FORECASTING ADULT WEIGHT STATUS TRENDS IN CANADA WITH LONGITUDINAL POPULATION-BASED DATA

by

Mark J. Staples

Submitted in partial fulfilment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
March 2015

Table of Contents

List of Tables

List of Figures

Abstract

The prevalence of obesity in Canada and other Western countries has increased dramatically over the last 30 years. This project modelled future prevalence of weight status trends under a Markov framework by estimating transition probabilities between categorical weight states. A microsimulation was developed to track individual weight status histories through discrete-time updates. Model and simulation parameters were derived from a nationally representative longitudinal survey which followed 17,000 participants over a 20 year period. A series of weight prevention and intervention strategies were simulated to determine their potential impact on weight status trends and population health utility.

List of Abbreviations Used

APPLE	Alberta Project Promoting active Living and healthy Eating
APPSIM	Australian Population and Policy Simulation Model
BMI	Body Mass Index
CFI	Canada Foundation for Innovation
CIHR	Canadian Institutes of Health Research
CRDCN	Canadian Research Data Centre Network
CVD	Cardiovascular Disease
ICER	Incremental Cost-Effectiveness Ratio
LOWESS	Locally Weighted Regression Smoothing
NPHS	National Population Health Survey
NSERC	Natural Sciences and Engineering Research Council
POHEM	Population Health Model
QALYs	Quality Adjusted Life Years
RDC	Research Data Centre
SSHRC	Social Sciences and Humanities Research Council
WHO	World Health Organization

Acknowledgements

I am indebted to my supervisors Dr. Peter Vanberkel and Dr. Stefan Kuhle, who have generously given their time, effort and knowledge throughout my degree. Their unending encouragement, advice and patience enhanced my learning experience greatly. The countless hours of project discussion and thesis editing were invaluable to my personal and academic growth.

I gratefully acknowledge the financial support I received through the Nova Scotia Health Research Foundation's Scotia Support Award, awarded to the Perinatal Epidemiology Research Unit of the IWK Health Centre in Halifax, NS. I am also grateful for the financial support received through NSERC funding awarded to Dr. Peter Vanberkel of Dalhousie's Department of Industrial Engineering.

I would like to acknowledge Heather Hobson of the Atlantic Research Data Centre, who welcomed all my questions and was always willing to share her expertise.

I also extend my thanks to Dr. Corinne Macdonald, a professor of mine during a pivotal year of my undergraduate degree. Without her teaching style I would not have gained the confidence and abilities to pursue this endeavor.

I am also very thankful for the support I received from my friends and family. Thanks to Stéphanie for her enduring love and patience during my academic studies. Finally, thank you to my parents for their unconditional love, confidence in my abilities, and encouraging me to pursue my ambitions.

The analysis presented in this thesis was conducted at the Atlantic Research Data Centre, which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the Atlantic Research Data Centre are made possible by the financial or in-kind support of the SSHRC, the CIHR, the CFI, Statistics Canada, and Dalhousie University. The views expressed in this paper do not necessarily represent the CRDCN's or that of its partners'.

Chapter 1: Introduction

Over the last decade overweight and obesity levels in Canada have grown to unprecedented levels, sparking debate on how government should intervene to help mitigate any further increase of its prevalence. In Canada from 1970-2004, data from a representative survey revealed an increase in adult obesity prevalence from 10 to 23% of the population, or a relative increase of approximately 200% [1]. In 2009-2011 obesity reached 27% while overweight rose to 40% according to the Canadian Health Measures Survey [2]. The obesity trend is similarly concerning for children and young adults, who are increasingly diagnosed with obesity-related conditions formerly exclusive to adults [3]. These findings are particularly concerning for Nova Scotians, as our rate of overweight and obesity ranks among the highest in Canada [4].

Climbing obesity rates have translated into higher prevalence of chronic health conditions such as type II diabetes, high blood pressure, heart disease and some forms of cancer. More specifically, a 2004 article found that 39% of type II diabetes cases, 45% percent of hypertension and 23% of coronary artery disease could be attributed to obesity [1]. A study estimating cardiovascular disease risk (CVD) concluded obesity and diabetes were the only risk factors of CVD which will not decrease in the next two decades [5]. A wealth of research indicates obesity is closely related to many undesirable health conditions and has appropriately become a top priority of the Canadian health agenda. A decade ago, obesity was cited as "rivaling smoking as the leading cause of premature death and illness" [6], and others suggest its health burden has indeed surpassed smoking in some respects [7][8].

Numerous analyses have estimated the overall cost of overweight and obesity to society. Janssen *et al.* [9] estimated the economic burden of Canadian obesity in 2001 to be $4.3 billion (direct costs $1.6B, indirect $2.7B) annually. In 2006 as obesity rates continued to climb, additional obesity-related comorbidities were identified and cost estimates increased to as high as 11 billion (direct costs $6B, indirect $5B) annually [10]. In a US study, Finkelstein *et al.* [11] reported that compared to normal weight individuals, those who are obese incur 27% more physician visits and outpatient costs, 46% more inpatient costs, and spend 80% more on prescription medication. Similarly an Ontario-based study concluded overall physician costs for obese adults were 18.2% higher compared to their normal weight counterparts [12]. The relationship was also determined to be age dependent, as obese adults over 60 were found to incur 28% higher physician costs than their normal weight counterparts.

The increase in obesity-related illnesses and healthcare costs provide Canadian policy makers with an incentive to reduce overweight and obesity prevalence. Poor nutrition and lack of physical activity are already established as major modifiable risk factors for overweight and obesity. Therefore, healthy diets and active lifestyles are the key foci of primary prevention and intervention strategies. Despite this apparent link, large-scale governmental weight management policies or programs have yet to be implemented due to uncertainty of their long term value and critiques of their cost-effectiveness [13]. These critiques, however, may begin to diminish as the weight of evidence in favour of prevention and intervention initiatives are substantiated.

Although adult intervention programs have come under scrutiny due to subsequent weight regain, Prince [14] argues a multilayered approach is needed rather than isolated weight loss, exercise or healthy eating interventions. Children's school prevention programs have shown more promising results [15], using the school environment to emulate a societal microcosm in which a broad range of policies, interventions and prevention programs can be tested. One example demonstrating successful obesity prevention is the Alberta Project Promoting active Living and healthy Eating (APPLE) within 10 selected schools [16].

The value of prevention programs lies not only in knowing their effectiveness, but also in weighing the cost of such programs against the future obesity outlook without changing the status quo. This thesis develops a simulation framework to forecast the trajectory of weight status prevalence in the coming decades. The proposed research is the first study to use Statistics Canada's longitudinal data sources to specifically model and simulate weight status trends in the Canadian population. Results from this simulation will also be extended to support the effective implementation of obesity management and prevention strategies.

Chapter 2: Problem Statement

Our health care system is structured to meet the immediate demands of those most ill, making it difficult to argue for investment in significant long-term obesity prevention or intervention programs. Furthermore, there is data available for Canada as to how the obesity epidemic might progress into the future. This gap leaves decision makers with a limited capacity to weigh potential future savings of obesity management programs against the future healthcare costs of the obesity epidemic.

Throughout the last several years researchers have produced a wide range of estimates trying to forecast the prevalence of obesity into the future. Wang *et al.* [17] extended current trends to predict 80% of American adults will be overweight or obese by 2022 and eventually all adults will reach either state by 2048. They used a simple linear regression model, based on prevalence data itself to forecast future trends. As Levy *et al.* [18] notes, predictions based on past prevalence trends are usually validated over a limited number of years, so extrapolating too far into the future will likely result in an inaccurate prediction. In contrast, others incorporated recent evidence that obesity rates were peaking and forecasted a stabilization of obesity prevalence in the coming years [19][20]. Other models range from a system dynamics caloric model to analyses of how social networks impact obesity, illuminating the complexity of the problem and heterogeneity of the modelling approaches.

To improve predictive performance microsimulation is often proposed. Guy Orcutt [21] is frequently credited for pioneering the field of tax policy microsimulation, a methodology that has since expanded to the healthcare field [22]. Microsimulation uses data gathered at the individual level to model and simulate individual's life histories within an environment representing demographic and social characteristics [23]. In essence, microsimulation theory is not fundamentally different from discrete-event simulation, a well-developed field in the operations research community. A simulated environment allows decision makers to apply hypothetical policy changes and analyze their impact on an individual or aggregated level. Simulating individual histories is generally preceded by a mathematical model which describes how individuals are expected to change based on their attributes. For instance, when modelling future disease rates researchers often assume individuals can be described through a Markov process, and then simulate them using a Monte Carlo approach [24]. These models use random sampling for a set of likelihoods ("transition probabilities") to determine an individual's movement through stages of a disease. Estimates for the transition probabilities are commonly obtained from longitudinal or cross sectional studies of the target population.

An Australian research project developed a microsimulation called the Australian Population and Policy Simulation Model (APPSIM), which incorporated several societal processes such as healthcare, social security, labour, and demographic change among others [25]. Lymer and Brown [26] further developed the microsimulation's health component to predict obesity prevalence and the resulting future health expenditures in Australia until 2050. Similarly, UK researchers developed a microsimulation framework to study obesity and other chronic diseases [27], which has since been extended to other European countries [28]. Although we can learn from foreign obesity prediction studies and microsimulations, it is rarely practical or feasible to simply transfer these models for use in Canada due to unique modelling assumptions and data-specific programming requirements of estimating the transition probabilities.

No Canadian study forecasts and explores the social or economic effects of the obesity epidemic through modelling individual's weight status changes. In 1994 Statistics Canada released the Population Health Model (POHEM) microsimulation, which included a body mass index (BMI) prediction module [29]. However, similar to Wang *et al.* [17], the module is based on linear projections of prevalence of BMI through time. Apart from a brief model description, no validation documents or report is available to further examine their obesity module. POHEM has since been used to simulate osteoarthritis [30] and physical activity [31] with respect to the BMI predictions. Therefore, an opportunity exists to supplement the current literature by offering an alternative modelling approach to forecast Canadian obesity.

As obesity is considered a preventable disease there appears to be substantial incentive to advance our understanding of these trends and if obesity could be prevented. The overarching goal of the proposed research is to forecast overweight and obesity prevalence in the Canadian population over the next several decades and assess the impact of existing and hypothetical weight management programs. The first phase of the project employs mathematical modelling to determine the probabilities that the weight status of an individual will change in the future given their current weight status, age and gender. Subsequently, these transition probabilities are input into a Monte Carlo simulation model to forecast the future development of overweight and obesity in Canada. Finally, model inputs are altered to simulate potential prevention and intervention efforts.

Chapter 3: Literature Review

In this section the main data sources are outlined for population-based obesity research and surveillance in Canada, followed by a more detailed discussion of the National Population Health Survey (NPHS), the chosen data source for this project. Subsequently the BMI indicator is discussed, followed by an overview of obesity modelling and microsimulation is given. Finally, potential prevention and intervention strategies are briefly reviewed while exploring their potential application to the proposed modelling framework.

There are several large Canadian surveys which have tracked and provided researchers with a means to model obesity trends. These surveys are gathered by either cross-section or longitudinally. Cross-sectional surveys generate a new sample of participants each collection cycle, whereas longitudinal surveys collect information on the same participants every cycle. For example, a cross-sectional survey called the Canadian Community Health Survey has gathered information related to health status, health service utilization and health determinants since 2009. The National Longitudinal Survey of Children and Youth monitored the development of individuals aged 0-25 over eight survey cycles from 1994 to 2008. A third major Canadian survey and the focus of this research, is the NPHS. It collected information longitudinally on 17,267 participants from 1994-2010.

Throughout the NPHS survey there are two factors which prevent it from having the desired characteristics of being truly random and a representative sample of the Canadian population. During the personal interview stage of the first cycle it was too costly to provide coverage for every geographic region. This results in an unequal chance to be chosen for the survey and therefore an underrepresentation of some demographic groups. A post-stratification weighting procedure was applied by Statistics Canada after each survey cycle to account for this unequal probability of selection. Secondly, the total number of survey participants decrease each cycle from attrition due to death or non-response. By 2010/2011, only 46.2% of the original sample remained. For a more detailed discussion of complex survey design issues, the reader is directed to the NPHS user guide [32].

3.1 Body Mass Index

The main measure of obesity is an individual's BMI and is calculated as weight divided by height squared. BMI was introduced in the early 20th century and has since become widely used in assessing weight-related mortality, morbidity and chronic illness. Depending on the modelling framework and available data, one may choose BMI as a continuous variable or separate it into discrete categories. The World Health Organization (WHO) [33] classifies individuals into one of six weight statuses (Table 1) with obesity broken into severity levels due to poorer health outcomes associated with increasing BMI levels.

Table 1 - WHO weight status classification

Weight Status	BMI (kg/m^2)
Underweight	<18.5
Normal weight	18.5 > 24.99
Overweight	25-29.99
Obese Class I	30-34.99
Obese Class II	35-39.99
Obese Class III	>40

Despite the widespread history of using BMI as a measure of health, it is an imperfect indicator. For example someone with an above average muscle mass will have an elevated BMI and possibly be misclassified in a higher risk category for developing health conditions. Katzmarzyk and Janssen [34] argue it would be more informative to measure BMI combined with waist circumference to more accurately measure an individual's health risks. However, measuring individuals' waist circumferences is an uncommon practice for population wide surveys and the use of BMI is considered sufficient to estimate population health in large-scale epidemiological studies [35].

To ensure physical health data is unbiased and of high quality, participants are ideally measured by a physician. However personnel and logistical efforts required to do this are very costly, particularly over several years of a longitudinal survey. Instead, large surveys often collect data by relying on individuals to self-report information. Survey participants tend to overestimate their height and underestimate their weight, which results in a self-reported BMI that is lower than their true BMI. In a report by the Canadian Institute for Health Information and the Public Health Agency of Canada [36] found an approximately 8% difference in measured (25.4%) and self-reported (17.4%) obesity prevalence as of 2009. Shields *et al.* [37] among others [38][39] have found using self-reported data for body mass index can lead to biased conclusions. Shields accounted for this bias in the NPHS by fitting a series of regression equations to adjust for self-reported BMI.

3.2 Simulation in Healthcare

Simulation is a powerful analytical tool in healthcare applications that supports the delivery of more efficient services while maintaining treatment quality [40]. Discrete-event simulation has been extensively applied to aid in operational healthcare decisions, such as waitlist management, surgery workload scheduling and patient flow research. For these operational simulations, the goal is to find a solution that meets a minimum level of service or quality standard with constrained resources. Simulations aiding in health care decisions for population health are also discrete-event simulations, but are more commonly called microsimulations. In microsimulations individual's life histories are simulated within an environment representing demographic and social characteristics of the population of interest. These models permit an examination of hypothetical scenarios in lieu of policy experiments which would be otherwise too costly or impractical to test [41]. Simulated individuals typically transition through a series of states according to observed or assumed economic and social behaviours. In the context of obesity, deriving this behaviour is an ongoing field of research, as will be discussed in the following section.

3.3 Obesity Modelling

This section is an overview of several past obesity forecasting studies and their research methods. The methodologies within the obesity modelling context are categorized in this thesis as: (1) Linear BMI modelling, (2) Transitional BMI modelling (3) Other approaches.

3.3.1 Linear BMI Modelling

Several studies have used regression to project the prevalence of continuous or categorical BMI in order to project trends into the future. Perhaps the most provocative obesity forecast is provided by Wang *et al.* [17], who projected all Americans will be either overweight or obese by 2050. Despite this unlikely outcome Wang's model operates under the same simplifying assumption many health prediction models make when using real data - that current trends will hold linearly, into the future. In this case the trend clearly cannot hold past 2050, but it nonetheless offers a concerning projection of American obesity.

In 1994 Statistics Canada released POHEM, a discrete-event, continuous-time microsimulation framework designed to simulate how individuals may react to hypothetical policies. Researchers at McGill University's Surveillance Lab created an obesity module for the POHEM framework that models change in categorical BMI with a series of linear regressions [29]. Their data consisted of five cycles of the NPHS longitudinal survey from 1994-2004 and the variables gender, age group, previous weight status, income, education and region of residence. The data was divided into 28 different groups conditioned on age group, gender, previous weight status, followed by 28 separate linear regressions. Although a linear model is used, an advantage is that their model need not assume the behaviour of future covariate distributions since other modules of POHEM predict them.

While linear regression modelling of BMI prevalence can provide policy makers with a "what-if" scenario of the future economic and health impacts of obesity, forecasts of a linear nature can be problematic if extrapolating over a longer horizon [18]. Von Ruesten *et al.* [42] noted obesity prediction based on prevalence trends is quite sensitive to the *a priori* modelling assumption of either a linear or a non-linear functional form.

3.3.2 Transitional BMI Modelling

As previously mentioned, BMI can be discretized into weight statuses. A transitional BMI model could estimate the likelihood, or transition probabilities, of individuals moving between these different weight statuses. Upon estimating these transition probabilities, one can simulate a population and aggregate the individual life histories to obtain the weight status distribution across time. Several methods of calculating transition probabilities are discussed in this section.

One means of calculating transition probabilities is through a Markov framework. A first order Markov model predicts an individual's next state conditional only on their current state, and is therefore considered "memoryless" [43]. The transition probabilities are typically estimated with one large set of transition observations; however longitudinal data is comprised of numerous individual trajectories over a relatively small number of observations. Kalbfleisch and Lawless [44] showed that longitudinal data can be used so long as the individuals are mutually independent. The authors developed a continuous-time Markov model, which estimates an individual's transitions through states as function of their current state and time spent in that state. A continuous Markov framework is advantageous in some cases because it can estimate the duration between events as well as accommodate multiple event occurrences when the order of occurrence is important. Alternatively, a discrete-time Markov model requires equally spaced, non-missing observations to properly estimate transition probabilities. Galler [45] showed that a discrete-time Markov model can be considered a special case of a continuous-time model if the longitudinal data structure has equally spaced observations and sufficiently short cycle durations. When these two conditions are met, Galler recommends using a discrete-time Markov model to reduce theoretical and modelling complexities when building a microsimulation.

One caveat to the previously discussed Markov models is that transition probabilities are stationary across time. Singer [46] developed a non-stationary Markov model for when this assumption does not hold. Another caveat is that subgroups of the population behave differently and likely do not have homogenous transition probabilities across socio and demographic strata. A US study by Basu [19] applied a Markov approach to predict individual weight status while incorporating several other variables such as age, gender and ethnicity. Without sufficiently disaggregating the different behaviours of subgroups there is a risk of introducing overestimation or underestimation bias [47][48]. However, Andreassen [49] notes that adding explanatory variables to a Markov model quickly expands the number of transition probabilities one must estimate and thus the amount of possible disaggregation is limited by the available longitudinal survey data.

Modelling individual categorized BMI is also possible through logistic regression where, for instance, normal and obese categories represent the binary states to be predicted. With longitudinal data the likelihood of moving between BMI weight statuses can be derived by pooling individual histories into a single dataset. When modelling longitudinal data it is common to use a transitional logistic model structure [50], which includes a time lagged dependent variable, such as weight status. Using the estimated model coefficients individual transition probabilities can be derived for use in a microsimulation model. For instance, individuals in APPSIM [25] transition between normal and obese weight statuses derived from transitional logistic regression.

Agresti [51] extended transitional logistic regression to transitional multinomial regression, for which three or more categories can be simultaneously modelled. Modelling health states with multinomial regression has been used in several health care contexts [52][53][54]. Xie and Zimmerman [55] have specifically addressed transitional multinomial regression for categorical data of longitudinal surveys. If longitudinal data is unavailable, it is possible to use cross-sectional data to derive transition probabilities. One such study uses a cross sectional survey and linear programming with a superadditive cost matrix to ensure individuals can only transition between neighbouring weight statuses [56]. As of 2010, Statistics Canada no longer maintains a longitudinal population health survey, so transition modelling with cross-sectional data may gain popularity as a viable alternative.

3.3.3 Other Approaches

Several methodologies other than linear and transitional BMI modelling have been employed to predict future weight status prevalence. This section describes several of these approaches.

A British research project, called the Foresight Report [27] found weight status prevalence was approximately linear from 1993-2004, however acknowledged that applying a linear model would erroneously lead to some groups exceeding 100% obesity prevalence. Instead, they used a monotonic non-linear function that is asymptotic to 0 and 100% to predict the prevalence of each BMI category through time. Results were then normalized to ensure prevalence at each time summed to 100%. The authors then converted these cross-sectional predictions to a longitudinal individual behaviour model via an inverse cumulative function from the regression results.

The Organization of Economic Co-operation and Development created a quantile regression model to project the continuous BMI distribution for several countries [57]. The model used covariates that the authors found to be highly correlated with BMI such as gender, age, marital status, ethnicity, and education attainment. The future BMI distribution was estimated by using population projections by age and gender, while other covariates were held at current levels. The authors cautioned that these types of projection models are dependent on the covariate data distributions remaining stationary through time. For Canada, they projected a further obesity rate increase of approximately 5% across all age groups from 2010 to 2020.

As Alimadad *et al.* [58] notes, obesity is a chronic disease that is often the result of long-term poor health behaviours. They argue a conventional memoryless first order Markov model may not be optimal for modelling obesity. Instead they estimate an individual's probability of transitioning between each weight status based on their maximum historical weight status for up to eight previous periods. For example, a person with a maximum weight of normal through only two survey periods has a 0.834, 0.159, and 0.007 chance of being normal, overweight or obese, respectively in the next period. Whereas, if an individual has maintained a maximum weight of normal through all eight cycles the conditional probabilities are 0.893, 0.104, and 0.003. The authors found as one maintains a normal maximum weight history through time their chance of remaining normal weight increases and their chance of transitioning to overweight or obesity decreases. There are however two challenges in building a microsimulation based on their method. Firstly, all individual's maximum weight histories are not known prior to the start of the survey. Lacking a technique to predict these histories the authors set each initial maximum weight history to the first survey response of each individual. The second challenge is how to extend the model beyond the eight available survey cycles. Since transition probabilities for a maximum weight status history of greater than eight cycles are not defined, it would be necessary to extrapolate or predict how the probabilities change as more time is spent in each weight status.

