

Finally, the details of the relationship between healthier people in a neighborhood and access to facilities have never been established. It is unknown how much of the association is because increased access results in individuals within the community becoming healthier, and how much is because individuals with healthy behaviors seek out neighborhoods with better access to facilities.

Strategy for modeling interventions:

The studies under this umbrella are diverse, and the effects are heterogeneous, such that a single quantitative summary across estimates has never been attempted. No group-randomized studies (in which some communities are randomly selected for improvement, and control communities are randomly selected for future design improvements) have been attempted.

The abstracted intervention trials report impact in a variety of ways. The most common was frequency of use following the building of a trail or enhancement to a current urban environment. ModelHealth: Obesity incorporates relationships between diet, physical activity, and BMI using continuous relationships. These relationships estimate the impact of a given increase in physical activity at an individual, or micro, level by examining it relative to each simulated individual's baseline level of physical activity. To determine the impact of this intervention in a comparable manner to the others that were modeled, the intervention impact needed to be expressed in terms of a change in energy balance. The final effect size was determined by assuming use of a two-mile trail, three times per week.

Description of the ModelHealth: Obesity microsimulation

Overview

ModelHealth: Obesity is a Markov microsimulation. A Markov microsimulation is a model in which simulated individuals/agents, age over time while facing period-specific "risks" of changing health behaviors and/or health outcomes. In each cycle (currently, the equivalent of one year), individuals may remain in their current state or transition to a different one. When modeling obesity and physical activity, it is difficult to define a set of discreet, mutually exclusive states, because both are continuous measures that dynamically change over time. The model tracks these factors as continuous measures at the individual level. Classification into categories, or states, is only done for purposes of reporting or assigning state-specific costs. In ModelHealth: Obesity, each individual agent's age, diet, physical activity, body mass index, and health status is tracked through time. Each individual's associated costs and disease outcomes are determined as a function of tracking those characteristics. All simulated agents in the model are independent: the actions of one individual do not affect those of another.

When simulated agents are introduced into the model, they are assigned to a population strata, or cohort. Each cohort is defined by a unique combination of initial age, sex, and ethnicity. Each cohort is equally sized in the model. These demographic characteristics specify the distribution from which initial BMI, diet and level of physical activity are drawn. These demographics also affect how these characteristics evolve over time, as well as outcomes such as disease risk.

Estimates from the models come from aggregating across agents the state-specific costs and benefits resulting from their individual actions. As noted in Figure 2, each user is assigned to a population stratum at initiation. To create estimates specific to a location or region, these strata can be dynamically weighted to specific areas of interest without requiring extensive reconfiguration and re-running of the model.

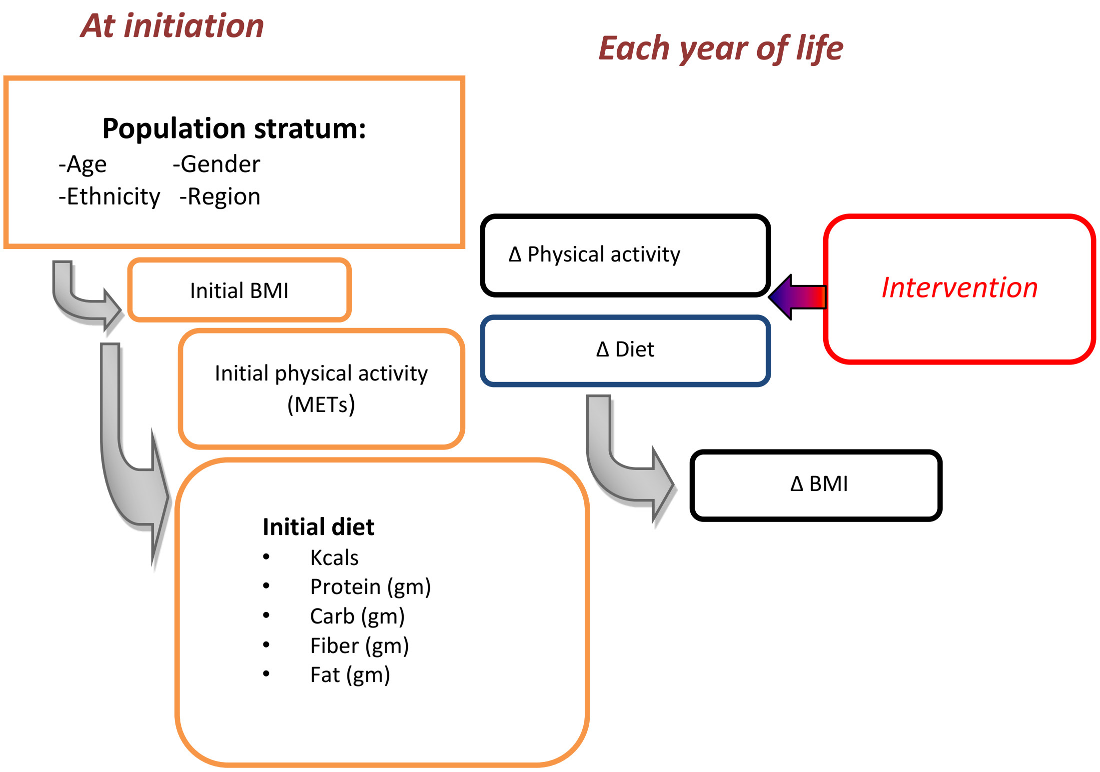
Markov microsimulation models do not require population-wide data and can be re-calibrated to specific sub-populations or targeted intervention groups. These models can become very complex because every initial action, resulting interactions, and ultimate health state must be explicitly modeled. Thus, care must be taken to protect against unintended linkages in and between agents and across interventions. Our approach to this is discussed in later sections. The next section presents the general model structure. This is followed by a more detailed discussion of the two modules comprising ModelHealth: Obesity: the youth module and the adult module. Then, details are provided on how ModelHealth: Obesity is implemented.

Model structure

The model starts by generating a population of simulated individuals, or agents. Each agent is assigned to a population strata defined by a unique combination of initial age, sex, race and ethnicity. In the model, each population strata is equally sized at initiation.

Once a simulated population is created, agents are aged through life in yearly increments, or cycles. During each cycle, diet, physical activity, BMI, cardiovascular risk factors and disease risk are evaluated.

Figure 2. Flow of ModelHealth: Obesity



Two discreet processes occur in ModelHealth: Obesity: initiation and progression. The following section discusses how agents are initialized into the model. A subsequent section shows how agents are aged through the youth and adult modules.

Model initiation

At model initiation, demographics (initial age, sex, race/ethnicity) are initialized according to the current stratum. Then, health behaviors are initialized according to those demographic values. This begins with the assumed time-invariant demographics of lifetime educational achievement and family income. Although there is some evidence suggesting a link between appearance and income – and obesity is a factor that can affect appearance – the current focus of analysis is not on evaluating the simultaneity of this relationship. Table 2 lists the agent-level parameters tracked by ModelHealth: Obesity, their conditioning factors, assumed distribution across the population, and the timing of initialization in the model.