Thomas *et al.* [20] offered another approach of incorporating individual history that combines system dynamics and a disease modelling framework called SIR ("Susceptible, Infected, Recovered"). A series of differential equations determine transition rates from a susceptible state (normal weight) to an exposed state in which individuals can become "infected" with overweight and obesity. A recovered state represents individuals who were previously overweight or obese, but have since returned to normal weight. These recovered individuals become more susceptible to re-entering overweight than their normal weight counterparts who have never been overweight. Another interesting feature of Thomas' research is the incoming population has a certain chance of being born into an 'obesogenic' environment in which they enter directly into the susceptible state. This is considered a feedback effect since the probability of birth into an obesogenic environment is estimated by the percentage of reproductive age women categorized as overweight or obese.

Another approach is agent-based modelling. These methods models each individual in terms of 'rational actors' who make decisions based on a complex set of rules depending on their characteristics and external environment. For example, an obesity agent-based model, by Bourisly [59] assumes rational actors will make dietary and physical activity decisions based on food prices, exercise facility location and their current weight status. The disadvantage of this type of modelling is that having so many interactions between individuals and the environment can force models into incorporating unverifiable or simplified decision rules and parameterizing the model while lacking empirical data [60].

Other notable obesity modelling approaches include mapping the spread of obesity through social networks [61] and simulating an individual's caloric imbalance [62]. Research is currently underway to improve the existing POHEM-BMI module [63], however no upcoming publications are declared and a release date for the model is unknown. Another project under development is the Canadian ACE-BMI model [64]. It is a Markov cohort model and borrows from the Australian-based Assessing Cost Effectiveness (ACE) obesity framework [65].

3.4 Microsimulation Characteristics

Upon reviewing healthcare microsimulation literature, Okhmatovskaia et al. [66] observed that there is a need to be more consistent when defining a model's technical terms. This section will clarify the terms used within this microsimulation.

The model and modelling parameters that describe each individual's behaviour are usually defined as either *deterministic* or *probabilistic*. Deterministic parameters are simply single point estimates of model parameters. Probabilistic parameters add a distribution around the point estimate and are used when there is a significant amount of variance in the estimates. For continuous and categorical regression techniques confidence intervals are readily available. For a Markov approach resampling techniques (bootstrapping) can be used, as in Basu [19], who estimated probabilistic weight status transition probabilities.

Another important consideration is whether individual's behaviour within a microsimulation is assumed to be *static* or *dynamic*. A *static* microsimulation is distinguished by unchanging behaviour or transition rules through time [19][67]. *Dynamic* microsimulations on the other hand, describe an individual's behaviour as a function of time or a function of the change in explanatory variables over time [11][27]. Although individual behaviour might be assumed static, this does not preclude other simulation elements from being dependent on time. For example in a Norwegian microsimulation [49] transition probabilities are stationary through time, yet mortality rates are adjusted as time progresses. Similarly, the obesity component of POHEM [29] is static through time, yet the explanatory variables that determine BMI are dynamically updated every simulation cycle. Li and O'Donoghue [68] note that as data availability and computational power increase, dynamic models have become a more realistic option for long-term population prediction models. However, they also caution that dynamic model development is usually considerably more time consuming than static models.

Microsimulation models are further distinguished between *cohort* or *population* models, depending on what output information is desired. A *cohort* microsimulation begins with an initial cohort born in the same year, or an initial sample of all ages. A group of individuals can be simulated through multiple prevention or intervention scenarios, and then compared to the reference scenario simulation results. Cohort simulations are terminated when either all participants have died or reach a specified age limit. For example, Lightwood *et al.* [69] used a cohort aged 35-64 to project the economic costs of obesity until 2050 with a Markov model simulation. Cohort models, however, do not account for the changing demographics over time due to the incoming and outgoing population. A *population* microsimulation can investigate demographic effects such as age and gender by adding births, deaths and immigration into the model. Although the Canadian age and gender demographic predictions are likely accurate for the decades ahead, the future distributions of other variables are estimated with less confidence or not at all. As a result, these future distributions of interest either rely on external forecasts or are assumed to remain static throughout the simulation.

The final simulation characteristic discussed here is whether an initial population is generated *synthetically* or with *cross-sectional* data. Sometimes an individual's entire life history is a necessary consideration in predicting a future status or event. However if a sample of individuals is captured during a relatively short survey timeframe there may be insufficient data to explore the historical significance of any variables. In this case a *synthetic* population can be generated. Calibration techniques can fill this information gap by retrospectively generating synthetic life histories for each individual using one or several external data sources [70]. If prior history is not necessary for a model, a *cross-sectional* approach can instead use survey data to define the initial population, as in Lymer and Brown's APPSIM research [26] and the Foresight Report [27].

According to these definitions of microsimulation characteristics, this project develops a simulation that has *deterministic* transition probabilities, is behaviourally *static* and begins with a *cross-sectional* initial population. The simulation develops both a *population* and *cohort* analysis component. For a further review of microsimulation characteristics and nomenclature, see Rutter *et al.* [70].

3.5 Intervention and Prevention

This section briefly describes various prevention and intervention research in Canada, and how this research can be adapted for this model. According to Sacks and colleagues [71], there are three main streams of obesity prevention and intervention:

1. Public policies which target a range of social and environmental factors related to obesity (i.e. subsidies on healthier foods, infrastructure promoting active lifestyles)
2. Community-level prevention and intervention programs that influence individual behaviour (i.e. education, school-based programs, workplace programs)
3. Individual health services and clinical interventions for those with high BMI and other comorbidities (i.e. weight loss programs, bariatric surgery)

The first stream is usually analyzed with agent based modelling or system dynamic techniques, which describe the relationships between an individual's environment, social status, and individual economic preferences. This research does not consider social or environmental policies decisions; however, I introduce a hypothetical nation-wide prevention policy which is assumed to alter the transition probabilities of moving between certain weight statuses.

The second stream consists of community-level prevention and intervention projects. For example, a national initiative called the PAN Canadian Health Living Strategies was created in 2005 [72]. It consists of a declaration by government on federal, provincial and territorial levels to prioritize the promotion of chronic disease prevention, such as obesity. However Prince [14] argues the initiative places much of the onus on adults to eat right and exercise by providing only basic means of education such as the Canadian food guide and the nutrition labelling program. Another community level obesity program was launched in British Columbia, called ActNow [73]. It is an initiative to support obesity prevention across all ages by targeting multiple risk factors. Some of the projects include 'Dial-A-Dietitian' and 'Health Check' to encourage healthy diet choices, as well as the implementation of healthy lifestyle programs into work environments. A third community level program, APPLE Schools, tracked and analyzed a prevention program within 10 selected schools [74]. Since the program's inception in 2008, follow-up research has confirmed its effectiveness [16] and simulated obesity-related costs savings of scaled-up versions of the program [75]. Results of APPLE Schools are incorporated into this research by assuming that the successful childhood obesity reductions will be maintained as they reach adulthood.

At the individual level research encompasses adult dietary interventions [76], physical activity interventions [77] and a combination of both [78]. In more extreme cases, clinical options are available such as bariatric surgery [79], or prescription medication [80]. An intervention scenario will be introduced into this research where 10,000 obese individuals per year receive a bariatric procedure.

Chapter 4: Project Methodology

This thesis is partitioned into three primary objectives: 1) manipulating the NPHS survey into a useable dataset, 2) modelling how individuals transition between each weight status, and 3) simulating the Canadian population to forecast and assess potential prevention programs. All of the analysis was performed in R, a programming language for statistical data analysis [81]. The first two objectives were completed within the Atlantic Research Data Centre (RDC) in Halifax N.S., where the NPHS is accessible to researchers. The third objective was completed outside of the RDC. The distinction between inside and outside the data centre is important because of strict rules governing which data can leave the data centre. Generally speaking, any data point which is derived from less than 5 participants cannot be removed from the RDC.

4.1 Data Preparation

The NPHS was chosen for this research to provide data and for internal validation purposes. The survey was conducted every two years between 1994 and 2010, and followed the participants for 18 years (or 9 cycles) making it a longitudinal survey. In total approximately 150 questions were asked each year of 17,276 Canadians. Information on the respondents was gathered through personal interviews initially and subsequent interviews were conducted on the telephone. Northern Canada (Yukon, Nunavut, North West Territories) and people living in health institution (e.g. nursing homes) were gathered in a separate survey and are omitted from this research.

4.1.1 Datasets

The NPHS is voluntary, so not all participants took part in every survey cycle, resulting in a significant number of missing weight or height responses. This thesis defines three different datasets: 1) "no exclusions", 2) "limited" and 3) "intermediate", each with varying amounts of missingness.

The first dataset (“no exclusions”) contains all individuals regardless of how much of the survey they have completed in each cycle. The second dataset (“limited”) contains those individuals who either participated in all survey cycles, or have died yet participated in all previous cycles before death. Note that datasets “no exclusions” and “limited” are respectively equivalent to the “Square” and “Full” datasets described in NPHS user guide [32]. The labels have been changed herein to improve clarity.

My research methodology does not require completely answered surveys (i.e. the “limited” dataset), but only that individuals have responded to any two *consecutive* survey cycles. The third dataset (“intermediate”) is then a subset of the “no exclusions” dataset, created in two steps. First, for each survey year in the “no exclusions” dataset, any instances where individuals are missing height or weight are removed since BMI cannot be calculated. Secondly, each individual’s profile is checked for consecutively complete cycles. All instances of consecutively complete survey data compose the “intermediate” dataset. This approach of subsetting consecutive data is described in the NPHS user guide as the “Cycle Twinning Approach” [32, pp. 63]. It is important to note that this approach assumes that missingness of the data occurs uniformly across all individuals and is therefore is ‘ignorable’ [82].

The “no exclusions” dataset consists of 140,290 records while the “limited” consists of 75,739 records. The (“intermediate”) dataset consists of 126,671 by salvaging 50,932 records from the “limited” dataset.

4.1.2 Survey Coding

The survey questions required to create the “no exclusions”, “limited” and “intermediate” datasets are listed below, along with the corresponding NPHS syntax.

Table 2 - Survey questions and corresponding NPHS codes

Question	Variable description	NPHS Survey Code (Primary Key)
What is your gender? (first cycle only)	Gender	SEX
What is your date of birth (first cycle only) and what is your age (subsequent cycles)	Age	DHC**n**_AGE
How much do you weigh? Was that in pounds or kilograms?	Weight	HWC**n**I3LB
How tall are you without shoes on?	Height	HWC**n**_IHT
It is important to know when analyzing health whether or not a person is pregnant. Are you pregnant?	Pregnancy Status	PHC**n**_4B
	Survey Status	SP3**n**_STA
	"no exclusions" sampling weights	WT64LS
	"limited survey" sampling weights	WT6DLF

The bold "n" found in some of the NPHS Survey Codes indicates the cycle. The survey cycles are indexed by n = {2,4,6,8,0,A,B,C,D} representing years 1994 to 2010 where n=2 is 1994, n=4 is 1996, etc. The answers to the height and weight questions are not recorded directly but rather using representative codes. For example, a height of 47 inches is recorded as "37". These conversions are available from Statistics Canada [83]. Survey status indicates whether an individual has missed a cycle or has deceased. Sampling weights for the "no exclusions" and "limited" datasets are discussed in the following section.

4.1.3 Survey Sampling Weights

When selecting participants for the NPHS survey, Statistics Canada does not select people proportionally equal from all geographic regions. This is done for practical reasons and to save costs. To correct for this bias, Statistics Canada calculated sampling weights to increase or decrease the influence of each respondent in the analysis. Further details on how these sampling weights are computed are available from the NPHS user guide [32, pp. 19].

While sampling weights for the “no exclusions” and the “limited” datasets are provided by Statistics Canada, weights for the “intermediate” dataset were not recomputed for this research. However, the NPHS user guide recommends using the sampling weights from the “no exclusions” dataset on all datasets derived with the Cycle Twinning Approach [32, pp. 69]. The “intermediate” dataset is derived with this approach and therefore its sampling weights can be approximated with the “no exclusion” weights. Additionally, all weights are rounded to the nearest integer for this research.

4.2 Data Filtering and Correction

Despite collection efforts and procedures to minimize errors, some data entries are incorrect and may introduce bias. During survey response collection there are two sources of error. NPHS interview personnel may either enter data incorrectly or the individual being interviewed may misreport answers. The subsequent sections describe how the “intermediate” dataset was further refined to account for errors specifically related to age, pregnancy, weight and height.

4.2.1 Age Range

Although data collection was planned to be performed in two year intervals, a discrepancy in age can occur if an individual is unable to be contacted on the scheduled collection date. The surveyor will attempt to contact participants up to four months after the collection date. If these participants are successfully contacted and have a birthday during the four month grace period, their age will be offset for that cycle. To ensure age always increases in two year intervals, I assume their data was collected on time and adjust their age accordingly. Furthermore the population of interest is only adults (18 and over), so participants under the age 18 were excluded.

4.2.2 Weight and Height

Statistics Canada performed error detection for height and weight values, and then imputes them to more likely values. Statistics Canada has corrected these errors for both the "limited" and "no exclusions" datasets. However, they acknowledge further outlier detection may be useful.

Height corrections were applied to two age groups by Statistics Canada to ensure individuals have plausible height profiles. Individuals from age 18-65 are assumed to be fully grown and their height should remain stable. The height of those above 65 was allowed to decline in height 1-3 inches for the remainder of their survey. Any heights violating these conditions have been imputed by Statistics Canada, the details of which can be found in the NPHS user guide [32, pp. 74].

For weight, errors are commonly caused by confusing measurement scales (kilograms vs pounds) or by transposing numbers (e.g. 160lbs is reported when the actual weight is 106lbs). In addition to error correction by Statistics Canada, two additional outlier procedures are used in this thesis to identify unlikely weight changes between survey cycles. The first procedure identifies individuals who have transitioned from underweight to obese, obese to underweight or overweight to underweight between cycles. The second procedure identifies individuals whose weight change exceeds 100lbs over a two year cycle. The outlying observations for each of these individuals are then corrected using the same weight imputation method as Statistics Canada [32, pp. 77], whereby the imputed weight is the average of the two nearest non-missing observations surrounding the outlier.

After applying the imputation methods, an individual's Body Mass Index (BMI) is calculated using,

$$BMI = 703 * \frac{Weight}{Height^2},$$

where weight is reported in pounds and height in inches. Subsequently their continuous BMI is transformed to one of four discrete states consisting of underweight, normal weight, overweight and obese with the classification scheme in Table 3.

Table 3 - Conversion table for BMI and weight status

Weight Status	BMI (kg/m^2)
Underweight	<18.5
Normal weight	18.5 > 24.99
Overweight	25-29.99
Obese	>30

4.2.3 BMI Correction

As reviewed in Section 3.1, Shields *et al.* [37] identified a bias in self-reported height and weight data and proposes the following corrections which I apply to the NPHS data.

$$Female\ Adjusted\ BMI = -1.08 + 1.08 * BMI$$

$$Male\ Adjusted\ BMI = -0.12 + 1.05 * BMI$$

4.2.4 Pregnancy

The NPHS user guide requires any weight-related research to exclude a women's cycle observation if they are between the ages 18 and 49 and answered "Yes", "Don't Know", or "Refusal" to the pregnancy question. Otherwise, their BMI for that cycle may be inflated which would result in an overestimation of the women's weight. Women older than 50 are assumed not to be pregnant. Removing pregnant women from the data reduces population representativeness, so it is important to note that results of this research omit pregnant women.

4.3 Modelling Transition Probabilities

The probability of an individual transition from a given state *i* to a new state *j* is defined as a transition probability. For simplicity and ease of understanding I start by describing transition probabilities for a homogenous population while ignoring age and gender. The model is a first order discrete-time Markov model. We can define a mutually exclusive finite set of states indexed by $M_t \in \{1,2,3,4\}$, representing underweight, normal weight, overweight and obese weight status, respectively, at time $t = \{0,1,2,...,T\}$. This provides a mathematical representation of how an individual's weight status changes over two year periods. For example the longitudinal survey data, $\{M_0, M_1, M_2, M_3\}$ could be realized as {2, 2, 2, 3}, where a person maintains their initial normal weight status until t = 2 at which time transition to overweight.

The defining feature of a Markov chain is the Markovian property or the memoryless property, which describes how individuals transition from a current state $i = M_t$, to a future state $j = M_{t+1}$. This property says that the probability of transitioning to any future state is dependent only upon the current state. Let p_{ij} be the probability that an individual in weight status i will transition to weight status j. Let $\boldsymbol{P}_{ij}$ be a matrix of all p_{ij}'s, as shown in Table 4. Note that the transition probabilities within the matrix must be non-negative and the sum across each row must equal one since an individual only exists and transitions within these four states.

Table 4 - Transition matrix with only four weight states

		Future State ($\boldsymbol{j}$)			
		1	2	3	4
	1	p_{11}	p_{12}	p_{13}	p_{14}
Current state ($\boldsymbol{i}$)	2	p_{21}	p_{22}	p_{23}	p_{24}
	3	p_{31}	p_{32}	p_{33}	p_{34}
	4	p_{41}	p_{42}	P_{43}	p_{44}

To account for a non-homogenous population gender and age are now incorporated into the formulation. The new state space at time t is defined as $X_t = \{M_t, A_t, G_t\}$ where,

$i = X_t$ is the current state

$j = X_{t+1}$ is the next state

$M_t \in \{1,2,3,4\}$ and indexes m weight statuses

$A_t \in \{18,19, \dots ,80\}$ and indexes a ages

$G_t \in \{0,1\}$ and indexes g genders; male (g = 0) and female (g = 1)

Due to data limitations, participants can only change states every other year leading to the following relationships between the state variables: $A_{t+2} = A_t + 2 \; \forall \; t$ and $G_{t+2} = G_t \; \forall \; t$. M_{t+2} is a function of M_t and $\boldsymbol{P}_{ij}$ and is computed with a random sampling procedure discussed in Section 4.4.

This structure of the resulting transition probability matrix is summarized in Table 5. Note that all ages between 18 and 80 are accounted for in the state space, however transitions only occur on a two year time step, as per the available data. As individuals age past 80 they are removed from the simulated population and are no longer tracked. The transition probability structure only shows the male gender (g = 0), however the structure is identical for females.

Three methodologies are used to estimate transition probabilities of this matrix structure. The first computes the transitions probabilities empirically from the NPHS survey data. The second uses local weighted scatterplot smoothing, and the third uses multinomial logistic regression.

Table 5 - Transition matrix disaggregated by age, gender and weight status

P_{ij}		Future state: $j = \{M_{t+1}, A_{t+1}, G_{t+1}\}$																								
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	…	1	2	3	4
		18	18	18	18	19	19	19	19	20	20	20	20	21	21	21	21	22	22	22	22	…	80	80	80	80
Current state: $i = \{M_t, A_t, G_t\}$	1,18	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	0	0	0	0	...	0	0	0	0
	2,18	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	0	0	0	0	...	0	0	0	0
	3,18	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	0	0	0	0	...	0	0	0	0
	4,18	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	0	0	0	0	...	0	0	0	0
	1,19	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	...	0	0	0	0
	2,19	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	...	0	0	0	0
	3,19	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	...	0	0	0	0
	4,19	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	0	0	0	0	...	0	0	0	0
	1,20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	...	0	0	0	0
	2,20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	...	0	0	0	0
	3,20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	...	0	0	0	0
	4,20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p_{ij}	p_{ij}	p_{ij}	p_{ij}	...	0	0	0	0
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮
	1,78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...	p_{ij}	p_{ij}	p_{ij}	p_{ij}
	2,78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...	p_{ij}	p_{ij}	p_{ij}	p_{ij}
	3,78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...	p_{ij}	p_{ij}	p_{ij}	p_{ij}
	4,78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...	p_{ij}	p_{ij}	p_{ij}	p_{ij}

4.3.1 Empirically Derived Probabilities

The empirically derived probabilities are computed directly from the NPHS data. Each p_{ij} is computed by summing the number of individuals in state i who transition into state j and then dividing by the total number of individuals leaving state i. Let n_{ij} be the number of individuals who transition from state i to j, weighted by the NPHS correction used to account for design effect and non-response (section 4.1.3). It follows that,

$$p_{ij} = \frac{n_{ij}}{\sum_{k=1}^{4} n_{ik}}, \forall i, j$$

This approach and its use with longitudinal data are discussed in Kalbfleisch [44].

For some combinations of age, gender and current weight status the NPHS data may not contain any observations of participants transitioning to the subsequent state. In these cases no transitions probabilities are estimated, and an individual who enters as this state could not transition elsewhere. For these missing transition probabilities their values are imputed to the transition probabilities from the previous state, or the first non-missing previous state.

4.3.2 Locally Weighted Regression Smoothing

If less than 5 individuals are recorded for any state transition, the data cannot be exported from the RDC due to the possibility of participant identification. Smoothing of the transition probabilities prevents any chance of participant identification and is exportable from the RDC. Additionally, smoothing removes some of the variability of the empirical transition probabilities caused by limited data. Locally weighted scatterplot smoothing (LOWESS) is a non-parametric regression method which uses a series of linear or non-linear regression models of the same scatterplot to predict the independent variable [84][85].

A single subset of transition probabilities is defined by gender, a current weight status, and a subsequent weight status. For example, there is one subset for normal weight males who are transitioning to normal weight males; another subset for normal weight males transitioning to overweight males, etc. In total there are four different current weight statuses, four subsequent weight statuses and two genders accounting for 4 × 4 × 2 = 32 subsets. The LOWESS regression technique, summarized below, is applied to each of the 32 subsets and is based on [51].

In each subset we have a single transition probability for each age, say P(*Z*). The LOWESS regression technique considers a set of weighted transition probabilities 'centered' on a single age *Z*. The regression is computed on this set and a new smoothed transition probability is computed. This procedure is repeated for all ages.

Formally, let *h* be the total number of data points considered for a local regression for a given age *Z*. In this thesis, the value *h*=39 is chosen by solving *h* = γ*(80-18) [51] where γ = 2/3. This is consistent with Agresti [51], who suggests choosing γ values that are between [0.65-0.75] to ensure the fitted line is not too irregular but does not miss interesting patterns. Let *H* be the set consisting of *h* data points a_k, where k=1,2,…,h and,

$$\mathrm{H} = \begin{cases} [a_1 = 18, a_2 = 19, \dots, a_h = 18 + h\,], & when\ Z - \left\lfloor \frac{h}{2} \right\rfloor < 18 \\ [a_1 = 80 - h,\ a_2 = 80 - h + 1, \dots, a_h = 80], & when\ Z + \left\lfloor \frac{h}{2} \right\rfloor > 80 \\ [a_1 = Z - \left\lfloor \frac{h}{2} \right\rfloor, a_2 = Z - \left\lfloor \frac{h}{2} \right\rfloor + 1, \dots, a_h = Z - \left\lfloor \frac{h}{2} \right\rfloor\,], & otherwise \end{cases}$$

Let P(a_k) be the empirical probability that a participant of age *a* transitions from the current to the subsequent weight status. Finally, each probability P(a_k) is then weighted to give a greater influence to those data points closest to the age *Z*. Let the new probabilities at each age a_k be $P(ak)'$ where,

$$P(a_k)' = P(a_k) * \left(1 - \left|\frac{Z - a_k}{h - Z}\right|^3\right)^3$$

For the given *H*, a quadratic regression is then performed on all $P(a_k)'$s. From this regression equation, the probability at age *Z* is then computed (called $P(Z)'$) which is our smoothed transition probability at age Z. For a formal account of this method, see Cleveland [84].

The LOWESS regression procedure is applied to each age in each of the 32 transition probability subsets. The resulting smoothed transition probabilities can now repopulate the original transition probability matrix structure of $\boldsymbol{P}_{ij}$ (Table 5). However, the new $P(Z)'$ values can be <0 or >1 since LOWESS does not require estimates to adhere to probability bounds. In such cases the values are imputed to 0 and 1 respectively. Furthermore, the sum of the probabilities for any given age can exceed 1. This is accounted for by normalizing the LOWESS predicted values across transition rows to ensure each row sums to 1.

4.3.3 Multinomial Regression

A multinomial regression is similar to logistic regression, but expanded for multiple categorical variables. Diggle *et al.* [50] developed multinomial regression specifically for longitudinal data analysis. The particular approach used is the 'transitional multinomial regression' [51], where the dependent variable is an individual's next weight status while the independent variables are an individual's current age, gender and weight status. A transformation of the current age to age^2 is also included as an independent variable to account for possible non-linear effects of age.

As with the LOWESS methodology, modelling transition probabilities with multinomial regression ensures participant's identities remain confidential. Unlike the LOWESS, the data does not need to be disaggregated into 32 different subsets prior to the regression. Additionally, renormalization is not required since the procedure naturally ensures all transition rows of the $\mathbf{P}_{ij}$ matrix sum to 1. Imputing missing transition probabilities is also not required since the regression can handle a certain amount of sparsity in the data. Too much sparsity existed to perform the regression with the limited dataset, prompting the use of only the intermediate dataset for this research.

The multinomial model is fit using the intermediate subset in order to maximize the number of available observations for the regression procedure. Additionally, the NPHS sampling weights are given to the multinomial function to account for unequal probability of selection. Probabilities for the transition matrix (Table 5) are then calculated using the estimated coefficients from the multinomial regression, generated within the RDC.

4.4 Microsimulation Model

A simulation of the NPHS survey is performed from 1994 to 2030 using the intermediate dataset to derive all transition probabilities and initial parameters. The microsimulation tracks participants and their weight statuses through time according to the following process:

At cycle t=0, the model is initialized by populating it proportionally with individuals in every age, gender and weight status category.

At cycle t+1, the following procedures are run:

I. The model advances 2 years in time and each individual is aged by 2 years. Those who are over the maximum age limited are removed.
II. A weight status is assigned to each individual based on their previous age, gender, and weight status. This is modelled by a random sampling procedure based on the transition matrix and is described below.
III. Each individual also has a probability of dying and is removed from the model following a random sampling procedure.
IV. New participants aged 18 and 19 are brought into the model.

Subsequent cycles: The steps described for cycle t+1 are then repeated for subsequent cycles until the simulation horizon is reached.

Post metric processing: Output metrics are then computed.

This simulation uses cumulative uniform random sampling to determine how individuals transition through weight statuses. For an individual's current state, i, the transition row of the P_{ij} matrix is first transformed to its cumulative distribution. For example, the transition row {0.2, 0.4, 0.3, 0.1} would be transformed to {0.2, 0.6, 0.9, 1}. Next a uniform random number, U, is generated between [0,1] and the subsequent weight status is determined by,

$$M_{t+1} = \begin{cases} Underweight & for \quad 0.0 \leq U < 0.2 \\ Normal & for \quad 0.2 \leq U < 0.6 \\ Overweight & for \quad 0.6 \leq U < 0.9 \\ Obese & for \quad 0.9 \leq U \leq 1.0 \end{cases}$$

A similar procedure is used to remove individuals who have died. Let $P(Death)$, be the probability of dying. When U' (a uniform random number between [0,1]) is less than $P(Death)$, the person dies and is removed from the model.

This model is applied to two different time periods. The first simulation period from 1994-2010 is used to validate the model by comparing the predicted weight status prevalence to the observed NPHS data. The second period, from 2010-2030, is used for forecasting and is based on only transition probabilities derived from the LOWESS and multinomial regression data (which may leave the secure RDC). There are a number of subtle changes to the above model description for each period which are described in the following sections.

4.4.1 Validation Period

The model is validated over the 1994-2010 NPHS survey period using each of the transition probability methodologies. The validation simulation begins in a populated state with all survey participants who were between the ages 18-80 in 1994. Therefore, in 1994 there is no difference between the actual and simulated states.

The incoming population consists of participants who were enrolled as children and turned either 18 or 19 in any survey year after 1994. These 'births' are appropriately merged into the population each year before applying the simulation. Rarely 'births' entering the simulation have missing weight status, in which case their initial observation is imputed to their next available weight status observation.

Since the actual population is being simulated we know in which cycle each participant dies. I account for this by setting $P(Death) = 1$ during the appropriate cycle. Furthermore, from the NPHS data we know when participants missed a cycle or were pregnant during the cycle. To reflect this in the simulation output, I retroactively remove simulated data points which correspond to these absences in data. Finally, the sampling weights for each individual are applied to each of their observations from 1994-2010 to account for regional sampling bias.

To formally compare each transition probability methodology the mean percentage error (MPE) is calculated for each weight status prevalence through time. The MPE can be used to compare forecasted versus observed time series data. It is defined as follows:

$$MPE_m = \frac{100}{T}\sum_{t=1}^{T}\frac{f_{t,m} - a_{t,m}}{a_{t,m}} \quad \forall\, m$$

Where,

$t \in 1,2,\ldots,T$ $simulation\ cycles$

$f_{t,m}$ $is\ the\ forecasted\ weight\ prevalence\ m\ for\ simulation\ cycle\ t$

$a_{t,m}$ $is\ the\ historical\ weight\ prevalence\ m\ for\ survey\ cycle\ t$

For example, if predicted normal prevalence at time *t* was 0.4 and the observed prevalence was 0.8, the percent error would be -50%. The MPE, as above, is the average of these percent errors across all *t*.

4.4.2 Forecasting Period

The forecasting period represents our reference scenario from 2010-2030, starting with an initial sample of 500,000 participants to represent the Canadian population. The age, gender and weight status for these participants is distributed according to the weighted 2010 data from the validation period. The initial distribution for ages 20-80 were simulated and could be removed from the RDC without restriction. The initial distribution for the incoming 18 and 19 year olds consists of real NPHS data, requiring that their data be aggregated over both ages to avoid disclosure restrictions. These two distributions are then combined outside of the RDC to form the initial forecasting distribution in 2010. Unlike the validation simulation, each of the 500,000 participants is assigned an equal sampling weight of 1 when computing the output metrics.

Births

For each cycle the incoming number of 18 and 19 year olds must be estimated. These simulation 'births' are estimated use the birth rate from 18 and 19 years prior to the current simulation year. This data is available from the World Development Indicators [86]. However, we are simulating only a fraction of the Canadian population and the number of births in each cycle must be proportional to our simulated population. This fraction is calculated by dividing the initial sample size (500,000) by the population (34,005,274) in 2010 [86], and is assumed to be constant through time. For instance, the number of modelled births is 2012 is 1,121, since the birth rate in 1994 was 0.0131 (births/person) and the population was 29,111,906, as shown below:

$$Births_{2012} = \frac{500{,}000}{34{,}005{,}274} * 0.0131 * 29{,}111{,}906 = 1{,}121$$

After calculating the number of births for each simulation cycle, each individual is assigned an age, gender and weight status according to the distribution for ages 18 and 19 (Table 6). This 'birth' distribution is derived from the 2010 NPHS data, for which 18 and 19 year olds are aggregated due to disclosure restrictions.

Table 6 - Incoming participant demographic distribution

Gender	Age	Weight Status			
		Underweight	*Normal Weight*	*Overweight*	*Obese*
Male	*18*	0.0132	0.1277	0.0677	0.0343
Male	*19*	0.0132	0.1277	0.0677	0.0343
Female	*18*	0.0100	0.1873	0.0407	0.0191
Female	*19*	0.0100	0.1873	0.0407	0.0191

Deaths

For all simulation cycles the probability of death for each individual is based on their age, gender and weight status. Two information sources are combined to calculate the probabilities of death for the Canadian population. The first source is the 2010 age-gender specific probabilities of death, $P(Death \mid a, g)$, from Statistics Canada [87]. The probability of death is assumed to remain at 2010 levels throughout the simulation. To accommodate the 2 year time step in the model, I combine the two death probabilities as follows,

$$P\ (Death' \mid a, g) = 1 - [1 - P\ (Death \mid a, g)] * [1 - P\ (Death \mid a - 1, g)]\ \forall\ a, g$$

The second source is from a study by Orpana *et al.* [88], which reports the relative risk (RR_m) of death based on weight status where normal weight is the reference category (Table 7).

Table 7 - Weight-related relative risks of death

Weight Status	Relative risk of death (RR_m)
Underweight	1.73
Normal weight	1.00
Overweight	0.83
Obese[1]	1.05

[1]Weighted average of relative risks for obese classes I and II.

Using these two data sources the death probabilities are updated every simulation cycle, by taking into account the population's weight status distribution in each cycle. We want to compute the death probabilities per age, gender *and* weight status $P\ (Death \mid m, a, g)$ that ensures both the age-gender specific probability of death, $P(Death' \mid a, g)$, and RR_m's scores are adhered too. First, consider that $P(Death' \mid a, g)$ can be expressed using the following equation where $P\ (Death \mid m, a, g)$ is unknown:

$$P(Death' \mid a, g) = \frac{\sum_M (w_{m,a,g} * P\ (Death \mid m, a, g))}{\sum_M w_{m,a,g}} \quad \forall\, m, a, g$$

where, $w_{M,A,G}$ represents the number of people in each age, gender and weight status stratum. These values are known during each simulation cycle. Rearranging, the probability of death for normal weight (which has a $RR_m = 1$), can be computed as follows:

$$P(Death \mid 2, a, g) = \frac{P(Death' \mid a, g) * \sum_M w_{m,a,g}}{\sum_M RR_m\, w_{m,a,g}} \ \forall\, m, a, g$$

Using this as a reference point, the remaining unknown death probabilities can be computed as follows:

$$P(Death \mid m, a, g) = P(Death \mid 2, a, g) * RR_m \quad \forall\, m, a, g$$

4.5 Simulation Metrics

The primary metric is the weight status prevalence for each of the weight categories in each year of the simulation. Weight status prevalence is the percentage of the population in a given weight status.

The secondary metric is Quality Adjusted Life Years (QALYs), which are calculated by multiplying time spent in a health state and the utility of that health state. Steensma *et al.*

[89] derived age, gender, weight status specific utilities using NPHS survey data. Individual utilities are determined from health related quality of life measures that assesses basic senses, mobility, dexterity, feelings, cognitive ability and pain among others. Utility as determined from the Health Utility Index falls between [-0.36, 1.00] where 0 represents death, the lower limit represents a state worse than death and the upper limit is perfect health [32]. The health utilities used to calculate each individual's QALYs are included in Appendix A.

Individuals value current health more than future health [90], so the discounted present value of the QALYs for each gender is calculated as,

$$DPV\left(QALYs_g\right) = \sum_{t=0}^{T} \frac{QALYs_{g,t}}{(1+r)^t} \ \forall\, g$$

where,
r is the discount rate
t is the year
g is the gender

I assume the discount rate is 3.5%, as per [91].

A common metric when deciding whether to implement a health care program is the incremental cost-effectiveness ratio (ICER) [92]. ICER weighs the incremental gain (or loss) in QALYs of a scenario against the costs required to achieve these benefits. ICER is calculated by,

$$ICER \ = \ \frac{C1 - C2}{Incremental\ QALYs} = \frac{C1}{Incremental\ QALYs}$$

where,

$C1$ is the cost of a prevention or intervention scenario

$C2 = 0$ is the cost of the baseline (i.e. taking no action)

$Incremental\ QALYs\ =$ Intervention scenario QALYs – Reference scenario QALYs

In our case *C2* is 0 since no action is taken during the reference scenario forecast. The costs of interventions are not precisely estimated for this research, so *C1* is not known. From Laupacis [92], we know that an intervention is considered very effective when ICER is not more than $20,000 / QALY. As such, for each proposed interventions I set ICER=$20,000 and compute *C1*. This implies that if the intervention can be completed at a cost of *C1* then it will be very effective. Similarly, I repeated this calculation for interventions deemed "moderately effective" and "poorly effective" (see Table 7). For the purposes of this research moderately effective is assumed to have an ICER of $60,000/QALY. Finally, for each scenario the intervention cost per person is calculated by dividing *C1* by the number of individuals taking part in the intervention.

Table 8 - Incremental cost effectiveness ratio thresholds

Intervention	ICER ($/QALY)
Very effective	<20,000
Moderately effective	20,000-100,000
Poorly effective	>100,000

4.6 Prevention and Intervention Scenarios

The simulation framework enables the analysis of potential interventions aimed at reducing overweight and obesity. These 'what if' scenarios are simulated by changing inputs. The first scenario projects the results of an existing school-based intervention program. The second scenario considers a bariatric surgery intervention option for a selected proportion of the population. The third scenario alters the transition probabilities across the whole population to mimic the effects of a nation-wide prevention program.

To determine the impact of each scenario the simulation metrics described in the previous section are computed and compared to the reference forecast. The weight status prevalence is reported using only one replication since prevalence results are not significantly different across multiple runs. The variation across replications for QALYs however, is considerably higher. The QALY metric is calculated by using 30 scenario replications and creating a confidence interval of the difference between each replication's QALYs and the reference scenario's QALYs. Assuming the 30 metric calculations are normally distributed, the student-t test is used to create confidence intervals as recommended by Rossetti [93].

4.6.1 School-based Intervention Scenario

The first scenario draws from the results of the evaluation of a children's school-based health promotion program [16]. The researchers found a 2.2 percentage point absolute obesity reduction in intervention schools compared to control schools. The 2.2% reduction of childhood obesity is assumed to carry forward to 18 and 19 year olds for this intervention scenario. Thus, the incoming birth distribution is adjusted so that the probability of entering as an obese 18 or 19 year old is reduced by 2.2 percentage points, while the probability of entering as a normal weight individual is increased by 2.2.

This scenario is run as a cohort simulation (no births or deaths) where an initial sample of 20,000 18 and 19 year olds is selected in 2010 and simulated until they reach age 80. The reference forecast initial sample is determined from the original distribution, whereas the scenario forecast is sampled from the adjusted distribution. To reduce the variation across each replication's initial sample the number of males and females as well as the number of 18 and 19 year olds are forced to be equal to that of the reference scenario's initial sample.

4.6.2 Bariatric Surgery Intervention

This scenario is modelled as a population simulation (includes births and deaths) with an initial sample size of 500,000 individuals. An adult intervention program is implemented where a number of obese individuals are selected for bariatric surgery each year. Only obese individuals aged 18-65 are considered eligible for this surgery [94]. The Canadian healthcare system is assumed to be capable of performing 10,000 per year for the entire Canadian population. This number of yearly surgeries is adjusted based on the proportion (0.01467) of Canadians being simulated in 2010 (see Section 4.4.2).

According to Richardson and colleagues [95] approximately 70% of bariatric patients achieve significant weight loss during the first three years after surgery. For this scenario 70% of surgery patients each year are assumed to transition to normal weight and remain in normal weight status for the length of the simulation. The other 30% are assumed to remain in the obese category and transition according to the original transition probabilities.

The variation of the initial sample for each scenario replication is eliminated by using the same sample as the reference forecast. The variation of incoming births is also eliminated for each replication by using the same sample of incoming births as the reference scenario. When calculating the scenario cost per person the intervention population is slightly different for each replication. To the simplify calculations the number of individuals chosen for surgery is taken to be the average number of individuals chosen across all replications.

4.6.3 Primary Population-wide Prevention

The final scenario modifies transition probabilities to assess the potential impact of a population-wide prevention program. This primary prevention would hypothetically encourage individuals to remain in their current weight status rather than transitioning to a higher one. The transition probabilities of remaining in normal and overweight categories $\forall\, A, G$ is increased by 2 percentage points. In addition, transition probabilities from normal to overweight, and overweight to obese are reduced by 2 percentage points of their estimated values $\forall\, A, G$.

As with the previous scenario, the initial sample and incoming births are set to be identical to the reference forecast scenario. The total population affected by the intervention is the sum of the initial sample (500,000) and the number of incoming births over the simulation horizon.

Chapter 5: Model Verification

Law and Kelton [96] recommend various simulation verification techniques. This section describes how several verification techniques are utilized to ensure the simulation model is performing as intended.

Technique 1

Building a simulation model and coding in general should be written in small segments and tested continuously as more detail is added. This helps reduce confusion while debugging and provides a meaningful structure for those who may review the code. For this thesis, separate functions were written for the various steps of data manipulation, transition modelling and the simulation. Law and Kelton also advise the modeller to build complexity into a model sequentially. For instance, the forecasting simulation was built sequentially. First, the initial population was simulated through weight statuses using only the transition probabilities. After verifying results were being generating as anticipated, births and deaths were added to the model and again checked for consistency with expectations.

Technique 2

Law and Kelton also recommend running the model under simplifying assumptions for which its true characteristics are known. While performing the validation simulation within the RDC I compared the number of individuals in the simulation to the observed NPHS data. This led to the discovery of syntax errors as well as unexpected NPHS data anomalies. For example a coding error in the age correction was identified, and unexpected missing weight statuses in the NPHS data were discovered.

Technique 3

Another verification technique suggested by Law and Kelton is ensuring that the simulation input distributions are consistent with the data from which they are derived. For the forecasting simulation a new cohort of 18 and 19 year olds enter the model each year. The age, gender and weight status distribution of these incoming individuals is verified by performing a chi-square test to compare the sampled and expected counts in each stratum (Appendix B). The test only includes 18 year old males and females since the distribution across 18 and 19 year olds is identical. The chi-square test results in a p-value of 0.22, indicating that we cannot reject the null hypothesis that these two distributions are the same. This confirms that our initial sample size is large enough to ensure incoming births are consistent with the NPHS data.

Technique 4

The final verification technique used in thesis is checking that simulation results are reasonable. For this technique, I verified that the modelled transition probabilities adhere to the assumption that they are stationary through time. Transition probabilities through time for each gender were recalculated (Appendix C) from the forecasting period simulation results. Actual transition probabilities by year from the observed NPHS data could not be released from the RDC due to disclosure restrictions.

Chapter 6: Results

This section first provides the derived transition probabilities in graphical form for LOWESS and multinomial methodologies. Next, results from the validation simulation period (1994-2010) are reported. Finally, results of the forecast simulation and scenario analyses are reported.

6.1 Descriptive Results

Figures 1-4 depict the probability of transitioning from a current weight status to a future weight status, stratified by age, gender and current weight status. The transition probabilities were derived from the LOWESS and multinomial methodologies, respectively (Section 4.3.2 - 4.3.3). Empirical transition probabilities were not released due to disclosure restrictions. When comparing the LOWESS and multinomial transition probabilities, it is important to note the scales on the x-axis are different. The scale for the LOWESS ranges from ages 18 to 80 years, while the scale for multinomial covers ages 18 to 104 years.