Table 2: Agent-level parameters in ModelHealth: Obesity

Agent-level parameter	Conditioning factors	Population distribution	Time of initialization	Fixed or variable	Source for baseline
Age	N/A	N/A	Model introduction	Variable	Current Strata*

Sex	N/A	N/A	Model introduction	Fixed	Current Strata*
Race/Ethnicity	N/A	N/A	Model introduction	Fixed	Current Strata*
Region	N/A	N/A	Model introduction	Fixed	Current Strata*
Education	Age, Sex, Race, Region	Dirichlet	Model introduction	Fixed	NCES, BLS[14,15]
Income	Age, Sex, Race, , Region Education	Dirichlet	Model introduction	Fixed	BLS. SIPPS[15,16]
Insurance Status	Age, Sex, Race, , Region, Education	Dirichlet	Model introduction	Fixed	BLS. SIPPS
BMI	Age, Sex, Race, Diet, PA	Gamma	Model introduction	Variable	BRFSS/NHANES[17,18]
Calories	Age, Sex, Race, BMI	Gamma	Model introduction	Variable	NHANES
Protein** (gms/day)	Age, Sex, Race, BMI	Normal	Model introduction	Variable	NHANES
Carbohydrates** (gms/day)	Age, Sex, Race, BMI	Normal	Model introduction	Variable	NHANES
Fiber** (gms/day)	Age, Sex, Race, BMI	Normal	Model introduction	Variable	NHANES
Fat** (gms/day)	Age, Sex, Race, BMI	Normal	Model introduction	Variable	NHANES
Sugar** (gms/day)	Age, Sex, Race, BMI	Normal	Model introduction	Variable	NHANES
Physical Activity** (METs/day)	Age, Sex, Race, BMI	Gamma	Model introduction	Variable	BRFSS/NHANES
Tobacco Smoke	Age, Sex, Race,	Binomial	Age 18 or introduction	Variable	NHIS/BRFSS[17,19]
SBP**	Age, sex, race, BMI	Gamma	Age 18 or introduction	Variable	NHANES[20]
DBP**	Age, sex, race, BMI	Gamma	Age 18 or introduction	Variable	NHANES
LDL**	Age, sex, race/ethnicity, BMI	Gamma	Age 18 or introduction	Variable	NHANES
HDL**	Age, sex, race/ethnicity, BMI	Gamma	Age 18 or introduction	Variable	NHANES

*ModelHealth: Obesity simulates equally sized cohorts, or strata, defined by unique combinations of initial_age, sex, ethnicity, and census region. These strata are then weighted to the analytic frame of interest (US population, specific country, city, etc.) to create tailored estimates of effect size.

**The primary focus of the model is on BMI and BMI change. Thus, all factors are conditioned on BMI to create consistency at model initiation.

Conditioned on the fixed basic demographic factors of age, sex, region and race/ethnicity; an agent's lifetime educational achievement is determined. These socio-economic factors are then used to determine each simulated agent's insurance status as noted in the insurance module technical documentation.

Time varying agent-level factors are initialed by random draws from a joint probability distribution. Depending on the initial age of the agent's stratum, these joint distributions are constructed in one of two ways. For agents whose initial age is less than 18, the youth module is used. In the youth module, the fixed conditioning factors are sex, ethnicity and initial age. The time-varying factors that are initialized and tracked in the youth module are: BMI, physical activity level (expressed in METs), and diet (total kilocalories and grams of protein, carbohydrates, dietary fiber, and fat). For individuals whose initial age is 18 or older, the adult module is used.

The final step in initializing an agent is to determine initial values for cardiovascular disease risk factors. This is done only in the adult module. For agents with an initial age younger than 18, initial values of cardiovascular risk factors are set when they reach age 18. For agents who enter the model as adults these are set at initiation into the model. Cardiovascular risk factors are set according to distributions conditioned on age, sex, race, tobacco smoke status, and BMI.

The youth module

Youth obesity is a relative concept. Considerable changes in BMI are a normal part of growth and development. A BMI of 20 is considered high for a 6-year-old, but normal for a 16-year-old. Thus, obesity and obesity risk are determined by comparing a youth's BMI to standardized BMI growth charts using standardized percentiles. Youth obesity is defined as a BMI at or above the 95th percentile, and a child is considered at risk of obesity if its BMI is between the 85th and 95th percentiles.

Overview of modeling baseline youth BMI values

The youth module of ModelHealth: Obesity reflects this by modeling the change in BMI over time in terms of percentiles of the standardized BMI distribution conditioned on agent age, sex, ethnicity, physical activity, and diet. Initial BMI is determined by a random draw from a distribution fit to continuous NHANES data from 2001-2010 and conditioned on age, sex, ethnicity, and the time-variant behaviors of physical activity and diet. In each cycle, BMI percentile is adjusted to account for three factors: 1) natural changes in BMI with age, 2) population BMI trends corresponding to age, and 3) individual variation.

Each of these factors reflects important aspects of youth BMI (Figure 3). The first captures the relationship between BMI and BMI percentile reflected in the standardized growth charts. For example, the BMI of an individual at age 2 who is at the 60th percentile of the BMI distribution will be higher than their BMI at age 5 if he/she remains at the 60th percentile. The youth module of ModelHealth: Obesity captures this by adjusting BMI each cycle according to trends contained in the standardized growth charts. The second factor, population BMI trends, regards observed trends in youth obesity related to age. For instance in 2012, 18% of children aged 6-11 were obese; this rate increased to 21% among adolescents[18]. To capture this age-related upward trend, each simulated agent's BMI is adjusted. This secondary BMI adjustment reflects the third factor surrounding youth BMI, each person's path is unique.

Factor 1: Natural BMI changes with Age

Age- and sex-specific BMI growth charts published by the Center for Disease Control provide monthly BMI percentiles from ages 2 (24 months) through age 20 (240 months). These charts also provide age (monthly) and sex-specific level (L), mean (M), and scale (S) parameter values to translate an observed BMI to a BMI z-score. Comparing a youth's BMI to established growth charts identifies those who are, or who are at risk of, obesity using the following formula:

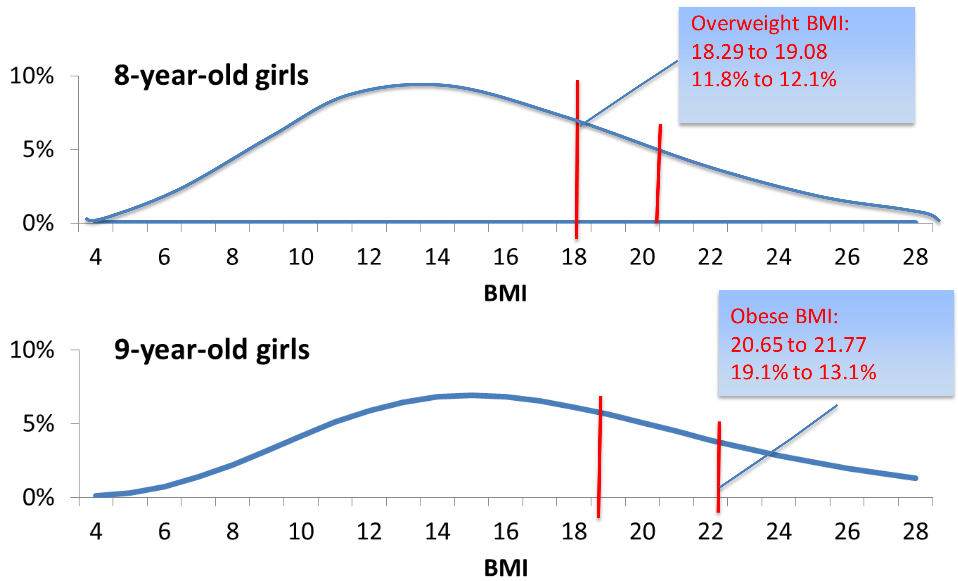
$$BMI_{\ Z-score} = \left[\left(\frac{BMI}{M_{age,sex}}\right)^{L_{age,sex}} - 1\right] \Big/ L_{age,sex} * S_{age,sex}$$