Transition Probabilities for Males

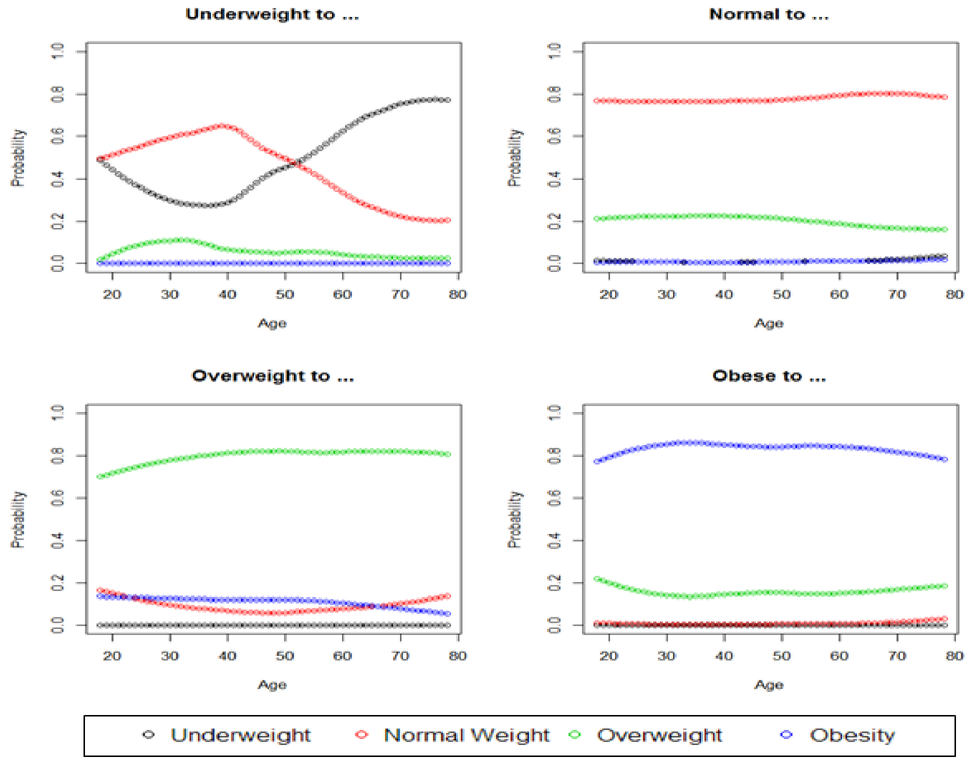


Figure 1 - Transition probabilities by age for males, derived from the LOWESS method

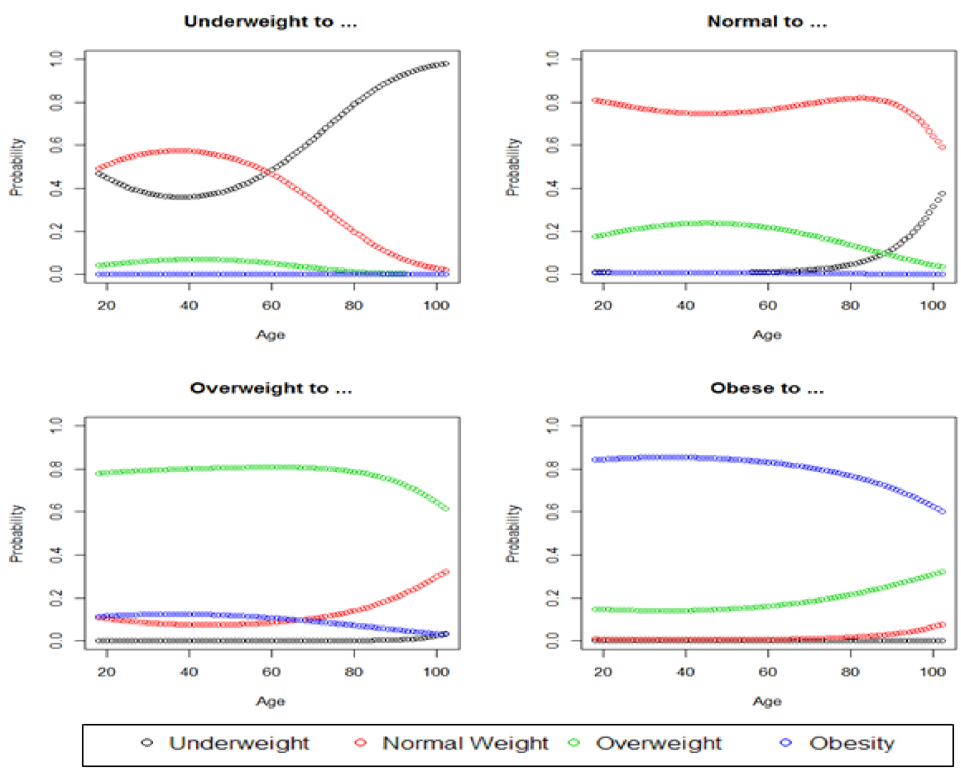


Figure 2 - Transition probabilities by age for males, derived from the multinomial method

Transition Probabilities for Females

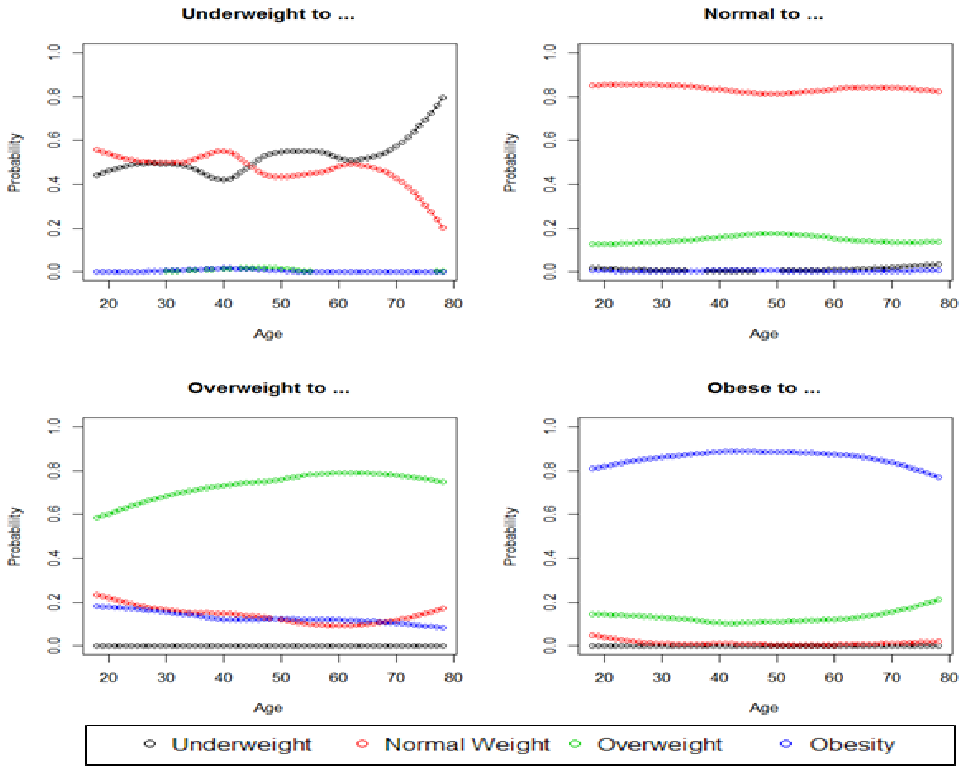


Figure 3 - Transition probabilities by age for females, derived from the LOWESS method

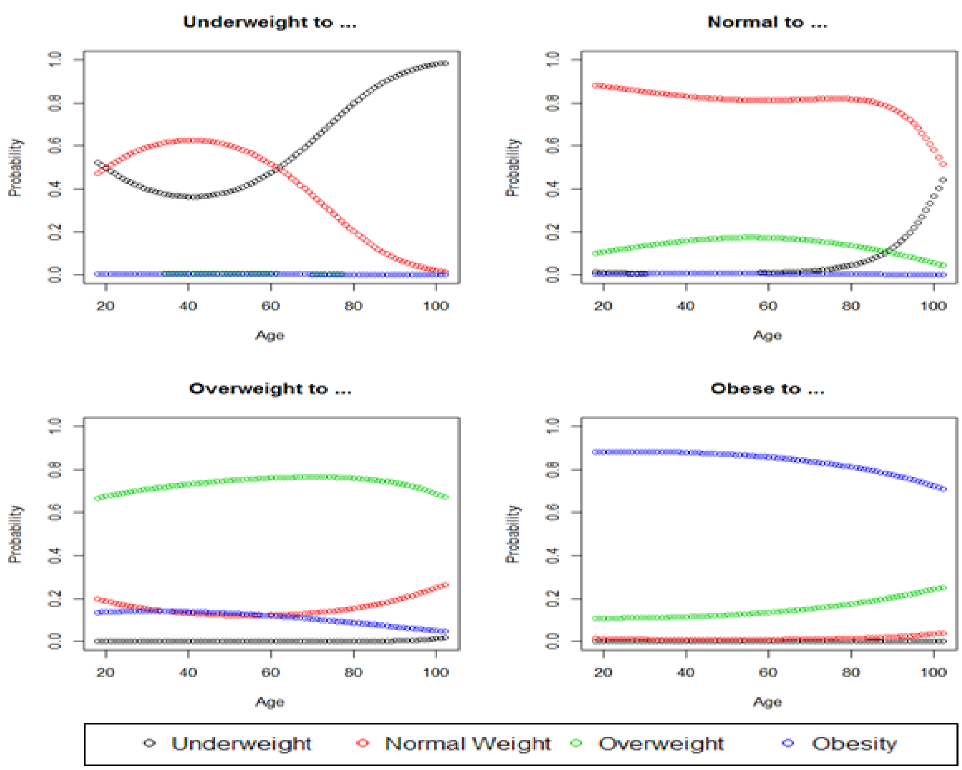


Figure 4 - Transition probabilities by age for females, derived from the multinomial method

The most evident difference across all graphs is observed in the current underweight transition probability profiles. Although the LOWESS and multinomial probabilities follow the same general pattern, they deviate by as much as 10%. This is attributed to a lack of empirical data for estimating the underweight transition behaviour. As age increases there is also less available data which contributes to large deviations in transition probabilities for those above 80. For both males and females, the transition probabilities of the youngest ages are also notably different when comparing the two approaches. The LOWESS overweight to overweight and obese to obese status transition probabilities are 3-4% higher than those of the multinomial procedure. This difference diminishes in the late 20's, except for female overweight to overweight status transitions which do not converge to similar values until approximately age 40.

In both approaches older participants tend to transition to lower weight statuses. For example, Figure 2 shows the declining transition probabilities for male obese to obese weight status. In terms of gender differences, the probabilities of remaining in the same current weight status for normal, overweight and obese are higher in females than in males. This indicates females are more resistant to weight change throughout the duration of adulthood than their male counterparts.

In addition to creating weight status prevalence time series plots, the weight status prevalence by age was examined. The weight status prevalence by age for the NPHS cannot be released from the RDC, so only simulated results are provided. Nonetheless, the same general trends are present in the observed NPHS data. Figures 5 and 6 are generated from the 1994-2010 validation period simulation using multinomial transition probabilities. Results for the LOWESS method were sufficiently similar.

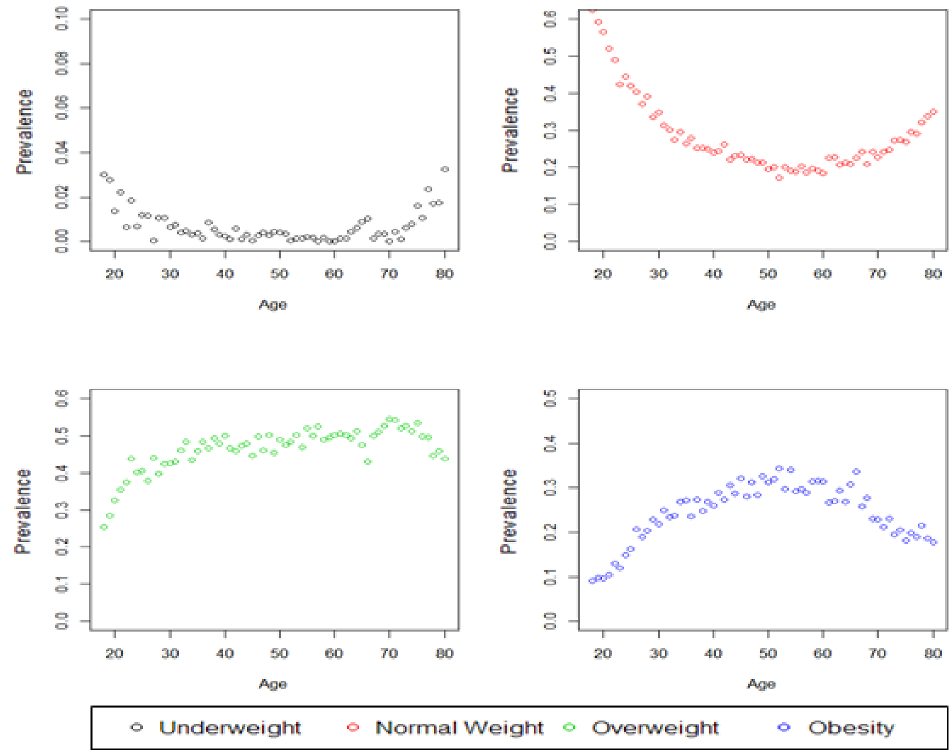


Figure 5 - Simulated weight status prevalence by age among males for the 1994-2010 validation period

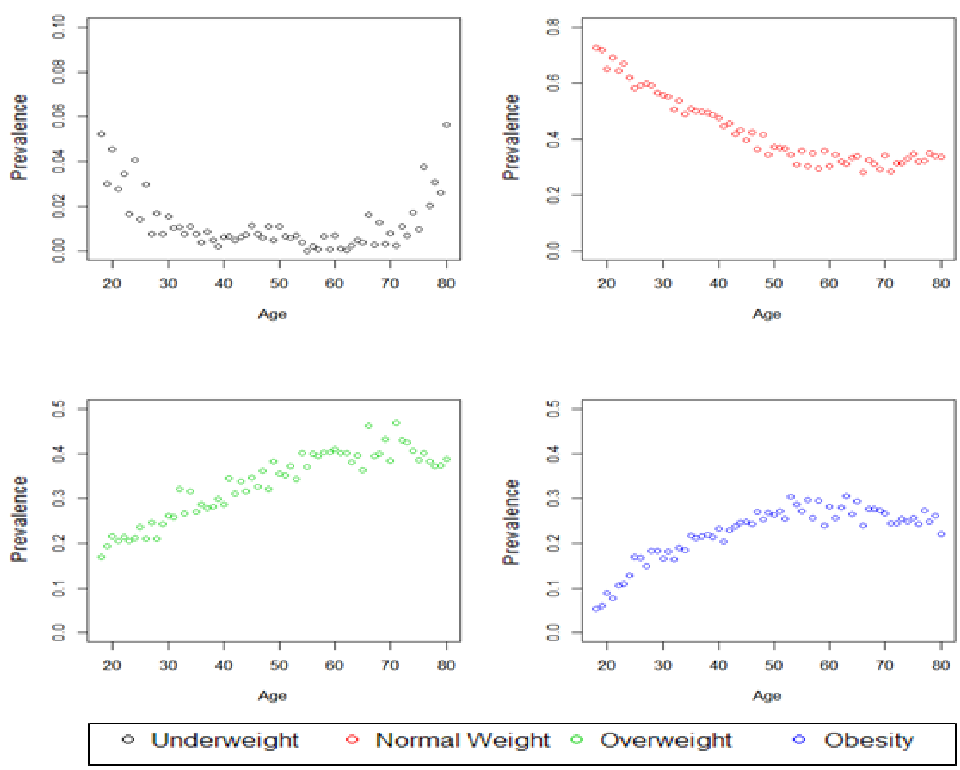


Figure 6 - Simulated weight status prevalence by age among females for the 1994-2010 validation period

For males aged 18-35 a sharp decrease in the prevalence of normal weight status is observed. This corresponds to significant increases in the prevalence of overweight and obesity over the same age range. Obesity in the male population peaks from ages 50-60 and declines from ages 60-80. For females, the prevalence of normal weight declines approximately linearly from age 18-50 while the prevalence of both overweight and obesity increase approximately linearly. The prevalence of both overweight and obesity continue to rise until plateauing at approximately age 50 and 60 for males and females, respectively.

The final descriptive data extracted from the RDC is the observed weight status prevalence by year. Figures 7 and 8 illustrate the NPHS weight status trends by gender from 1994-2010.

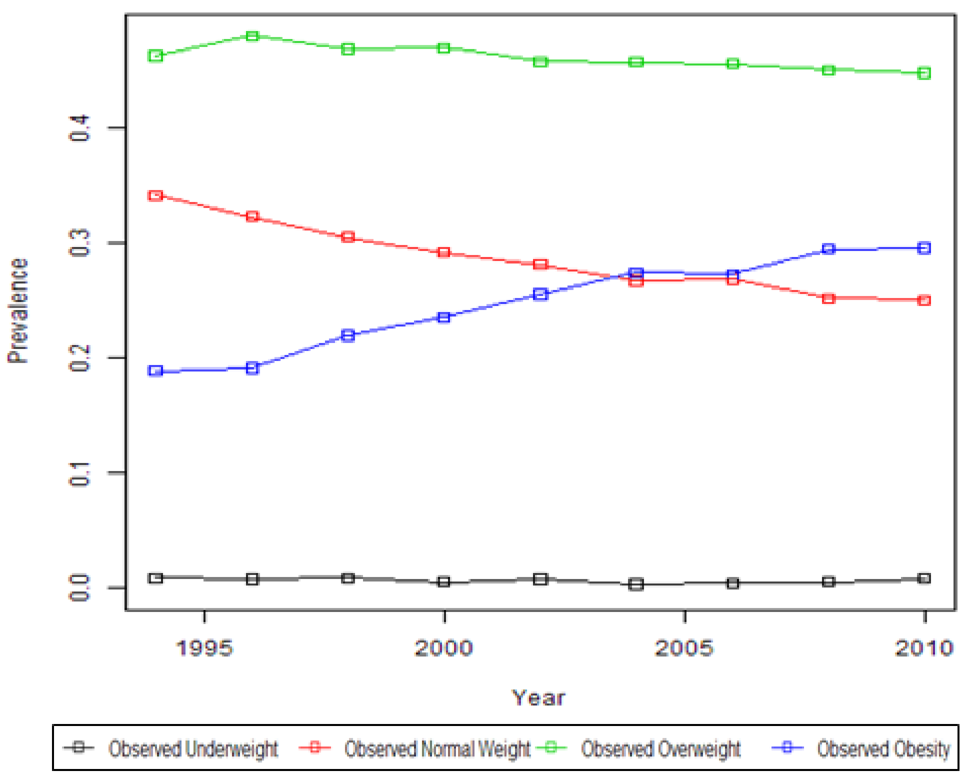


Figure 7 - Observed NPHS weight status prevalence by year among males from 1994-2010

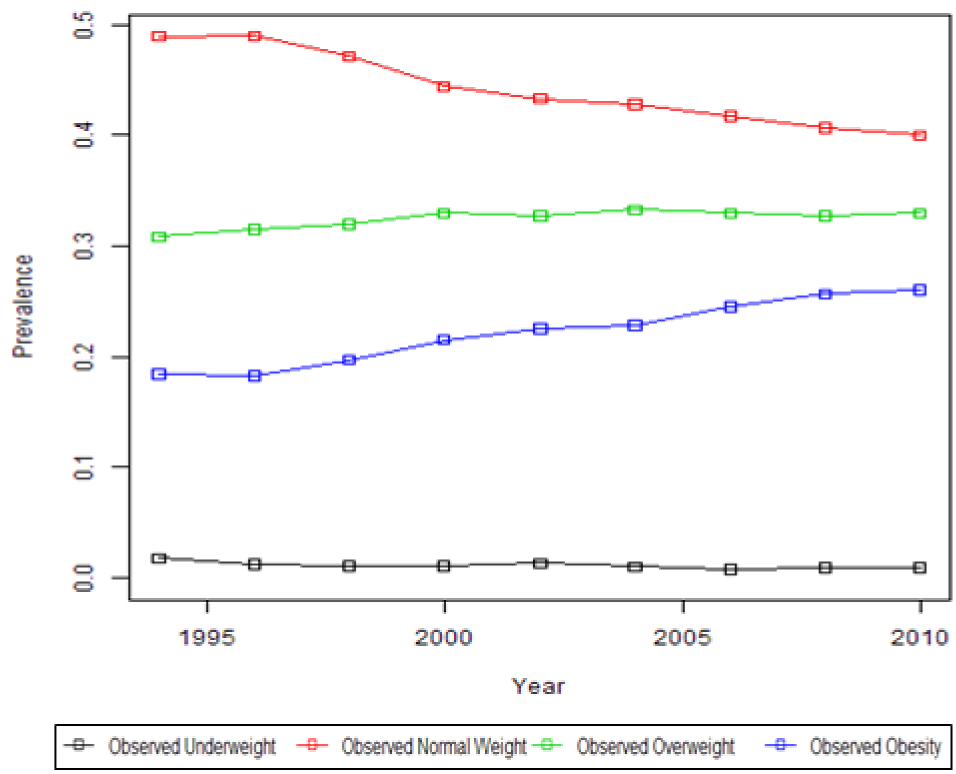


Figure 8 - Observed NPHS weight status prevalence by year among females from 1994-2010

For both male and females the prevalence of normal weight has decreased over time while that of obesity has increased. The prevalence of overweight in males has declined over time, while for females it has risen. For males, the observed obesity prevalence in the NPHS surpassed the prevalence of normal weight in 2004. For females the prevalence of normal weight, overweight and obesity appear to be converging, but at a slower rate.

6.2 Validation Simulation

The validation simulation for ages 18-80 was run using each the empirical, LOWESS and multinomial transition probabilities within the RDC. The simulation output is the prevalence of each weight status stratified by age and gender. The NPHS observed weight status prevalence (stratified by age and gender) is also overlaid on each figure for comparison. Another simulation was run for the validation period with individuals up to age 104 since the range of both the observed NPHS data and the multinomial method can be extended (Appendix D).

Empirical Method Validation

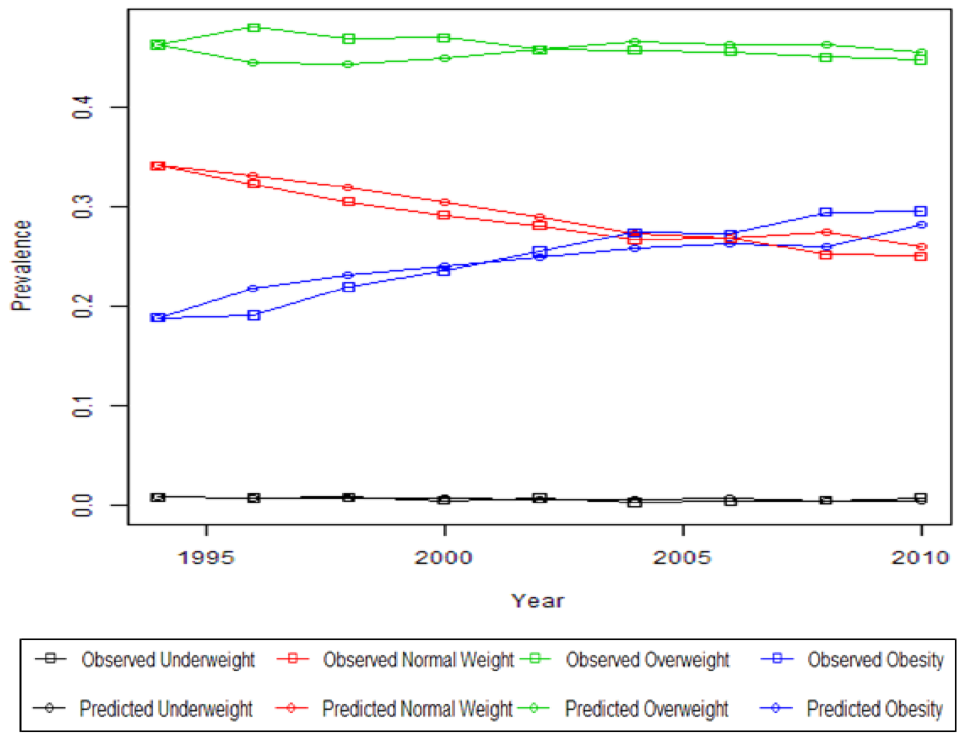


Figure 9 - Predicted versus observed weight status prevalence among males from 1994-2010 using empirical transition probabilities

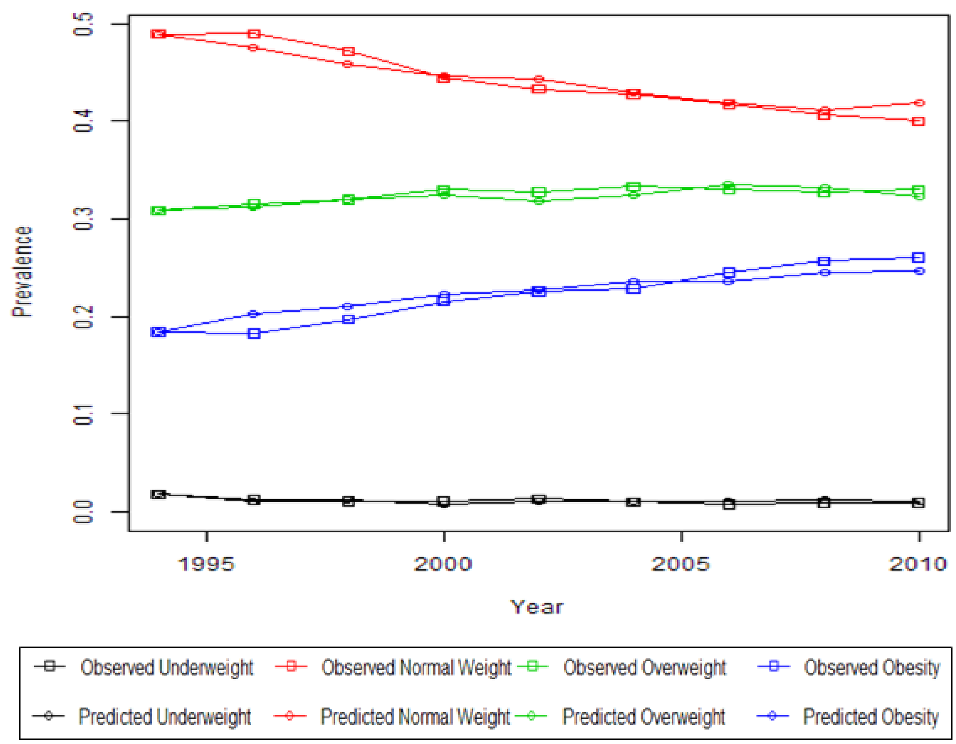


Figure 10 - Predicted versus observed weight status prevalence among females from 1994-2010 using empirical transition probabilities

LOWESS Method Validation

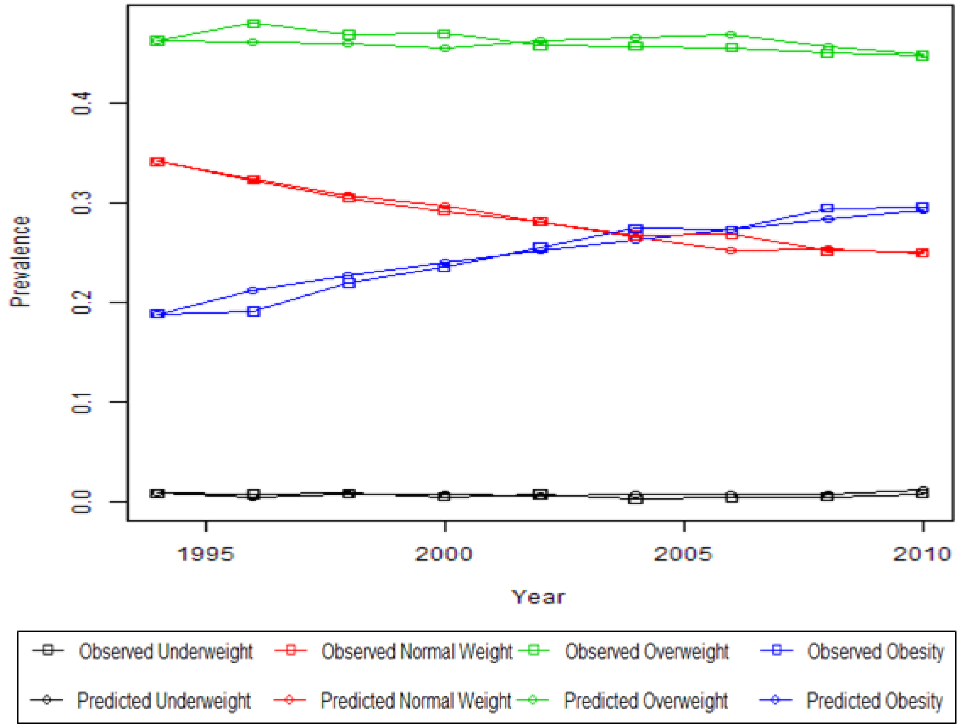


Figure 11 - Predicted versus observed weight status prevalence among males from 1994-2010 using LOWESS transition probabilities

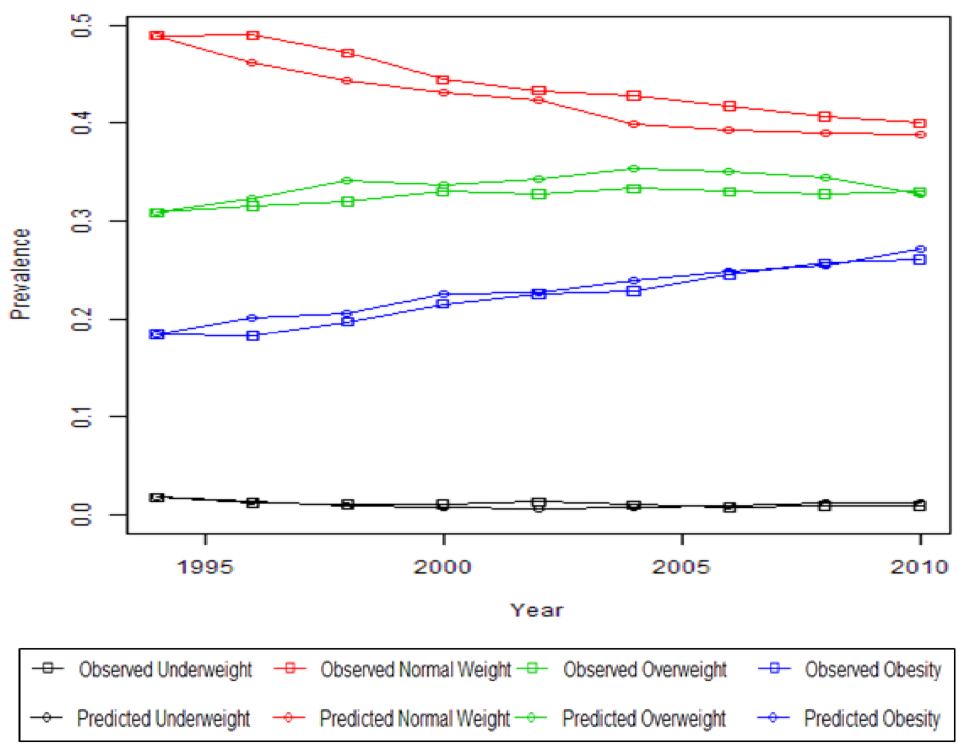


Figure 12 - Predicted versus observed weight status prevalence among females from 1994-2010 using LOWESS transition probabilities

Multinomial Method Validation

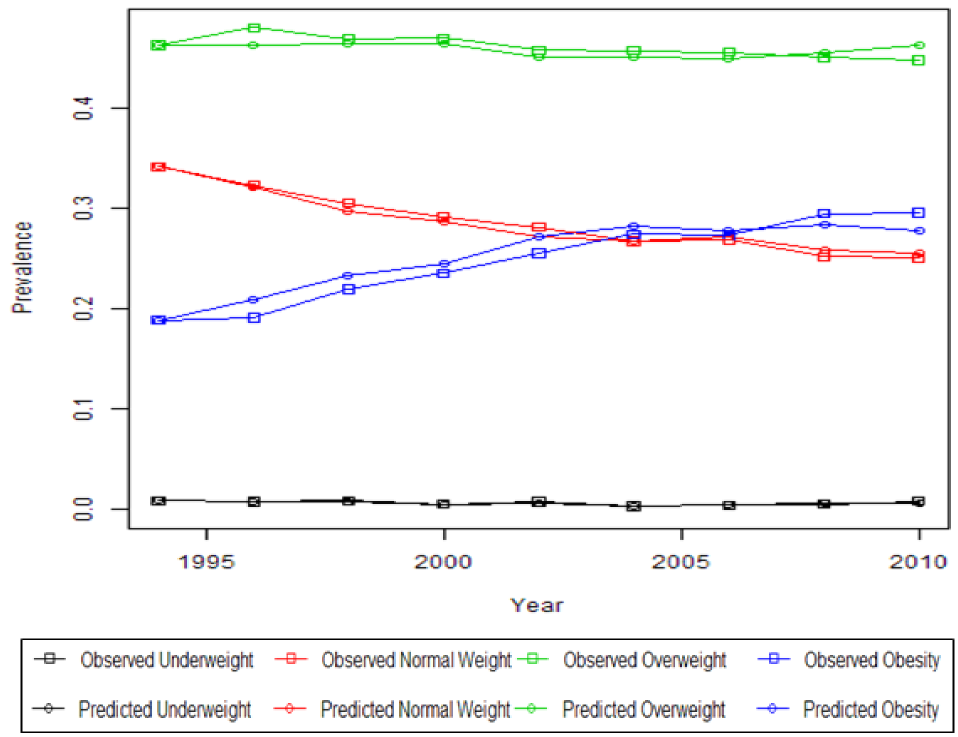


Figure 13 - Predicted versus observed weight status prevalence among males from 1994-2010 using multinomial transition probabilities

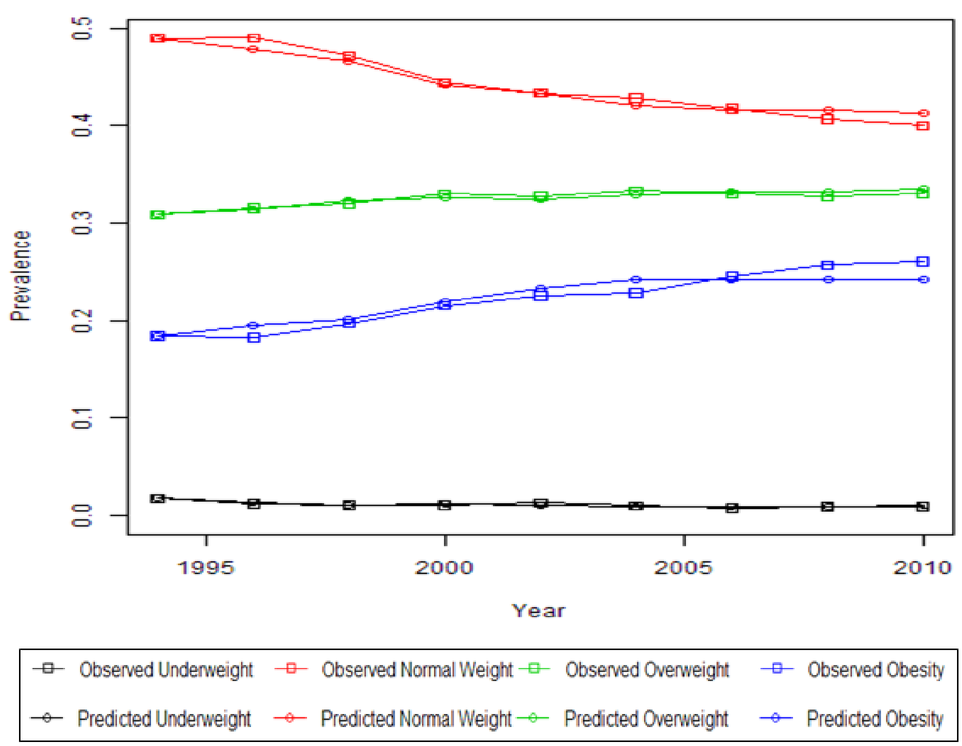


Figure 14 - Predicted versus observed weight status prevalence among females from 1994-2010 using multinomial transition probabilities

By visual inspection, all three methods appear to replicate the observed NPHS data closely. To compare this formally the mean percentage forecast error, as described in Section 4.4.1 is calculated for each transition probability method and each weight status prevalence forecast. The following tables illustrate the MPE for males and females.

Table 9 - Mean percentage forecast error for each transition probability method and weight status for males from 1994-2010

Transition Probability Method	Mean Percentage Forecast Error (%)			
	Underweight	*Normal*	*Overweight*	*Obese*
Empirical	-17.58	-3.73	1.13	1.03
LOWESS	-32.22	0.46	0.13	-0.81
Multinomial	4.45	0.38	0.70	-2.60

Table 10 - Mean percentage forecast error for each transition probability method and weight status for females from 1994-2010

Transition Probability Method	Mean Percentage Forecast Error (%)			
	Underweight	*Normal*	*Overweight*	*Obese*
Empirical	-5.34	-0.42	0.96	-1.30
LOWESS	1.92	4.57	-4.02	-3.59
Multinomial	-5.97	0.16	-0.10	-0.60

The MPE table for males reveals that the most evident discrepancy is the underweight prediction error. The highest error is an under prediction (-32.22%) of underweight males when using the LOWESS transition probabilities. In other words, the predicted underweight prevalence across all cycles was on average 32.22% less than the observed underweight prevalence. This large discrepancy is to be expected with the underweight category since sample sizes for deriving underweight transition probabilities are much lower than the other weight statuses. Other than the underweight prediction error, all three transition probability methods for the males predict the observed prevalence with a high degree of accuracy.

For females, the empirical and multinomial transition probabilities predict weight prevalence with the highest accuracy. A relatively larger error is observed for the

LOWESS procedure, which over predicts normal weight by 4.57% on average and under predicts overweight and obesity by 4.02% and 3.59% on average, respectively.

6.3 Forecast Simulation

The forecast simulation is a continuation of the validation simulation. This simulation is run once with an initial population of 500,000. Only one replication is reported because the yearly prevalence of each weight status was found to be practically identical across simulation runs (Appendix E). The following figures combine the validation period simulation results (1994-2010) and the forecasted results from 2010-2030.

LOWESS Method Forecast

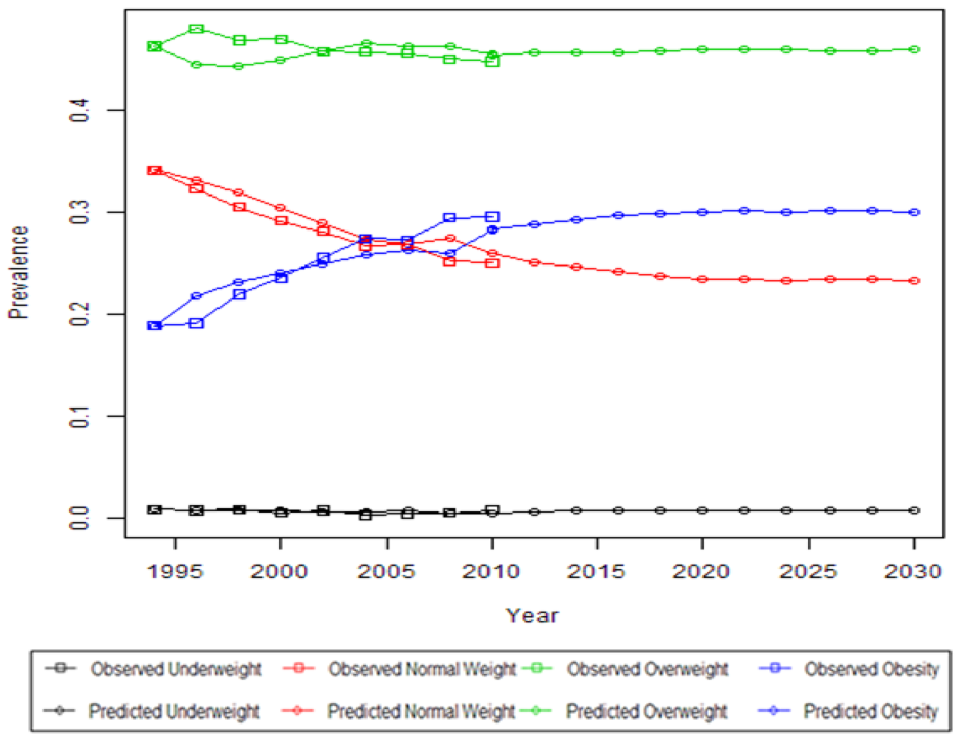


Figure 15 - Validation simulation (1994-2010) and forecasting simulation (2010-2030) weight status prevalence among males using LOWESS transition probabilities

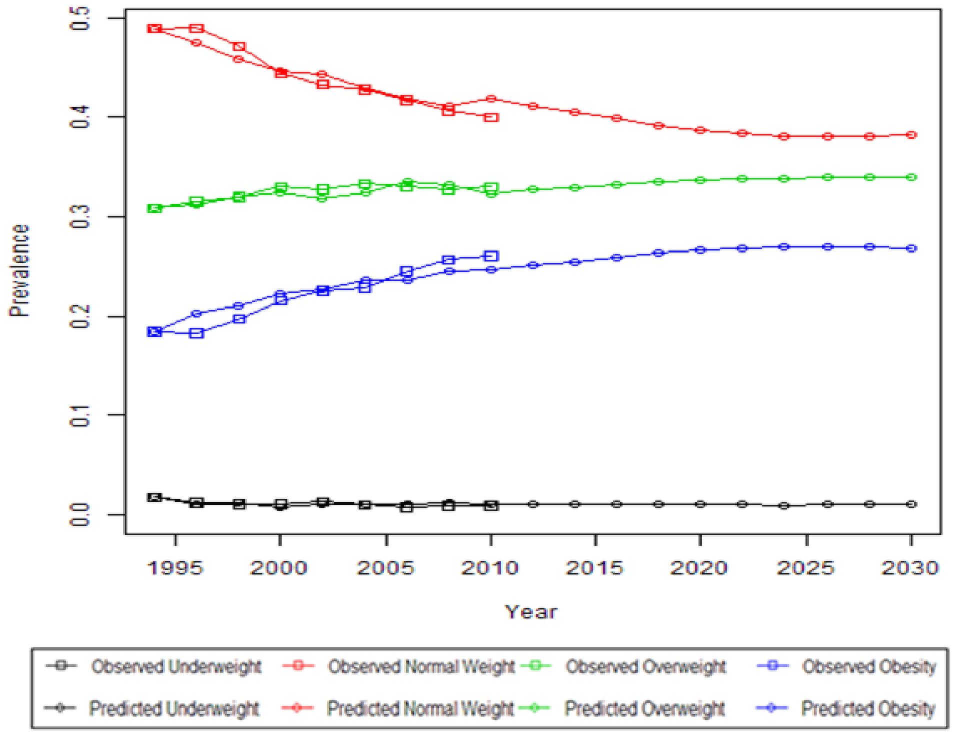


Figure 16 - Validation simulation (1994-2010) and forecasting simulation (2010-2030) weight status prevalence among females using LOWESS transition probabilities