This BMI z-score equation was algebraically manipulated to solve for BMI in the following way:

$$BMI = M_{age,sex} * \left(L_{age,sex} * S_{age,sex} * BMI_{\ Z-score} + 1\right)^{1/L_{age,sex}}$$

This second expression allows the model center on a given $BMI_{Z\text{-score}}$ while still evolving that person's actual BMI by age.

Figure 3: Illustration of baseline BMI change in youth



Factor 2: Population BMI trends

Recent studies using a longitudinal dataset of United States youth provided by the National Center for Educational Statistics (NCES) indicate there are certain ages where the incidence of youth obesity is more prevalent [21]. This implies an upward drift in youth BMI z-scores over time as an increasing number of youth's BMI exceed the corresponding age and sex-specific 95th percentile of the CDC growth chart. Unfortunately, the NCES data did not provide as comprehensive data on physical activity and diet as NHANES. Thus, NCES data were used to calibrate incidence of overweight and obesity, but NHANES were used to link to BMI to physical activity and diet. The calibration is discussed in a later section.

Factor 3: Individual Variation

While the general pattern, or center, of the age-related population-wide upward drift in overweight and obesity is informed by data and published studies. Within ModelHealth: Obesity, each person's path is unique. After accounting for the baseline $BMI_{Z\text{-score},}$ a secondary BMI adjustment is determined by a random draw from a distribution defined by the difference between the population-wide BMI distribution at their current age and the next year's population-wide BMI distribution. This process is illustrated in the following figure.

The definition of overweight is the 85th percentile of the BMI growth charts. For eight-year-old girls, this value is a BMI of 18.29, and 11.8% of all 8-year-old girls have a BMI at or above this value. For nine-year-old girls, the BMI corresponding to the 85th percentile is 20.65, and 19.1% of all 9-year-old girls have a BMI at or above this value. Also, as shown in the figure, the BMI distributions for both ages are skewed. Within the ModelHealth: Obesity, a method-of-moments estimation is used to parameterize age-specific gamma distributions. In this example, these distributions represent 8- and 9 year-old girls, respectively.

By comparing these two distributions, the age-based trends (i.e. increasing rates by age) of obesity and overweight are captured. Each modeled agent is the assigned a unique year-to-year change through a random draw from the difference between these two gamma distributions.

Determining baseline physical activity and diet from NHANES data

To have the largest sample possible when developing relationships between behavioral change and youth $BMI_{Z\text{-Score}}$, several years of NHANES data were combined. **Table 3** summarizes the NHANES variables that were used, by study year. For variables measuring youth physical activity, the value used to convert self-reported minutes of physical activity to metabolic equivalents is noted in the corresponding formula.

Table 3 Summary of NHANES datasets, variables* and formula used

Description	1999-2000	2001-2002	2003-2004	2005-2006+	2007-2008+	2009-2010+
Sex	riagendr	riagendr	riagendr	riagendr	riagendr	riagendr
Age in years	ridageyr	ridageyr	ridageyr	ridageyr	ridageyr	ridageyr
Age in months at exam	redageex	redageex	redageex	redageex	redageex	redageex
Race/ethnicity	ridreth1	ridreth1	ridreth1	ridreth1	ridreth1	ridreth1
Education for youth (grade level)	dmdeduc3	dmdeduc3	dmdeduc3	dmdeduc3	dmdeduc3	dmdeduc3
Household income	indhhinc	indhhinc	indhhinc	indhhinc	indhhin2	indhhin2
Body Mass Index (kg/m**2)	BMXBMI	BMXBMI	BMXBMI	BMXBMI	BMXBMI	BMXBMI
Do you now smoke cigarettes	SMQ040	SMQ040	SMQ040	SMQ040	SMQ040	SMQ040
During the past 30 days, on how many days did you smoke cigarettes?	SMQ640	SMQ640	SMQ640	SMD641	SMD641	SMD641
How old were you when you smoked a whole cigarette for the first time?	smq630	smq630	smq630	smd630	smd630	smd630
How old when [you/s/he] first started to smoke cigarettes fairly regularly?	SMD030	SMD030	SMD030	SMD030	SMD030	SMD030

METs from biking physical activity (bike_METs)	pad080	pad080	pad080	pad080	paq640*pad645*4/Week	paq640*pad645*4/Week
ME Ts from yardwork home tasks or other activity	paq050q*pad080*4 METs	paq050q*pad080*4 METs	paq050q*pad080*4 METs	paq050q*pad080*4 METs	NA (as 99999)	NA (as 99999)
METs from general activity (gen_METs)	pad120*pad160*4.5 METs/30.5	pad120*pad160*4.5 METs/30.5	pad120*pad160*4.5 METs/30.5	pad120*pad160*4.5 METs/30.5	vig_gen_METs = 8METs*paq610*pad615/wk mod_gen_METs = 4 METs*paq625*pad630/wk gen_METs = vig_gen_METs + mod_gen_METs	vig_gen_METs = 8 METs*paq610*pad615/wk mod_gen_METs = 4 METs*paq625*pad630/wk gen_METs = vig_gen_METs + mod_gen_METs
METs from strengthening activities (mus_METs)	pad460*60*4 METs/30.5	pad460*60*4 METs/30.5	pad460*60*4 METs/30.5	pad460*60*4 METs/30.5	NA (as 99999)	NA (as 99999)
METs during TV or inactivity (tv_METs)	30*1.2 METs	30*1.2 METs	30*1.2 METs	30*1.2 METs	pad590*60*1.2 METs	pad590*60*1.2 METs
METs from play or recreational activity (play_METs)	paq560/wk 60*7 METs	paq560/wk 60*7 METs	paq560/wk 60*7 METs	paq560/wk 60*7 METs	vig_play_METs = 8 METs*paq655*pad660/ wk mod_play_METs = 4 METs*paq670*pad675/ wk play_METs = vig_play_METs+mod_play_METs	paq706
METs from IAF physical activity	padMETs*paddurat*padtimes/30	padMETs*paddurat*padtimes/30	padMETs*paddurat*padtimes/30	padMETs*paddurat*padtimes/30	NA (as 99999)	NA (as 99999)
DAILY METs (From above variables)	bike_METs + yard_METs + gen_METs + mus_METs + tv_METs + play_METs + iaf_METs	bike_METs + yard_METs + gen_METs + mus_METs + tv_METs + play_METs + iaf_METs	bike_METs + yard_METs + gen_METs + mus_METs + tv_METs + play_METs + iaf_METs	bike_METs + yard_METs + gen_METs + mus_METs + tv_METs + play_METs + iaf_METs	bike_METs + gen_METs + tv_METs + play_METs	bike_METs + gen_METs + tv_METs + play_METs
Energy (kcal)	DRXTKCAL	DRXTKCAL	DR1TKCAL+DR2TKCAL/2	DR1TKCAL+DR2TKCAL/2	DR1TKCAL+DR2TKCAL/2	DR1TKCAL+DR2TKCAL/2
Protein (gm)	DRXTPROT	DRXTPROT	DR1TPROT+DR2PROT/2	DR1TPROT+DR2PROT/2	DR1TPROT+DR2PROT/2	DR1TPROT+DR2PROT/2

Carbohydrate (gm)	DRXTCARB	DRXTCARB	DR1TCARB+DR2TCARB/2	DR1TCARB+DR2TCARB/2	DR1TCARB+DR2TCARB/2	DR1TCARB+DR2TCARB/2
Total fat (gm)	DRXTTFAT	DRXTTFAT	DR1TTFAT+DR2TTFAT/2	DR1TTFAT+DR2TTFAT/2	DR1TTFAT+DR2TTFAT/2	DR1TTFAT+DR2TTFAT/2
Total saturated fatty acids (gm)	DRXTSFAT	DRXTSFAT	DR1TSFAT+DR2TSFAT/2	DR1TSFAT+DR2TSFAT/2	DR1TSFAT+DR2TSFAT/2	DR1TSFAT+DR2TSFAT/2
Total monounsaturated fatty acids (gm)	DRXTMFAT	DRXTMFAT	DR1TMFAT+DR2TMFAT/2	DR1TMFAT+DR2TMFAT/2	DR1TMFAT+DR2TMFAT/2	DR1TMFAT+DR2TMFAT/2
Dietary fiber (gm)	DRXTFIBE	DRXTFIBE	DR1TFIBE+DR2TFIBE/2	DR1TFIBE+DR2TFIBE/2	DR1TFIBE+DR2TFIBE/2	DR1TFIBE+DR2TFIBE/2
Total sugars (gm)	NA (.)	DRXTSUGR	DR1TSUGR+DR2TSUGR/2	DR1TSUGR+DR2TSUGR/2	DR1TSUGR+DR2TSUGR/2	DR1TSUGR+DR2TSUGR/2

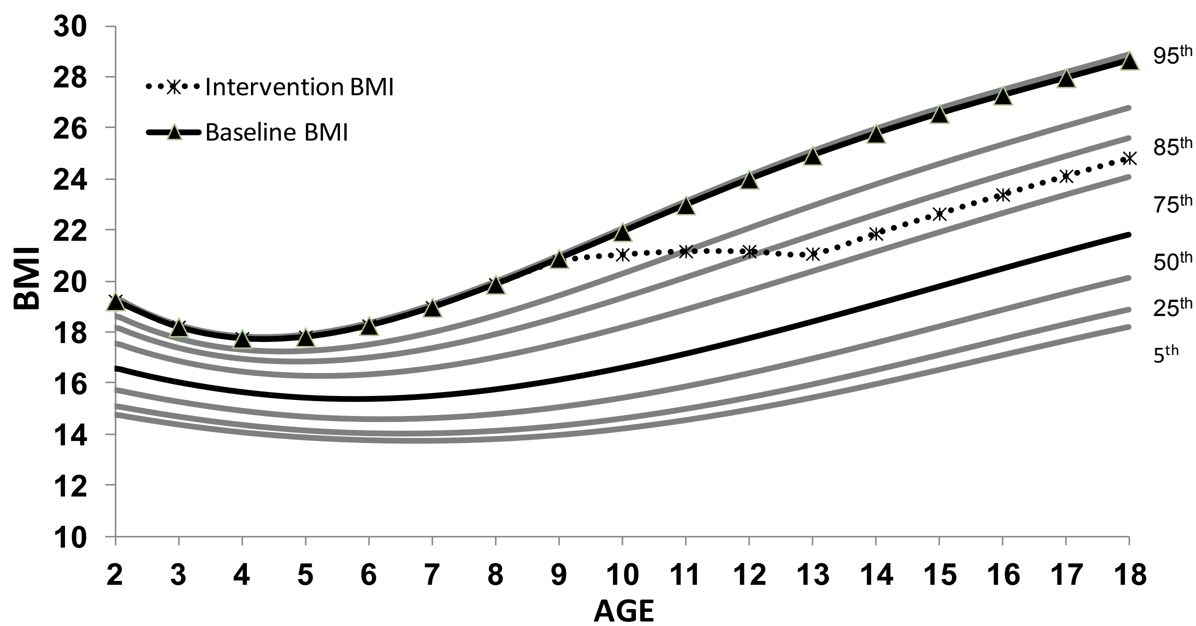
*Survey questions and variable names varied across study years.

+A two-day food frequency questionnaire was used starting in 2008.

Determining the impact of an intervention on youth BMI values

Figure 4 illustrates the BMI growth path of a potential agent and plots that agent's BMI against CDC BMI-for-age growth charts. This chart pertains to the life of a simulated non-Hispanic white male from ages 2-18 and illustrates how baseline BMI and the impact of interventions are modeled. The individual is introduced into the model at age 2 with a BMI of 19.5. While this BMI corresponds to the 84th percentile of a BMI distribution conditions on sex and race, it corresponds to the 89th percentile of the population-wide distribution represented by the CDC BMI growth charts.

Figure 4: BMI Growth and Policy Impact



At age 3, his BMI is decreased to 18.5, which corresponds to the same percentile of the sex and ethnicity conditioned distribution at age 3 as a BMI of 19.5 did at age 2. This process is repeated until age 18 and, assuming no significant behavioral changes, results in the BMI path shown by the black line with triangles.

The dotted line in Figure 4 reflects the impact of an intervention. In this instance, the intervention was a school-based physical activity intervention initiated at age 9 and continuing until age 18 the year of graduation. Thus the intervention was continuously applied and resulted in a sustained, elevated level of physical activity. As a result of the increased physical activity, the individual's BMI at age 10 is reduced from 21.9 to 21, and the corresponding change in BMI percentile is illustrated. At age 11, there is an additional decrease in BMI and the pattern continues until age 14 where a BMI path consistent with the increased level of physical activity is reached.