Multinomial Method Forecast

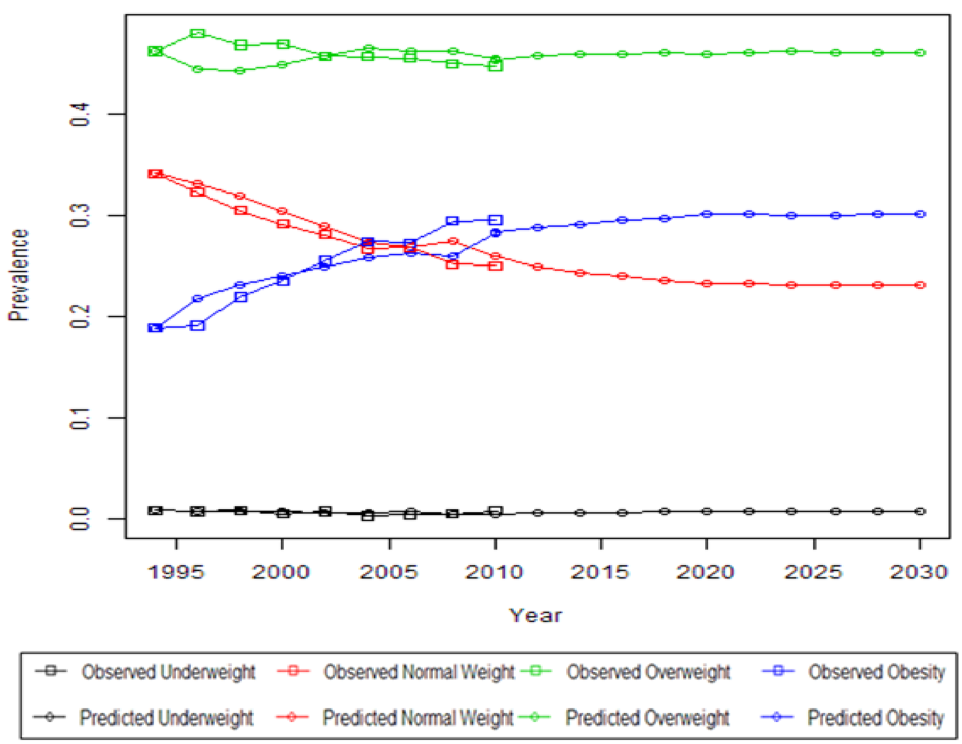


Figure 17 - Validation simulation (1994-2010) and forecasting simulation (2010-2030) weight status prevalence among males using multinomial transition probabilities

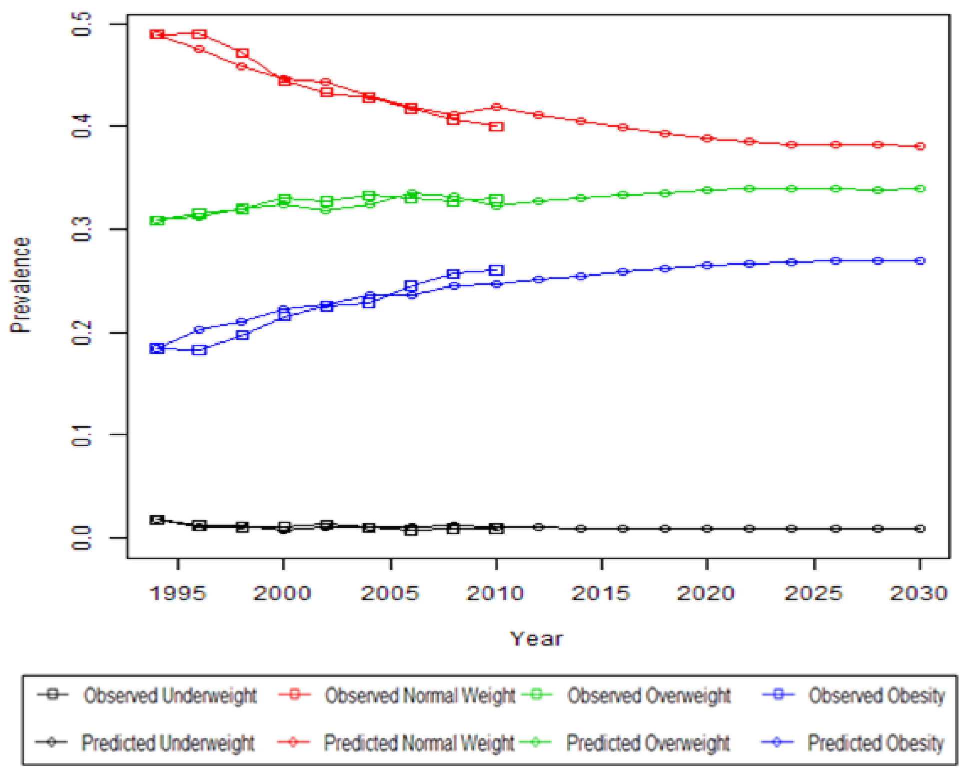


Figure 18 - Validation simulation (1994-2010) and forecasting simulation (2010-2030) weight status prevalence among females using multinomial transition probabilities

Results from Figures 15-18 show that the obesity prevalence for both males and females will continue to increase for a number of years before stabilizing. For males, the prevalence of overweight and obesity appear to stabilize after 2020. For females the prevalence of overweight and obesity continues to converge and appears to stabilize after 2026.

An additional analysis was performed to explore how the plateauing of the prevalence of obesity may be related to the Canadian age demographic. First, the following simulated data illustrates that the prevalence of obesity peaks between approximately ages 40 and 60 in males and between 50 and 70 in females. Note that the following results (Figures 19 and 20) represent only the simulation generated by the multinomial transition probabilities. The simulation results using LOWESS transition probabilities were not noticeably different.

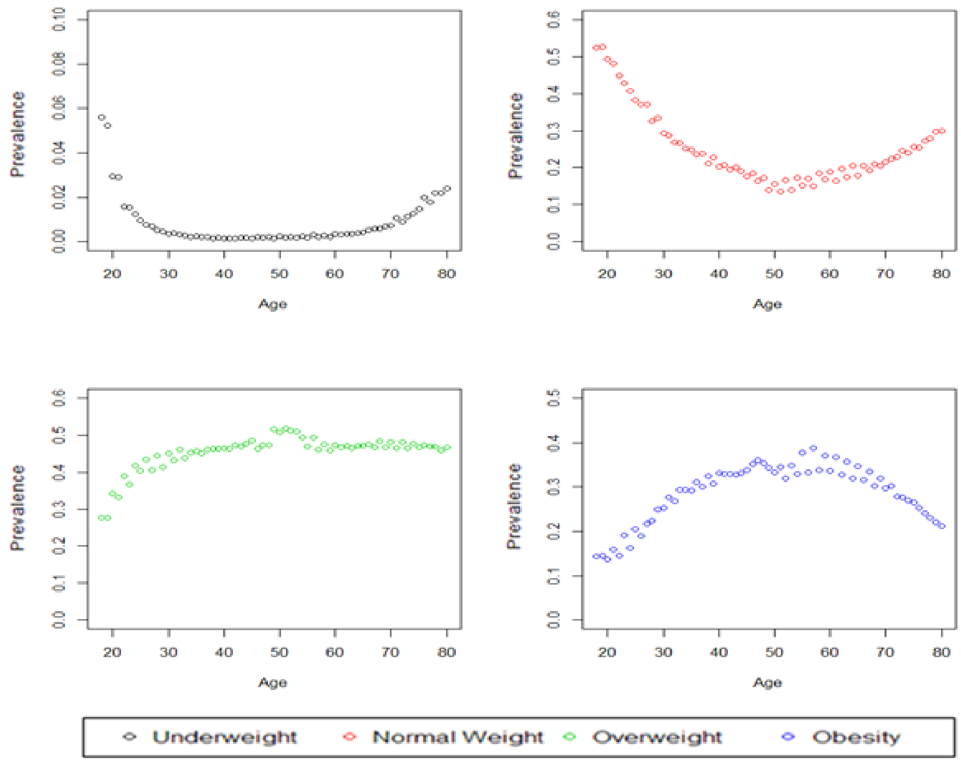


Figure 19 - Predicted weight status prevalence by age among males for the 2010-2030 forecasting period

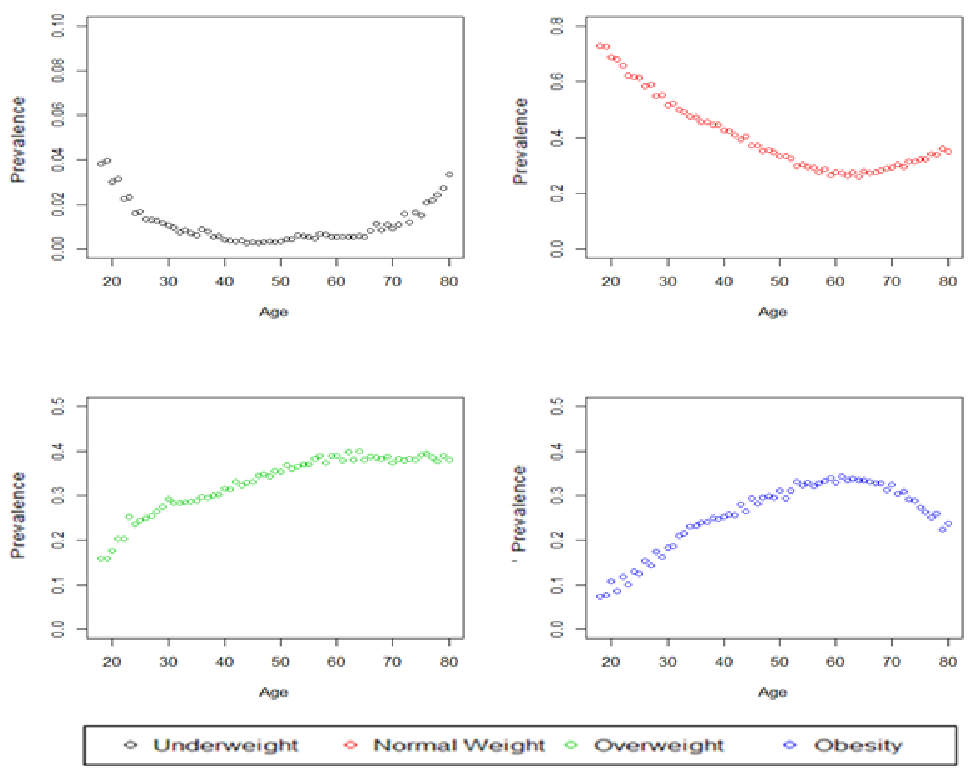


Figure 20 - Predicted weight status prevalence by age among females for the 2010-2030 forecasting period

Subsequently, we can consider the simulated age demographic data from 2010-2030 in Figures 21 and 22.

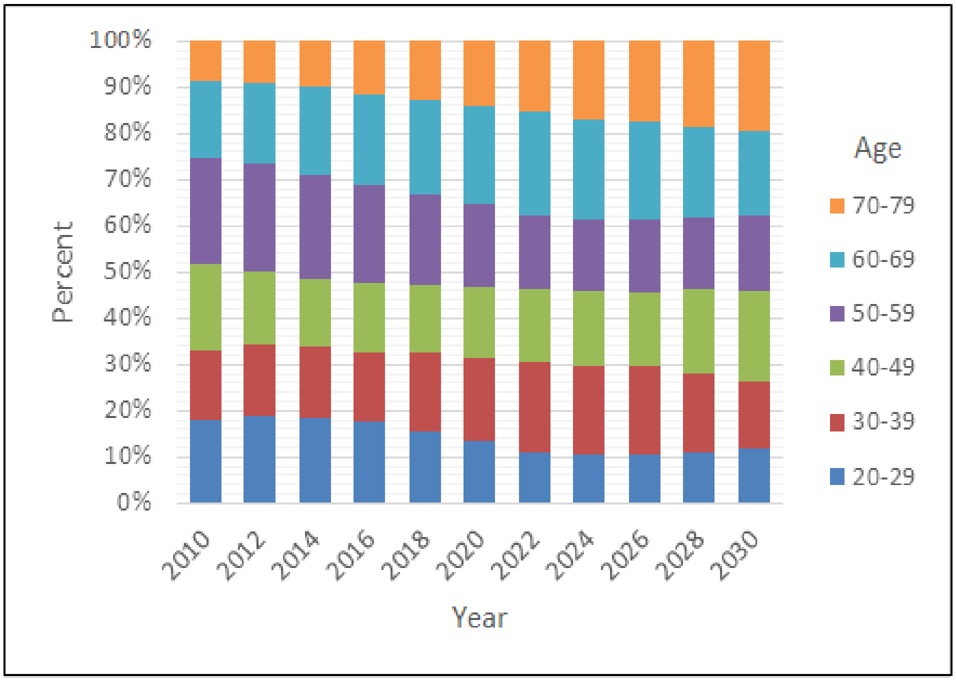


Figure 21 - Predicted age demographic distribution by year among males over the 2010-2030 forecasting period

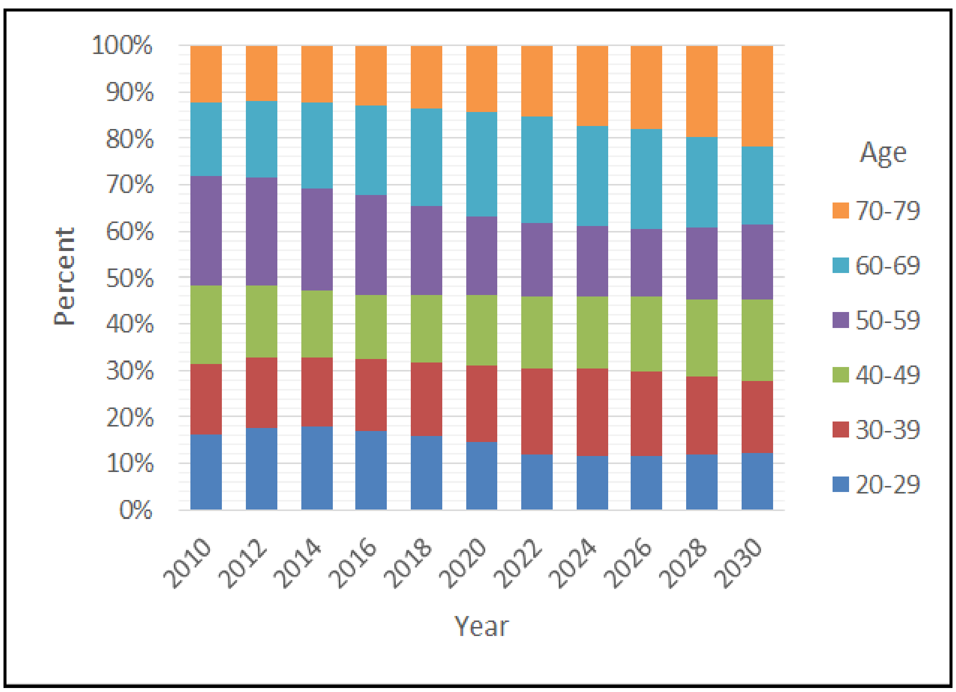


Figure 22 - Predicted age demographic distribution by year among females over the 2010-2030 forecasting period

These figures reveal that the proportions of males and females aged 40-70, who have the highest obesity rates (Figures 19 and 20), decreased approximately 6% from 2010-2030. Furthermore the proportion of 70-80 year olds, who are more likely to transfer to lower weight statuses, is increasing (Figures 19 and 20).

To further investigate the impact of the age demographic age-standardized prevalence for each weight status was calculated. The prevalence from years 2012-2030 was standardized based on the initial 2010 sampled values.

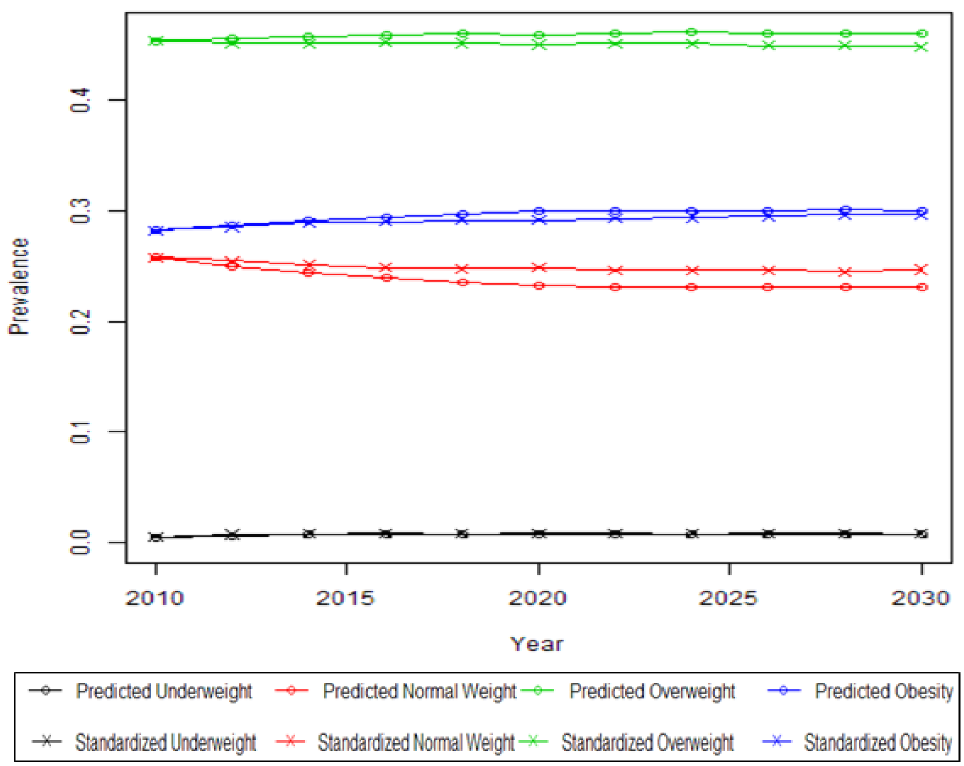


Figure 23 - Predicted versus age-standardized weight status prevalence among males from 2010-2030

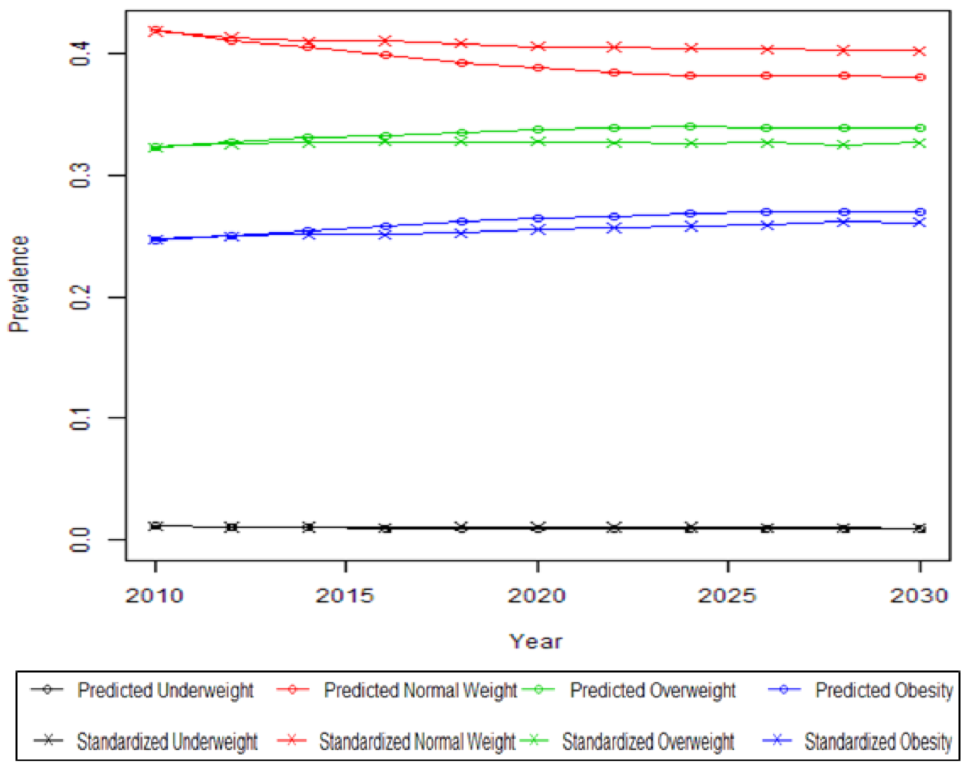


Figure 24 - Predicted versus age-standardized weight status prevalence among females from 2010-2030

Each of the weight status prevalence time series plots from 2010-2030 are more level than their corresponding forecasted prevalence. The largest adjustment between predicted and standardized graphs is that of the prevalence of normal weight, where the absolute difference is approximately 1% for males and 2% for females. Although the age demographic changed considerably across age groups (Figures 21 and 22), the age-standardized results (Figures 23 and 24) ultimately indicate this change does not contribute significantly to the observed stabilization of trends.

6.4 Scenario Results

This section presents the forecasted weight prevalence and the incremental QALYs for each of the proposed scenarios. The forecasted weight status prevalence is reported using one replication. The difference in QALYs between the reference simulation and 30 scenario replications are reported with 95% confidence intervals. Computation times[2] for each of the simulations and their results are summarized in the following table. Processing times for scenarios 1, 2 and 3 also include results generation.

Table 11 - Simulation computation times for each scenario

	Time (Minutes)	Initial Population	Replications
Scenario 1 reference forecast	30	20,000	30
Scenario 1	40	20,000	30
Scenario 2 and 3 reference forecast	450	500,000	30
Scenario 2	550	500,000	30
Scenario 3	550	500,000	30

6.4.1 School-based Intervention Scenario

This scenario models a cohort of 18 and 19 year olds with a reduced chance of entering as obese, and compares the results to a reference cohort simulation which uses the original incoming weight status distribution. The following figures overlay the weight prevalence results of the scenario and the reference forecast.