A multivariate, weighted regression estimated the relationship between youth BMI percentile (i.e. BMI z-score) and targeted youth behaviors (physical activity and diet expressed as an energy balance). This analysis used a log specification in order to estimate BMI z-score sensitivity to a percentage change caloric consumption and physical activity, respectively. In addition, the estimated BMI z-score elasticity to discreet changes to macronutrients whose daily consumption is measured within NHANES by gram was estimated. The macronutrients are: protein, carbohydrates (complex and simple), fiber, saturated fat and trans fat. The following general specification was used:

$$BMI_{Z-Score} = \alpha + \beta_{ag}{*}age{+}\beta_r{*}race{+}\beta_s{*}sex{+}\beta_M{*}ln(METS){+}\beta_K{*}ln(Kcal)$$
$$+\beta p{*}protein{+}\beta c{*}carb{+}\beta SF{*}SFat{+}\beta TF{*}TFat{+}\beta sug{*}sugar$$
$$+\beta_f{*}fiber{+}\beta{*}INTERACTIONS$$

The term "INTERACTIONS" represents several interactions found significant at the .05 level. The monotonic transformations applied to both physical activity (measures in metabolic equivalents or METs) and kilocalories allows the estimates of their corresponding coefficients to be interpreted as the change in BMI z-score that corresponds to a given percentage change in either physical activity or total energy consumption. These two later estimates drive changes observed within the Community Health Advisor as a net energy balance approach is used to model the impact of interventions.

Table 4 contains estimated results from the final specification that was incorporated into the ModelHealth: Obesity model. As noted, the data used came from the continuous NHANES data.

Table 4: BMI Transition Equations for Youth Module

Included Factor	Estimated Value	[95% Conf. Interval]	
Constant	1.572351	-4.43253	7.577234
2003-04	0.1072308	-0.01639	0.230848
2005-06	0.0319784	-0.10381	0.167765
2007-08	0.0226489	-0.10495	0.150247
2009-10	0.0991812	0.001308	0.197054
Black	0.2348633	0.162821	0.306905
Hispanic	0.256809	0.185804	0.327814
Other	-0.20231	-0.31039	-0.09424
Age	0.2679141	-0.1346	0.670423
Age: 6-12	0.0802496	-0.02549	0.185988
Log Kilocalories	0.577096	0.222726	0.931466
xAge	-0.086279	-0.11057	-0.06199
Protein (gm)	0.0050325	0.003479	0.006586
Carbohydrates (gm)	-0.000527	-0.0014	0.000348
Fiber (gm)	-0.011272	-0.01686	-0.00568
Sugar (gm)	0.0012338	0.000205	0.002263
Saturated Fat (gm)	-0.007902	-0.013	-0.00281
Trans Fat (gm)	0.0109678	0.005468	0.016467
Physical Activity (Log METs)	-0.810288	-1.56747	-0.05311
xAge	0.0530394	0.001293	0.104786

In addition to the specific factors indicated in the prior general specification, some additional control predictors were added to improve model fit. A separate intercept was estimated for each year of the NHANES survey. This was done to capture mean level shifts due to slight differences attributable to question rewording or measurement changes. For instance, in 2008-2009, a two-day food frequency questionnaire replaced a one-day questionnaire. Similarly, questions regarding types of physical activity varied slightly from year to year.

We also included an indicator, or dummy, variable for ages younger than 12. This variable was added for similar reason as the survey year intercepts. Prior to age 12, physical activity and diet were measured through parental interview. Certain questions such as the detailed daily activity questionnaire were not completed. This dummy variable was included to adjust for the slightly different measurement criteria. No significant interactions with either physical activity or energy consumption (kilocalories) were found.

The Adult Module

Initiation or transition into the adult model

Simulated individuals may be introduced into the adult module in two ways: either they are transitioned from the youth module or they are initialized as adults, depending on their initializing stratum. For those initialized as adults, initial BMI is determined by a random draw from an age, sex, ethnicity distribution fit to 2009 BRFSS data[17]. (The 2009 self-report BRFSS data were adjusted for self-report bias using established methods. BRFSS is used due to its larger sample size and region-specific data. Alternative specifications, using NHANES data, are also available.)

Those transitioning from the youth model do so at age 18. As Figure 1 indicates, both the CDC BMI growth charts and the NHANES data allow tracking of BMI growth percentile (standardized BMI z-score) up to age 20. These three years of overlap are used to calibrate each agent's BMI growth path. This is done in the following steps:

- **Creation of adult behavioral profile:** At age 18, a set of alternative behavioral values (BMI, Diet, and Physical Activity) consistent with the self-reported BRFSS data.
- **Simultaneous tracking:** From age 18-20, each simulated agent is tracked using both the youth and adult modules.
- **Comparison of changes:** BMI values at the end of each cycle are compared and averaged.
- **Final transition:** At age 20, agents are fully transitioned to the adult model.

When an agent transitions into the adult model at age 18, new values for diet and physical activity consistent with the BRFSS data are generated. This is done using predictive equations that adjust for age, sex, ethnicity, and current BMI. BMI is incorporated into these equations to produce a set of behavioral values (diet and physical activity) consistent with that agent's current BMI and the measurement scales drawn from the BRFSS data.

The reason for this transitional period is to protect against any dramatic BMI shifts attributable to the model algorithm or differences in underlying data sources and not actual behavior change. This potential for unintended BMI shifts exists, because of differences in the two data sources supporting the youth and adult modules, respectively. The NHANES data, which supports the youth module, is a smaller dataset, but it contains clinically measured BMI values, and its dietary measures come from either a one or two day food frequency questionnaire. In comparison, the BRFSS data, which supports the adult model, is considerable larger but contains self-reported BMI values. Similarly, its dietary measures are drawn from a set of self-reported questions. Both surveys also use a different set of questions to measure physical activity and exercise.

BMI Progression in the Adult module

BMI progresses in the adult module using a two-step process. Empirically, this process is estimated by use of a multinomial hurdle model. In this model, each simulated individual's BMI change category is first determined. Individuals are assigned to one of four BMI change categories: Weight Loss, Stable Weight, Drifter, and Weight Gain. The probabilities for a specific individual are determined by a random draw from a distribution derived from the results of a multivariate, multinomial logistic regression adjusting for age, sex, ethnicity, physical activity, and diet. Table 5 contains estimates from this estimation. (Consistent with BRFSS, the dietary components of these equations represent servings of certain food groups and not specific macronutrient quantities. These relationships are currently being re-worked using NHANES data to represent macronutrient quantities similar to the equations of the Youth module.)

Table 5: BMI Transition Equations for Adult Module

	Drifter		Weight Gain		Weight Loss	
Parameter	**Value**	**Std. Error**	**Value**	**Std. Error**	**Value**	**Std. Error**
Intercept	-0.29023	0.827943	0.647805	0.631845	-2.03113	0.73879
Male	-0.38772	1.363775	6.05225	0.88177	4.335525	1.036675
Black	0.104385	0.047082	0.756035	0.037636	0.448183	0.037685
Hispanic	-0.03022	0.04833	0.422473	0.03901	0.312098	0.03904
Other	0.016285	0.05025	0.322985	0.042525	0.247508	0.042442
Age	-0.0033	0.037157	0.012701	0.028382	0.016952	0.033092
AGE>23*	0.010765	0.039383	-0.05452	0.030191	-0.04064	0.034928
Age>40*	-0.00972	0.007928	0.032945	0.006696	0.014804	0.006883
Age>50*	-0.00346	0.006543	-0.02251	0.006053	0.003847	0.005954
Age*Male	-0.01407	0.060663	-0.28952	0.039534	-0.21873	0.04625
Age>23*Male	0.010982	0.062282	0.287768	0.040991	0.224463	0.047655
Age>50*Male	0.014695	0.007433	0.013101	0.006588	0.010079	0.006263
Prior_BMI	-0.0093	0.002401	-0.01539	0.002205	0.059824	0.001917
Ln(METs)	4.89E-05	2.66E-05	-6E-05	2.55E-05	-0.00022	2.29E-05
Ln(METs)*PriorBMI	-1.9E-06	9.81E-07	2.91E-07	9.33E-07	8.61E-06	7.76E-07

*These variables represent linear splines specified as Maximum of Age-XX or 0.