[2]Using an i-7 3520M processor (4 cores - R uses only one), 8GB RAM

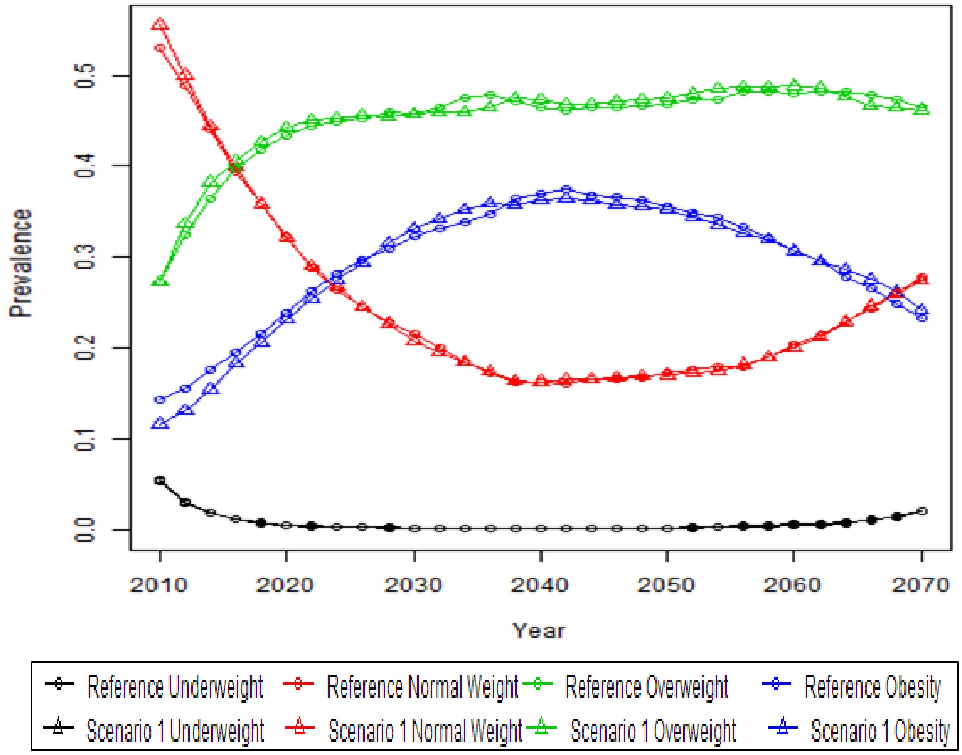


Figure 25 - Reference versus scenario 1 forecasted weight status prevalence among males from 2010-2030

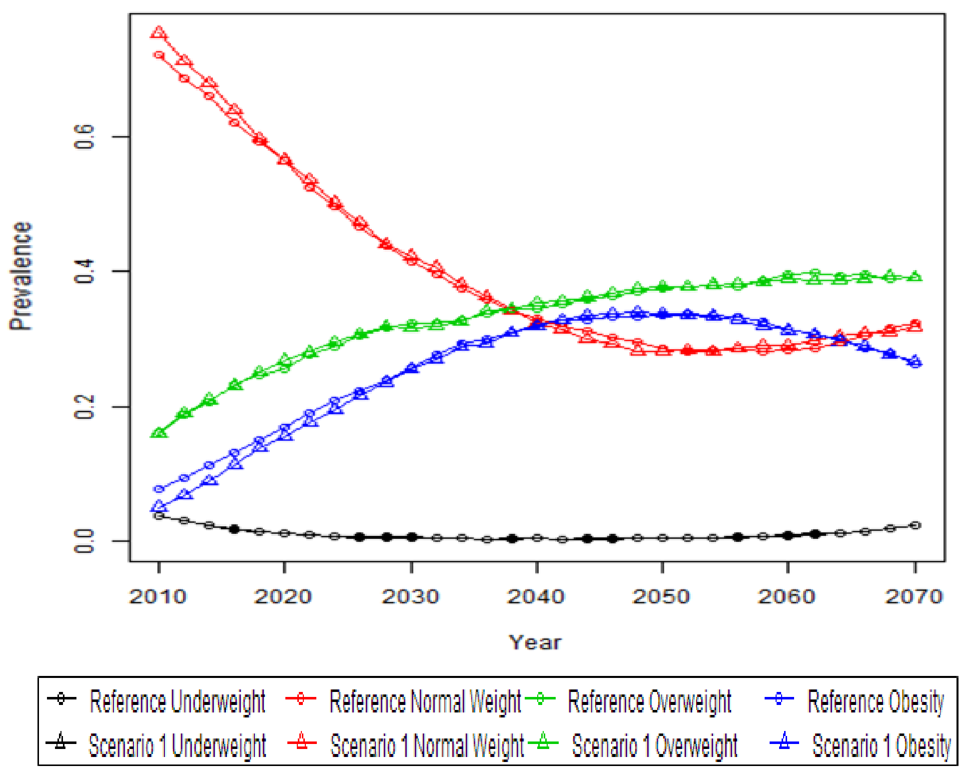


Figure 26 - Reference versus scenario 1 forecasted weight status prevalence among females from 2010-2030

The cohort intervention simulation for both males and females closely resembles the reference scenario. Although obesity prevalence was initially reduced for a number of years, this difference does not appear to endure throughout the simulation. To substantiate how this change in weight status impacts the quality of life for individuals, the QALY metric is used. The QALYs gained and the number of individuals affected by the intervention is presented in Table 12.

Table 12 - Scenario 1 versus reference QALYs simulation results

	Males	Females
Incremental QALYs	**64.98** [60.93, 68.85]	**113.19** [103.13, 123.25]
Intervention Population	9,728	10,272

According to Laupacis [92] a health improvement program is considered good value for money when one QALY can be gained for each $20,000 invested (Section 4.5). Applying this ratio to our case means that this school-based intervention is good value if it can be achieved at $133.40 [125.26, 141.55] per male and $220.39 [200.80, 239.98] per female. Table 13 summarizes this result, and the scenario cost per person which would be considered a moderate (60,000$/QALYs) and poor (100,000$/QALYs) investment.

Table 13 - Scenario 1 intervention cost per person

	Male ($/Person)	Female ($/Person)
Good Value	**133.40** [125.26, 141.55]	**220.39** [200.80, 239.98]
Moderate Value	**400.21** [375.77, 424.64]	**661.17** [602.40, 719.93]
Poor Value	**667.01** [626.29, 707.74]	**1,101.94** [1,004, 1,199.88]

For this scenario the incremental QALYs confidence intervals indicate that females benefit more from this intervention than males. The cost per person for each effectiveness level also reveals that we would be willing to spend significantly more on females than males for the same school-based childhood intervention.

6.4.2 Bariatric Surgery Scenario

The second scenario represents a bariatric surgery program. Each year 10,000 obese Canadians are chosen to participate, of which 70% remain in normal weight status for the length of the simulation. The remaining participant's weight statuses are unaffected by the intervention and continue transition through each cycle normally. As in the previous scenario the following figures overlay the weight status prevalence results from the reference forecast and Scenario 2.

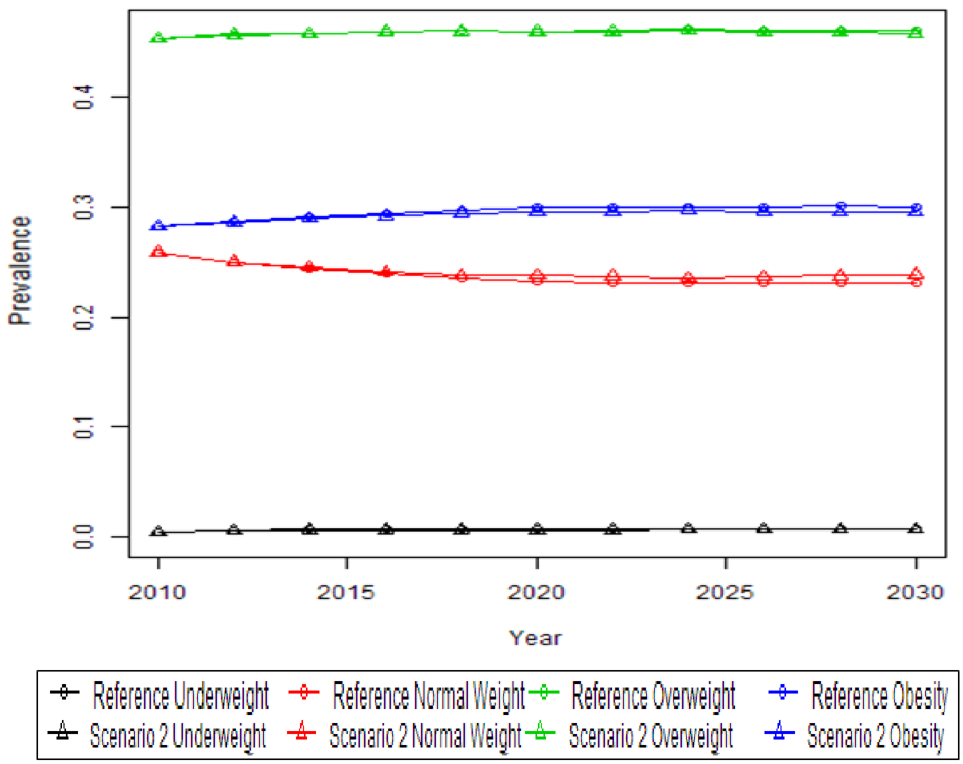


Figure 27 - Reference versus scenario 2 forecasted weight status prevalence among males from 2010-2030

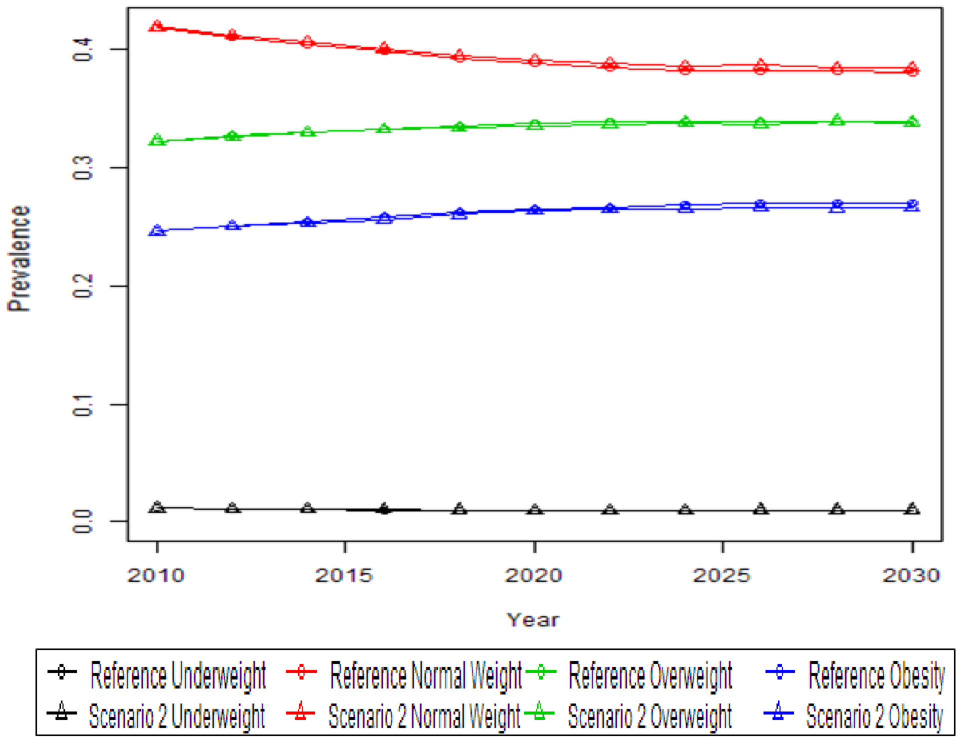


Figure 28 - Reference versus scenario 2 forecasted weight status prevalence among females from 2010-2030

These results indicate 10,000 bariatric surgeries would have a minimal effect on obesity prevalence for both males and females. However the difference in QALYs for this scenario compared to the reference scenario is more pronounced, as shown in the following table.

Table 14 - Scenario 2 versus reference QALYs simulation results

	Male	**Female**
Incremental QALYs	**727.17** [400.77, 1,053.58]	**661.97** [400.34, 923.60]
Intervention Population	1,609	1,332

The acceptable scenario cost per person is calculated for each level of effectiveness:

Table 15 - Scenario 2 intervention cost per person

	Male ($/Person)	Female ($/Person)
Good Value	**9,038.81** [4,981.58, 13,096.03]	**9,939.49** [6,011.06, 13,867.92]
Moderate Value	**27,116.42** [14,944.74, 39,288.10]	**29,818.47** [18,033.17, 41,603.77]
Poor Value	**45,194.04** [24.907.91, 65,480.17]	**49,697.4**5 [30,055.28, 69,339.61]

For this scenario the confidence intervals of cost indicate there is a difference between each level of intervention effectiveness since none of the intervals overlap. However there is no significant difference between males and females in terms of incremental QALYs or the acceptable intervention cost per person.

6.4.3 Primary Population-wide Prevention

The final scenario represents a hypothetical prevention program that affects all Canadians. It reduces the probability of transitioning from normal to overweight and from overweight to obese. As in the two previous scenarios the following figures overlay the weight prevalence results from the reference forecast and Scenario 3.

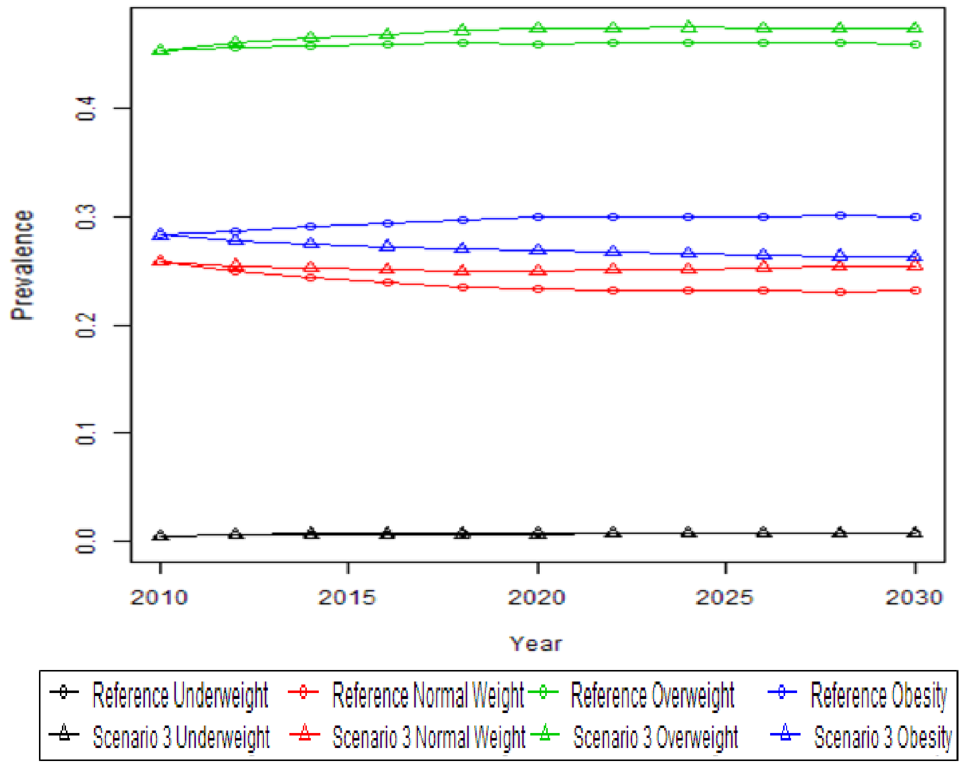


Figure 29 - Reference versus scenario 3 forecasted weight status prevalence among males from 2010-2030

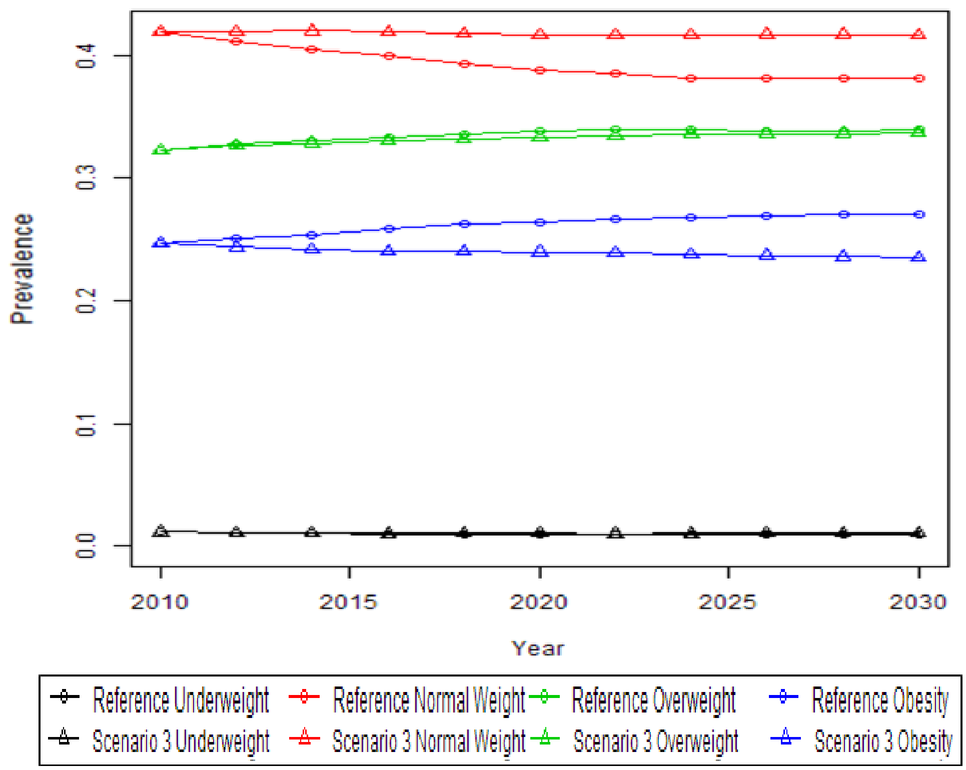


Figure 30 - Reference versus scenario 3 forecasted weight status prevalence among females from 2010-2030

For males the prevalence of obesity immediately reverses its upward trend and remains in decline until 2030. Normal weight continues to decline until 2020 and then returns to its 2010 prevalence level by 2030. The trend in the prevalence of overweight among males does not change, rising at a slightly faster rate than the reference scenario forecast. For females, the trend of obesity is also reversed and declines slightly over the course of the simulation period. Normal weight remains stable until 2030 and the overweight trend follows the same path as the reference scenario. Results of this scenario are summarized in the following tables:

Table 56 - Scenario 3 versus reference QALYs simulation results

	Male	**Female**
Incremental QALYs	**2,676.33** [2,362.41, 2,990.25]	**5,799.84** [5,499.02, 6,100.67]
Intervention Population	300,115	305,824

Table 17 - Scenario 3 intervention cost per person

	Male ($/Person)	**Female ($/Person)**
Good Value	**178.35** [157.43, 199.27]	**379.29** [359.62, 398.97]
Moderate Value	**535.06** [472.3, 597.82]	**1137.88** [1078.86, 1196.90]
Poor Value	**891.77** [787.17, 996.37]	**1896.46** [1798.10, 1994.83]

For this scenario the incremental QALYs confidence intervals indicate that females would benefit more than males from a nation-wide prevention program.

Chapter 7: Discussion

This section is divided into several parts that discuss various aspects of this research. First the forecast and scenario results are further discussed. Next, several limitations of the entire modelling process are reviewed. Subsequently the choice of a Markov behaviour model is examined. Finally, I describe the simulation's capability to run under a range of different input parameters.

Forecasting Simulation

The trajectory of Canadian overweight and obesity prevalence trends and whether large-scale action is necessary is an ongoing debate. A report by Twells *et al.* [4] predicts increases in obesity prevalence for all provinces and advocates for a nation-wide prevention and intervention strategy. In contrast, a report by Esmail and Basham [13] challenges the magnitude of the obesity problem and questions whether government can intervene effectively. A recent report by Sassi and colleagues [97] indicates that the prevalence of obesity in Canada is still increasing, but at a slower rate than previous years. My simulation results similarly indicate that the weight status prevalence distribution will stabilize through time. However, while Esmail and Basham suggest this stabilization is imminent or perhaps has already occurred, the forecasted obesity growth rate in this research does not fully diminish until approximately 2020 for both males and females. Regardless of the precise timing of weight prevalence stabilization, the proposed intervention and prevention programs have the potential to offer gains in quality of life for reasonable costs per person. The calculated intervention costs per person may even be conservative since any cost savings associated with preventing weight-related illness are not considered.

The first scenario follows a cohort of individuals aged 18 and 19 who have been assumed to participate in the APPLE Schools program in Alberta [74]. The weight status prevalence distribution appears to return to the same levels as those of the reference cohort who did not receive the intervention. However, in terms of QALY cost-effectiveness per child, the scenario results indicate that the APPLE school-based program is good value if no more than $133.40 [125.26, 141.55] and $220.39 [200.80, 239.98] is allocated to each male and female, respectively. Compared to an approximate $190 per child for the APPLE Schools program [98], these findings predict a positive return on investment for females if we desire good value from the intervention. For males, spending $190 per child is warranted if moderate intervention value is an acceptable outcome. A limitation of the scenario is that adult transition probabilities of those who received the intervention and those who did not are identical. In reality, a child who attends an APPLE School may have a reduced chance of weight gain throughout adulthood. However, not enough data is yet available to assess the long-term impact on health behaviours of these students.