The estimates of Table 5 are based on using an unchanged BMI as the reference category. They parameterize the probability of an individual falling into one of four weight change categories. Multiple non-linearities in the relationship between BMI change category and age were found when developing this empirical model. These shifts in the relationship were incorporated as a series of interactions or knots occurring at ages 23, 40, 50 and 60. For males, additional interactions were added at age 23 and 50. Figure 5 illustrates the distribution of BMI change categories by age.

Figure 5: Distribution of Adult BMI Change Categories

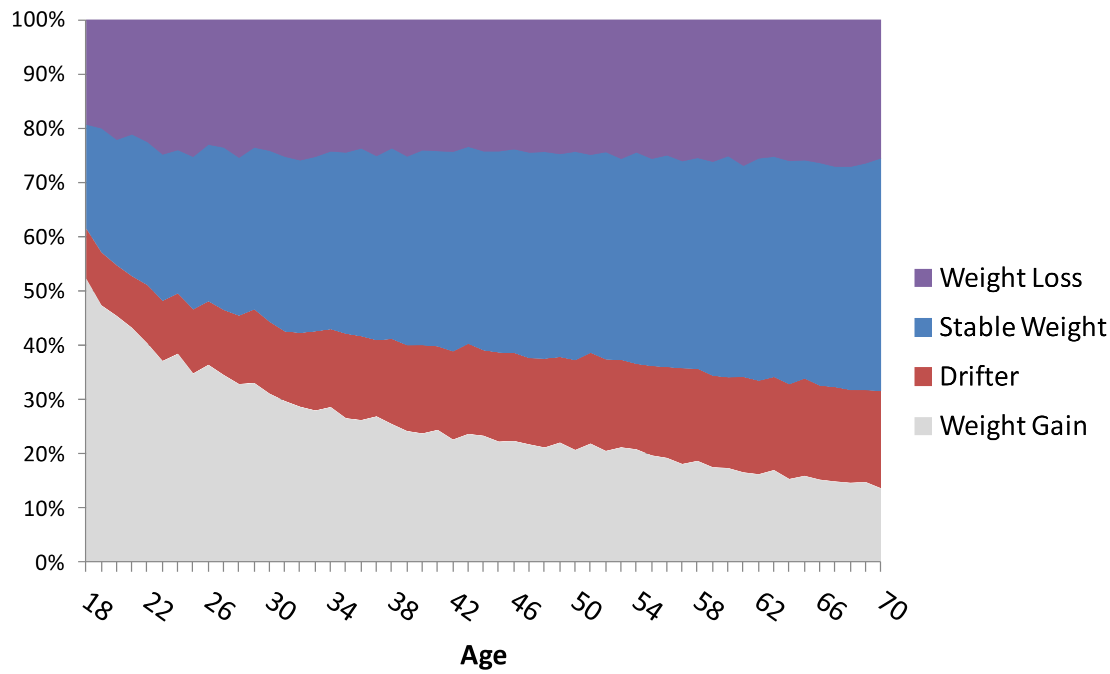


Figure XXXX: Distribution of Adult BMI Change Categories

Once an agent's BMI change category is identified, the magnitude of that agent's BMI change is determined by the predicted values from a continuous, multivariate regression that again adjusts for fixed and time-variant factors. Four separate regression models were developed, one for each BMI change category. As the magnitude of BMI change was typically skewed in each category, a generalized linear model with gamma distributed errors and identity link was used. The estimated coefficients from each model are contained in following table 6.

Table 6: Magnitude of BMI Change in Adult Module

	Drifter		Weight Gain		Weight Loss	
Parameter	**Value**	**Std. Error**	**Value**	**Std. Error**	**Value**	**Std. Error**
(Intercept)	4.06E-04	9.18E-03	4.15E+00	2.57E-01	-3.11E-01	1.07E-01
AGE	-6.71E-04	3.09E-04	-6.63E-02	8.14E-03	-7.53E-03	1.47E-03
Age>25	*Excluded*	*Excluded*	*Excluded*	*Excluded*	*Excluded*	*Excluded*
Age>30	8.21E-04	3.27E-04	*Excluded*	*Excluded*	*Excluded*	*Excluded*
Age>35	*Excluded*	*Excluded*	1.93E-02	3.29E-03	*Excluded*	*Excluded*
Age>40	*Excluded*	*Excluded*	*Excluded*	*Excluded*	9.21E-03	1.83E-03
Male	-1.89E-02	1.67E-02	-2.14E-01	1.30E-02	-3.23E-01	1.17E-02
AGE*Male	1.19E-03	5.84E-04	*Excluded*	*Excluded*	*Excluded*	*Excluded*
Age>30*MALE	-1.69E-03	6.22E-04	*Excluded*	*Excluded*	*Excluded*	*Excluded*
Black	1.91E-02	1.88E-03	2.40E-01	2.03E-02	-7.15E-02	1.99E-02
Hispanic	1.86E-02	1.89E-03	1.72E-01	2.12E-02	-3.19E-02	2.10E-02
Other	1.29E-02	1.84E-03	9.69E-02	2.30E-02	3.84E-02	2.30E-02
BMIyrago	8.18E-03	1.05E-04	-1.53E-01	1.17E-02	1.07E-01	4.03E-03
BMIyrago > 20	*Excluded*	*Excluded*	2.25E-01	1.26E-02	*Excluded*	*Excluded*
BMIyrago >= 25	*Excluded*	*Excluded*	-2.88E-02	5.41E-03	5.54E-02	4.86E-03
BMIyrago >= 30	*Excluded*	*Excluded*	*Excluded*	*Excluded*	*Excluded*	*Excluded*
BMIyrago >= 40	*Excluded*	*Excluded*	2.97E-02	1.23E-02	-8.89E-02	6.45E-03
Ln(METs)	-2.46E-06	9.92E-07	2.19E-05	8.37E-06	2.98E-05	9.62E-06
*BMIyrago	1.00E-07	4.05E-08	*Excluded*	*Excluded*	*Excluded*	*Excluded*
*(BMIyrago >= 25)	*Excluded*	*Excluded*	*Excluded*	*Excluded*	1.98E-05	4.40E-06
*AGE	*Excluded*	*Excluded*	-7.61E-07	1.78E-07	-8.48E-07	1.80E-07
F&V Servings	4.02E-03	2.39E-03	3.86E-02	1.23E-02	-3.39E-02	1.03E-02
Meat Servings	4.37E-03	2.62E-03	3.86E-02	1.23E-02	-3.05E-02	9.28E-03
Beverage Servings	4.37E-03	2.62E-03	3.67E-02	1.19E-02	-2.77E-02	8.44E-03
PREGNANT	*Excluded*	*Excluded*	7.56E-01	5.14E-02	*Excluded*	*Excluded*

As with the likelihood of BMI change, several age-based non-linearities were found. These were incorporated into the final model as indicator age-specific indicator variables, or knots, that shift the relationship between age and BMI for specific age groups. Interactions between these age-based shifts and prior BMI as well as current physical activity were also found and incorporated. No such relationships were found with race or dietary parameters.

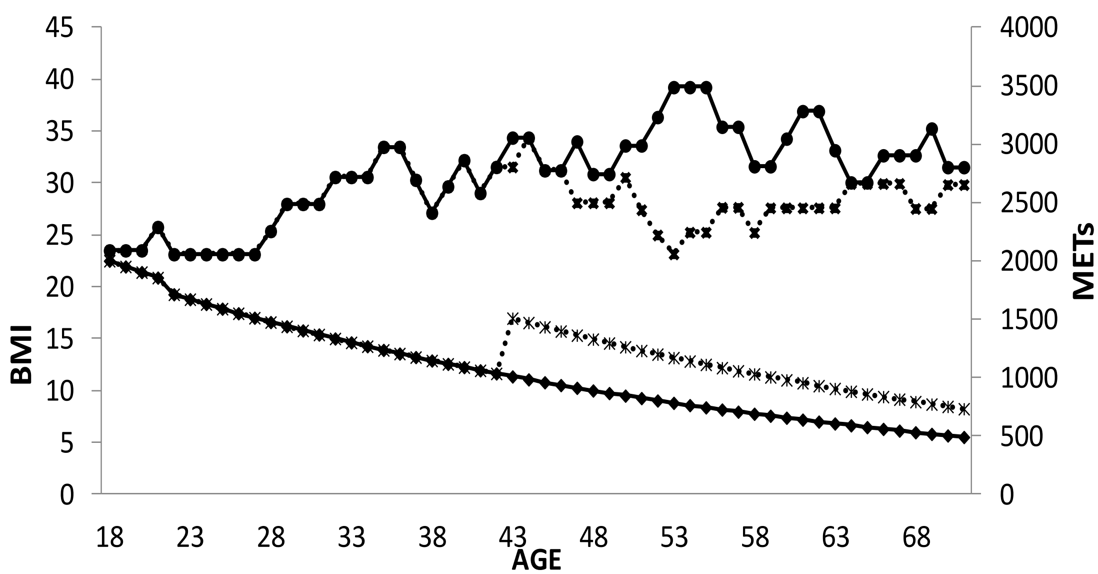
As with the youth module, physical activity – as measured in metabolic equivalents – was log-transformed to allow estimated coefficients to be interpreted as the change in BMI corresponding to a percentage change in physical activity. Again, this relationship drove the intervention effects observed within the Community Health Advisor.

Example of Lifetime Weekly Physical Activity and BMI for Simulated Agent

In ModelHealth: Obesity, BMI is the primary behavioral outcome of interest. Further, it is the primary driver of changes in cardiovascular outcomes and healthcare-related costs. The prior sections discuss how BMI is tracked throughout each simulated agent's life. This section presents a brief illustration of how this works in an actual simulated life.

Figure 6 tracks the path of adult BMI and physical activity for a white, sedentary, college-educated male as well as the impact of a one-year physical activity intervention initiated at age 42. The black lines track baseline physical activity and BMI. The dotted lines show these values following an intervention. The individual enters the model at age 18 with a BMI of 23 and an activity level of 2,000 METs/day. While his level of physical activity steadily declines, his BMI remains relatively stable until age 28, when it begins to sporadically move upward. Without the intervention, the pattern of increasing BMI continues until age 53, where it peaks at a BMI of 40 and then stabilizes. With the intervention—which dramatically increases lifetime physical activity—the individual's BMI first trends downward and then stabilizes in the overweight range.

Figure 6: Example of adult BMI and physical activity



Determining obesity-attributable costs

Individual-level MEPS data from the years 2001-2010 were used to estimate obesity-attributable costs for adults with appropriate weighting and cost deflation applied. No obesity-attributable costs for youth or adolescents are incorporated at this time (Table 7).

Rather than focusing on disease-specific incidence and ongoing cost, a function relating total medical expenditures across a wide range of BMI was developed. Further, as the goal of the empirical investigation was developing a smooth function surface, age and BMI-related knots were used to capture structural shifts in the response surface.

Two baseline specifications were estimated using a generalized linear model (GLM) specification. The first specification focused on overall costs. The second examined payer-specific differences. Interactions significant at the 5% level were retained. The following table contains estimated coefficients from the GLM models that specify obesity-attributable costs.

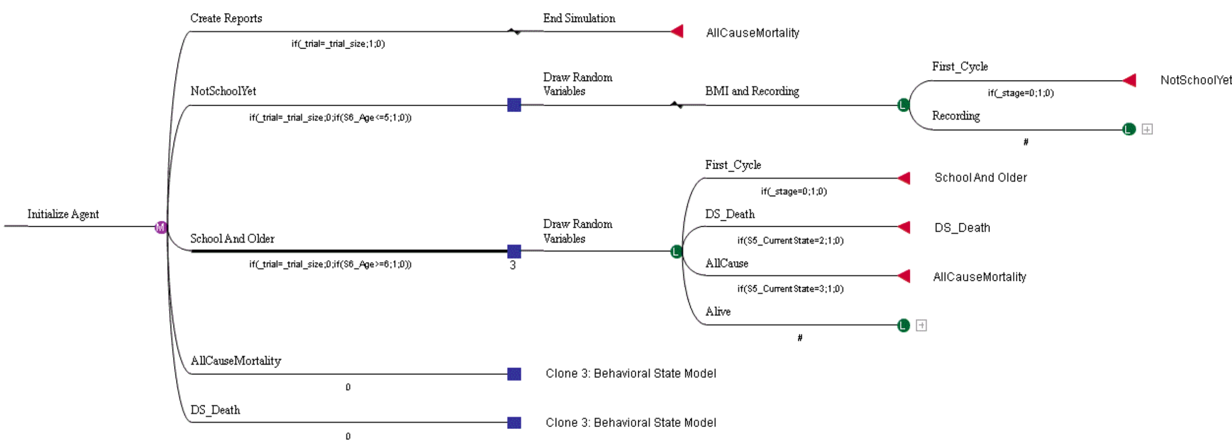
Table 7. Estimates of overweight and obesity-attributable costs

Index	Logistic	GLM
Intercept	0.186	6.029
Uninsured	-1.194	-0.885
Medicaid	0.119	0.521
Medicare	0.039	0.446
Other	0.501	0.655
Female	1.765	1.232
Black	-0.631	-0.168
African-American	-0.323	-0.154
Hispanic	-0.797	-0.383
Flu	0.904	0.498
Married	0.004	-0.052
MSA	-0.027	-0.006
Current Smoker	0.377	0.210
Age	0.023	0.039
≥ 50	0.019	-0.023
≥ 70	-0.004	-0.008
x female	-0.018	-0.018
x smoke	-0.012	-0.003
BMI > 25	-0.039	0.016
x age	0.002	0.000
BMI > 30	-0.011	0.000
x female	-0.008	-0.005
BMI > 35	0.092	0.019
x female	-0.006	-0.012
x age	-0.002	-4.000

Overview of Model Structure

Rather than create a complex decision tree comprising a diverse set of decision nodes that each represent a different combination of disease state, ModelHealth: Obesity treats its primary variables of interest (BMI and physical activity) as continuous. Thus, the layout of the decision tree reflects processes occurring in the model and not health status, or "state". Figure 7 presents the decision tree of the model.

Figure 7: The Basic Decision Tree of the ModelHealth: Obesity model



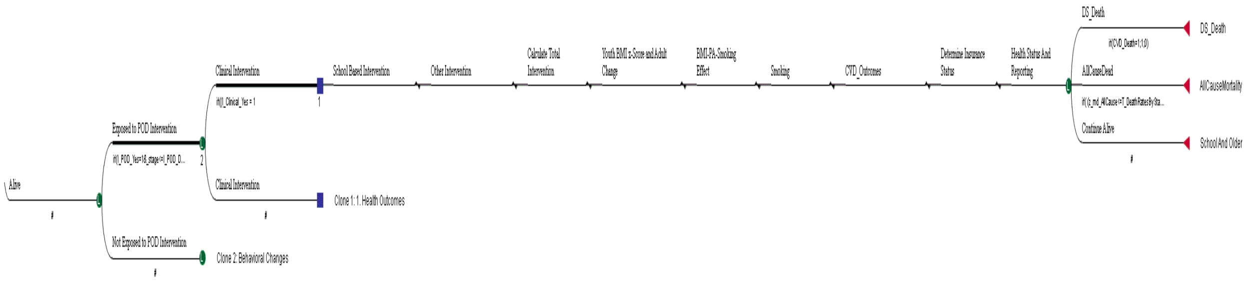
After initialization, and depending on their initial age, agents are either eligible for intervention (School and Older) or not (NotSchoolYet). Regardless of their initial age during the model's first cycle, all agents undergo an initialization cycle (First_Cycle). No interventions are applied, nor are any changes in age, health parameters or outcomes made during that cycle. Following the initialization cycle, agents are aged through life. During each cycle, each agent's behavior, health outcomes and related mModelHealth: Obesity morbidity and mortality impacts are evaluated.

ModelHealth: Obesity has two absorbing states, AllCauseMortality, and DS_Death (disease specific death). On entering these states, outcomes for the agent are no longer recorded. However, the model continues to draw the random numbers that would have been used to evaluate these outcomes, were the agent still alive.

Behaviors, interventions, disease outcomes and insurance status are evaluated in the behavioral change sub model. The structure of that module is presented below in Figure 8.

At the completion of the final trial, the model writes off three reports to tab-delimited text files. This is done at the Create Reports and End Simulation node.

Figure 8: The Behavioral change sub-module



Within ModelHealth: Obesity, if the agent is alive at the start of the cycle, the impact of the interventions and subsequent cardiovascular incidence and burden are evaluated. Each intervention is evaluated separately, and then the effects are combined. Thus, the model is flexible enough to allow for simultaneous assessment of multiple interventions; however, this is rarely done due to limitation in published effect sizes.

Once interventions are evaluated, BMI is adjusted. The method used depends on the current age of the agent. If the agent is still a youth, the youth module is engaged. If the agent is an adult, the adult module is engaged. Following evaluation of the intervention, smoking status is also evaluated. Smoking status is included in order to more precisely determine cardio-vascular risk.

References

1. Community Preventive Services Task Force. Guide to Community Preventive Services. Behavioral and social approaches to increase physical activity: enhanced school-based physical education. http://www.thecommunityguide.org/pa/behavioral-social/schoolbased-pe.html.
2. Wikipedia. Obesity in the United States. http://wikipedia.org/wiki/Obesity_in_the_United_States. Accessed 2014.
3. Center for Disease Control and Prevention. Vital Signs: Progress on Childhood Obesity http://www.cdc.gov/obesity/index.html. Accessed 2014.
4. National Collaborative on Childhoold Obesity Rearch. Statistics on Adolescent Obesity in the U.S. 2014; http://www.nccor.org.
5. Barton, M. Screening for obesity in children and adolescents: US Preventive Services Task Force recommendation statement. *Pediatrics.* Feb 2010;125(2):361-367.
6. Summaries for patients. Screening for and management of obesity in adults: U.S. Preventive Services Task Force Recommendation. *Ann Intern Med.* Sep 4 2012;157(5):I-32.
7. United States Preventive Services Task Force. Final Recommendation Statement Obesity in Adults: Screening and Management. *Final Recommendation Statement* 2012; http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/obesity-in-adults-screening-and-management#Pod2.
8. LeFevre, M.L. Behavioral counseling to promote a healthful diet and physical activity for cardiovascular disease prevention in adults with cardiovascular risk factors: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* Oct 21 2014;161(8):587-593.
9. Recommendations to increase physical activity in communities. *Am J Prev Med.* 2002 May 2002;22(4 Suppl):67-72.
10. Community Preventive Services Task Force. Environmental and Policy Approaches to Increase Physical Activity: Point-of-Decision Prompts to Encourage Use of Stairs. 2002; http://www.thecommunityguide.org/pa/environmental-policy/podp.html Accessed 2014.
11. US Energy Information Administration. *Commercial Buildings Energy Consumption Survey* 2014 http://www.eia.doe.gov/emeu/cbecs/public_use.html, 2014.
12. National Transit Database. http://www.ntdprogram.gov/ntdprogram/data.htm.
13. United States Census Bureau. American Community Survey 2014; http://www.census.gov/acs/www/.
14. National Center for Education Statistics: Data Tools in Sciences. 2014. http://nces.ed.gov/datatools.
15. Bureau of Labor Statistics. Occupational Employment Statistics. *May 2013 National Occupational Employment and Wage Estimates United States* http://www.bls.gov/oes/current/oes_nat.htm.
16. Survey of Income and Program Participation. 2008. http://www.census.gov/programs-surveys/sipp/. Accessed 2014.
17. Behavioral Risk Factor Surveillance System Survey Data (2009). Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2010. http://www.cdc.gov/brfss/annual_data/annual_2009.htm.
18. Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey Data (2010-2011). 2012; http://www.cdc.gov/obesity/data/childhood.html.
19. Center for Disease Control and Prevention. National Health Interview Survey. 2010; http://www.cdc.gov/nchs/nhis/tobacco/tobacco_questionnaires.htm.
20. National Health and Nutrition Examination Survey. 2000. http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm.
21. Cunningham, S.A., Kramer, M.R., Narayan, K.M. Incidence of childhood obesity in the United States. *N Engl J Med.* Jan 30 2014;370(5):403-411. https://www.ncbi.nlm.nih.gov/pubmed/24476431