The second scenario results reveal that performing bariatric surgery on 10,000 individuals per year will have a negligible effect on obesity prevalence. This is expected since the number of surgeries each year is very small relative to the population size. Assuming good value ($20,000/QALY) we would be willing to spend up to $9,038.81 and $9,939.46 per person for males and females, respectively. Although 80% of surgeries patients are currently women [79], the amount we would be willing to spend on female bariatric surgery is not significantly different than males. The cost of one bariatric surgery ranges from $12,000 to $24,000 [79]. According to the $/QALY effectiveness scale (Section 4.5) this intervention can be considered good to moderate value if a positive return on investment is desired. An important limitation of this scenario is that little data exists to support the assumption that successful patients will remain in normal weight status for their entire life. This optimistic estimate of long-term effectiveness may result in an overestimation of the QALYs gained.

The final scenario simulates a hypothetical primary prevention program. Once individuals reach overweight or obese weight status the probability of transitioning to a lower BMI category is small (Figures 1-4). Therefore this hypothetical program focused on reducing the probability of transitioning to overweight or obesity status. The prevention scenario had a considerably larger effect on the weight status prevalence than the previous two scenarios. The results may also be an underestimate as a population-wide prevention program designed to maintain a healthy weight will likely also favourably influence the transition probabilities toward a healthier weight. A limitation of this scenario is that the parameter choice of 2% is arbitrary, and unlike the previous two scenarios no cost benchmark exists to compare whether such a program would be economically sensible. Nonetheless, this simulation reveals that up to 178.35 [157.43, 199.27] per male and 379.29 [359.62, 398.97] per female can be spent to obtain good value from this intervention.

Modelling Limitations

Many elements of this microsimulation model operate under the assumption that current trends will hold into the future. For example, transition probabilities were found to be relatively stable from 1994-2010 period their behaviour past 2010 is uncertain. Likewise, mortality probabilities are based on data from 2009-2011 [87], however changing health behaviour and healthcare advancements are continually influencing Canadian's mortality rates. Similarly the age, gender and weight status specific Health Utility Index [89] is derived from 2000-2005 data and is assumed to remain unchanged through time. However, health utilities are also likely gradually changing as healthcare improvements offer better treatment of obesity-related diseases and conditions.

Another important limitation of this research is the exclusion of the recommended NPHS variance estimation technique throughout the validation and forecasting simulations. Typically simulation model outputs are accompanied by confidence intervals, derived from multiple simulation replications. However, confidence intervals derived from these standard methods are less reliable due to the NPHS's complex survey design [32]. Instead, Statistics Canada recommends the bootstrapping technique for correct variance estimation [99]. Briefly, bootstrapping consists of subsampling the initial 1994 NPHS data multiple times (with replacement) and recalculating the sampling weights for each new subsample. Each of these new bootstrap subsamples would contribute a different set of initial conditions per simulated replication, resulting in more accurate confidence intervals. This method of bootstrapping could also be used to derive probabilistic transition probabilities (as described Section 3.4). Applying bootstrapping in a future project would permit a more accurate internal validation and scenario analyses of the simulated weight status prevalence and QALYs.

A demographic consideration not modelled in this research is the additional population gained from immigration and how weight behaviour varies between the various ethnic groups. However, Goel et al. [100] and other found that the majority of immigrants from non-Western countries to Canada tend to adopt a lifestyle similar to that of the host country, and eventually develop a similar obesity prevalence and mortality risk as Canadian-born individuals.

Due to the use of a utility scale to calculate QALYs, the inherent weaknesses of preference scales should also be noted. Using QALYs to assess cumulative population health relies on two assumptions. The first is that individual utility is additive across time. In regards to health, this assumption may be violated if an individual with an adverse health condition perceives their condition differently through time. For example someone may eventually adapt their perception positively to their condition, while others may perceive the condition's impact on health as worsening through time [90]. The second assumption is that utility is additive across individuals. A person with identical health conditions and characteristics may perceive health entirely differently. With respect to this research, individuals within the same gender, age and weight status group may have differing perceptions of health. However, the risk of misclassifying a person's health utility may be reduced in our case because each group is an average health utility derived from NPHS data. Although these assumptions simplify the complexity of individual health preference, Weinstein suggests QALYs are still a useful means of decision making in health care settings [90].

Markov Model

The structure of this model aggregates the population into age, gender and weight status strata which serve as the explanatory variables for each individual. Faissol *et al.* [47] found that when computing transition probabilities with aggregated data, bias increases with the amount of data aggregation. For example, aggregating male and female data then calculating the transition probabilities would fail to describe any gender differences. Too much disaggregation of the data reduces samples sizes within each stratum. This causes three problems: First small samples sizes fail to truly represent the population, and second, results from a small subsample may not be released from the RDC due to violation of Statistics Canada disclosure guidelines, and third, they may cause the Markov model state space to become too large and intractable (the "curse of dimensionality"). Thus, estimating transition probabilities with the NPHS survey must strike a balance between an aggregated but potentially biased model and a disaggregated model with data derived from smaller sample sizes.

A more disaggregated model might separate obesity into class I, II and III obesity. This would reduce the amount of aggregation bias in the transition probabilities and account for differences of health utility in the higher BMI levels. An example of aggregating the variables is to collapse age into fewer categories. Future work might include programming a more flexible version capable of running the model with differing levels of aggregation and comparing the predictive success and forecast simulation results. Predictive performance may also be improved by introducing more variables into the Markov model such as education or income. This thesis introduces only age and gender for the microsimulation since Statistics Canada provides a reliable distribution of these variables into the future, while projections of other variable's distributions are less certain. Projecting these distributions is the study of more complex microsimulations such as Canada's POHEM [29] or Australia's APPSIM [25].

One could use a continuous-time Markov model which would remove the restriction that individual's data points must be consecutive. In such a model, transitions occur after an exponentially distribution time period. Such time periods could be determined from the NPHS data sets and may not be hampered by data missing from a given cycle. However, it would require that the time between weight status changes be exponentially distributed which may not be the case. This type of model could be built using the Multi-state Model package for R [101], which includes fitting and statistical analysis of a continuous-time model for longitudinal data.

For the LOWESS smoothing of the empirical probabilities, the smoothing parameter as well as the local regression span could be modified to explore if this would yield a closer simulated fit to the NPHS data. Future analyses could also introduce an *ordinal* multinomial regression, in which the naturally ordering of weight status is built into the model. However ordinal multinomial regression requires the proportional odds assumption to be met, which assumes each of weight status coefficients have identical slopes. This would reduce the number of parameters to estimate (and perhaps increases the capacity to input more independent variables), however the proportional odds assumption might not be met for our data.

Simulation Application

Spielauer [22] notes that microsimulation models derived from confidential data are often considered 'black boxes' because they are difficult to validate and interpret independently. This research also must be considered a 'black box' in terms of independent validation since it would require access to the RDC. However interpretation of this microsimulation is possible outside of the RDC since it allows users to change input parameters and initial assumptions. Such changes may include analyses of how using LOWESS versus multinomial model-based probabilities impacts the scenario results. The prevention or intervention scenario parameters can also be modified for investigation into sensitivity analyses or optimization of the simulation metrics. Additionally the user can extend the simulation horizon past 2030 as well as change the discount rate for QALYs. Finally, to quantify each scenario's effect on the elderly the age range can be extended to 104. However, low sample sizes of the elderly population may result in unstable estimates. Analyses of the elderly population would ideally recalculate the transition probabilities using additional data from the health institutions (e.g. nursing homes) component of the NPHS survey.

A future project could also further develop the code to implement multiple scenarios during the same forecast period. For example, children who attend the APPLE Schools program schools currently are assumed to transition through adulthood with the initially derived transition probabilities. A new forecast simulation could consider a combination of the first and third scenario which assumes a lower incoming obesity prevalence and accounts for any prevention benefits of the program that extend into adulthood. General usability of the simulation may also be improved by creating a web-based application for users [102].

Chapter 8: Conclusion

This research used the NPHS to derive the weight status transition behaviour of individuals according to their age, gender and previous weight status. A microsimulation model was created and validated from 1994-2010. Subsequently, the simulation was used to forecast the weight status prevalence from 2010-2030.

Results of the forecast simulation revealed a stabilization of weight status prevalence trends over the next 20 years. However, despite a leveling of overweight and obesity prevalence this research finds weight reduction programs would have an important positive impact on the future health of Canadians. In particular, the school-based (APPLE Schools) scenario simulation revealed that based on QALYs alone, the simulated intervention cost per child is approximately equal to the existing program cost. Incorporating healthcare cost savings into all scenarios would further increase the estimated amount we would be willing to spend on these intervention or prevention programs.

The lack of wide-scale prevention and intervention programs thus far is not entirely surprising, due to the difficulty in assessing their long-term effectiveness and net benefits. As primary research determines the effectiveness of such programs there is also the need to develop simulations in parallel to understand their potential benefits. Microsimulations such as the one presented in this thesis are a particularly promising avenue of research due to their ability to flexibly simulate intervention and prevention scenarios under a range of different assumptions.

Bibliography

[1] W. Luo *et al.*, The burden of adult obesity in Canada. *Chronic Dis. can. 27(4),* pp. 135-144. 2007.

[2] Statistics Canada, "Body composition of Canadian adults, 2009 to 2011," Minister of Industry, Ottawa, Tech. Rep. 82-625-X, 2012.

[3] O. T. Raitakari, M. Juonala and J. S. Viikari, Obesity in childhood and vascular changes in adulthood: Insights into the cardiovascular risk in young Finns study. *Int. J. Obes. (Lond) 29 Suppl 2*pp. S101-4. 2005.

[4] L. K. Twells *et al.*, Current and predicted prevalence of obesity in Canada: A trend analysis. *CMAJ Open 2(1),* pp. E18-26. 2014.

[5] D. Manuel *et al.*, "Projections of preventable risks for cardiovascular disease in Canada to 2021: A microsimulation modelling approach," *Canadian Medical Association Journal Open,* vol. 2, 2014.

[6] C. Le Petit and J. M. Berthelot, "Obesity: A growing issue," Minister of Industry, Ottawa, Tech. Rep. 82-003-XPE, 2006.

[7] H. Jia and E. Lubetkin, "Trends in quality-adjusted life-years lost contributed by smoking and obesity," *American Journal of Preventative Medicine,* vol. 38, pp. 138, 2010.

[8] J. P. Moriarty *et al.*, The effects of incremental costs of smoking and obesity on health care costs among adults: A 7-year longitudinal study. *J. Occup. Environ. Med. 54(3),* pp. 286-291. 2012.

[9] I. Janssen *et al.*, Waist circumference and not body mass index explains obesity-related health risk. *Am. J. Clin. Nutr. 79(3),* pp. 379-384. 2004.

[10] A. H. Anis *et al.*, Obesity and overweight in Canada: An updated cost-of-illness study. *Obes. Rev. 11(1),* pp. 31-40. 2010.

[11] E. A. Finkelstein *et al.*, Obesity and severe obesity forecasts through 2030. *Am. J. Prev. Med. 42(6),* pp. 563-570. 2012.

[12] I. Janssen *et al.*, Influence of overweight and obesity on physician costs in adolescents and adults in Ontario, Canada. *Obes. Rev. 10(1),* pp. 51-57. 2009.

[13] N. Esmail and P. Basham, "Obesity in Canada: Overstated problems, misguided policy solutions," Fraser Institute, 2014.

[14] S. A. Prince, "A population health approach to obesity in Canada – Putting the problem back into context." *Transdisciplinary Studies in Population Health Series,* vol. 1, pp. 22, 2009.

[15] R. Verstraeten *et al.*, Effectiveness of preventive school-based obesity interventions in low- and middle-income countries: A systematic review. *Am. J. Clin. Nutr. 96(2),* pp. 415-438. 2012.

[16] C. Fung *et al.*, From "best practice" to "next practice": The effectiveness of school-based health promotion in improving healthy eating and physical activity and preventing childhood obesity. *Int. J. Behav. Nutr. Phys. Act. 9*pp. 27-5868-9-27. 2012.

[17] Y. Wang *et al.*, Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. *Obesity (Silver Spring) 16(10),* pp. 2323-2330. 2008.

[18] D. T. Levy *et al.*, Simulation models of obesity: A review of the literature and implications for research and policy. *Obes. Rev. 12(5),* pp. 378-394. 2011.

[19] A. Basu, Forecasting distribution of body mass index in the United States: Is there more room for growth? *Med. Decis. Making 30(3),* pp. E1-E11. 2010.

[20] D. M. Thomas *et al.*, Dynamic model predicting overweight, obesity, and extreme obesity prevalence trends. *Obesity (Silver Spring) 22(2),* pp. 590-597. 2014.

[21] G. H. Orcutt, "A new type of socio economic system" *Review of Economics and Statistics,* vol. 58, pp. 773, 1957.

[22] M. Spielauer, "What is dynamic social science microsimulation?" Ministry of Industry, Ottawa, 2011.

[23] L. Mitton, *Microsimulation Modelling for Policy Analysis : Challenges and Innovations.* Cambridge: Cambridge University Press, 2000.

[24] F. A. Sonnenberg and J. R. Beck, Markov models in medical decision making: A practical guide. *Med. Decis. Making 13(4),* pp. 322-338. 1993.

[25] A. Harding *et al.*, "Validating a dynamic population microsimulation model: Recent experience in Australia," *International Journal of Microsimulation,* vol. 3, pp. 46, 2010.

[26] S. Lymer and L. Brown, "Developing a Dynamic Microsimulation Model of the Australian Health System: A Means to Explore Impacts of Obesity over the Next 50 Years" *Epidemiology Research International,* vol. 2012, 2012.

[27] K. McPherson *et al.*, "Tackling obesities: Future Choices. modelling future trends in obesity and the impact on Health," Government Office for Science, London, 2007.

[28] L. Webber *et al.*, "The future burden of obesity-related diseases in the 53 WHO European-Region countries and the impact of effective interventions: a modelling study," *BMJ Open,* vol. 4, 2014.

[29] McGill University Surveillance Lab., "Obesity module of POHEM:IHD,".

[30] J. A. Kopec *et al.*, Development of a population-based microsimulation model of osteoarthritis in Canada. *Osteoarthritis Cartilage 18(3),* pp. 303-311. 2010.

[31] C. Nadeau *et al.*, Development of a population-based microsimulation mode of physical activity in Canada. *Health Rep. 24(10),* pp. 11-19. 2013.

[32] Statistics Canada., "National Population Health Survey: Household Component Cycles 1 to 9 (1994/1995 to 2010/2011) Longitudinal Documentation," 2012.

[33] World Health Organization. BMI classification. 2015. Retrieved January 20, 2013, from http://apps.who.int/bmi/index.jsp?introPage=intro_3.html

[34] P. T. Katzmarzyk and I. Janssen. The economic costs associated with physical inactivity and obesity in Canada: An update. *Can. J. Appl. Physiol. 29(1),* pp. 90-115. 2004.

[35] D. M. Hall and T. J. Cole. What use is the BMI? *Arch. Dis. Child. 91(4),* pp. 283-286. 2006. . DOI: 10.1136/adc.2005.077339.

[36] Canadian Institute for Health Information and Public Health Agency of Canada., Eds., *Obesity in Canada: A Joint Report from the Public Health Agency of Canada and the Canadian Institute for Health Information.* Ottawa: Public Health Agency of Canada, 2011.

[37] M. Shields *et al.*, Bias in self-reported estimates of obesity in Canadian health surveys: An update on correction equations for adults. *Health Rep. 22(3),* pp. 35-45. 2011.

[38] A. J. Hayes *et al.*, Change in bias in self-reported body mass index in Australia between 1995 and 2008 and the evaluation of correction equations. *Popul. Health. Metr 9*pp. 53-7954-9-53. 2011.

[39] S. Connor Gorber *et al.*, A comparison of direct vs. self-report measures for assessing height, weight and body mass index: A systematic review. *Obes. Rev. 8(4),* pp. 307-326. 2007.

[40] P. T. Vanberkel and J. T. Blake, "A Comprehensive Simulation for Wait Time Reduction and Capacity Planning Applied in General Surgery," *Health Care Management Science,* vol. 10, pp. 373, 2007.

[41] M. Lagergren, "What is the role and contribution of models to management and research in the health services? A view from Europe," *European Journal of Operational Research,* vol. 105, pp. 257, 1998.

[42] A. von Ruesten *et al.*, Trend in obesity prevalence in European adult cohort populations during follow-up since 1996 and their predictions to 2015. *PLoS One 6(11),* pp. e27455. 2011.

[43] F. Hillier and G. Lieberman, *Introduction to Operations Research.* McGraw-Hill Higher Education, 2010.

[44] J. D. Kalbfleisch and J. F. Lawless, "The Analysis of Panel Data under a Markov Assumption," *Journal of the American Statistical Association,* vol. 80, pp. 863, 1984.

[45] H. Galler, "Discrete-time and continuous-time approaches to dynamic microsimulation reconsidered," National Centre for Social and Economic Modelling (NATSEM), Tech. Rep. Technical Paper No. 13, 1997.

[46] B. Singer, "Estimation of Nonstationary Markov Chains from Panel Data" *Sociological Methodology,* vol. 12, pp. 319, 1981.

[47] D. M. Faissol, P. M. Griffin and J. L. Swann. Bias in markov models of disease. *Math. Biosci. 220(2),* pp. 143-156. 2009.

[48] R. J. Kryscio and E. L. Abner. Are markov and semi-markov models flexible enough for cognitive panel data? *J. Biom Biostat. 4(1),* pp. 10.4172/2155-6180.1000e122. 2013.

[49] L. Andreassen *et al.*, "The future burden of public pension benefits A Microsimulation study" Statistics Norway, Tech. Rep. Discussion Papers 115, 1994.

[50] P. Diggle et al., *Analysis of Longitudinal Data*. Oxford, NY: Oxford University Press, 1994.

[51] A. Agresti, *An Introduction to Categorical Data Analysis.* New Jersey: John Wiley & Sons, 2007.

[52] F. Yu *et al.*, Use of a markov transition model to analyse longitudinal low-back pain data. *Stat. Methods Med. Res. 12(4),* pp. 321-331. 2003.

[53] J. Jung, "Estimating Markov Transition Probabilities between Health States in the HRS Dataset", 2006.

[54] H. Y. Kang et al., Results of a markov model analysis to assess the cost-effectiveness of statin therapy for the primary prevention of cardiovascular disease in korea: The korean individual-microsimulation model for cardiovascular health interventions. Clin. Ther. 31(12), pp. 2919-30; discussion 2916-8. 2009.

[55] Y. Xie and D. L. Zimmerman, Antedependence models for nonstationary categorical longitudinal data with ignorable missingness: Likelihood-based inference. *Stat. Med. 32(19),* pp. 3274-3289. 2013.

[56] J. Kassteele, *et al.*, Estimating net transition probabilities from cross-sectional data with application to risk factors in chronic disease modeling. *Stat. Med. 31(6),* pp. 533-543. 2012.

[57] F. Sassi *et al.*, "The obesity epidemic: Analysis of past and projected future trends in selected OECD countries," Organisation for Economic Co-operation and Development, France, Tech. Rep. 45, 2009.

[58] A. Alimadad *et al.*, "A novel algorithm for describing population level trends in body weight " *Scientific Research Open Access,* vol. 4, pp. 1514, 2012.

[59] A. Bourisly, "An obesity agent based model: A new decision support system for the obesity epidemic," in *13th International Conference on Systems Simulation,* Singapore, 2013, pp. 37.

[60] M. Spielauer, "Potential of dynamic microsimulation in family studies: A review and some lessons for FAMSIM+," Austrian Institute for Family Studies, Tech. Rep. 18, 2002.

[61] N. A. Christakis and J. H. Fowler. The spread of obesity in a large social network over 32 years. *N. Engl. J. Med. 357(4),* pp. 370-379. 2007.

[62] J. Homer *et al.*, "Obesity population dynamics: Exploring historical growth and plausible futures in the U.S. ," in *24th International System Dynamics Conference,* Nijmegen, 2006.

[63] S. Ross *et al.*, "Complementary Approaches to Modeling Projected Health and Economic Impacts of Obesity in Canada," *Canadian Journal of Diabetes,* vol. 37, pp. S218, 2013.

[64] P. Lau *et al.*, "Projecting the Burden of the Increasing Body Mass Index Trend in Canada Over the Next 25 Years," *Canadian Journal of Diabetes,* vol. 37, pp. S244, 2013.

[65] R. Carter *et al.*, Assessing cost-effectiveness in obesity (ACE-obesity): An overview of the ACE approach, economic methods and cost results. *BMC Public Health 9*pp. 419-2458-9-419. 2009.

[66] A. Okhmatovskaia *et al.*, SimPHO: An ontology for simulation modeling of population health. Presented at Simulation Conference (WSC), Proceedings of the 2012 Winter. 2012.

[67] S. Sutherland and F. Figari, "EUROMOD: the European Union tax-benefit microsimulation model ," *International Journal of Microsimulation,* vol. 6, pp. 4, 2013.

[68] J. Li and C. O'Donoghue, "A survey of dynamic microsimulation models: uses, model structure and methodology," *International Journal of Microsimulation,* vol. 6, pp. 3, 2013.

[69] J. Lightwood *et al.*, Forecasting the future economic burden of current adolescent overweight: An estimate of the coronary heart disease policy model. *Am. J. Public Health 99(12),* pp. 2230-2237. 2009.

[70] C. M. Rutter *et al.*, Dynamic microsimulation models for health outcomes: A review. *Med. Decis. Making 31(1),* pp. 10-18. 2011.

[71] G. Sacks *et al.*, Obesity policy action framework and analysis grids for a comprehensive policy approach to reducing obesity. *Obes. Rev. 10(1),* pp. 76-86. 2009.

[72] Public Health Agency of Canada. (2010). Overview of the pan-Canadian healthy living strategy. Retrieved January 5, 2015, from http://www.phac-aspc.gc.ca/hp-ps/hl-mvs/ipchls-spimmvs-eng.php

[73] R. Geneau *et al.*, "Mobilizing intersectoral action to promote health: The case of ActNowBC in british columbia," Public Health Agency of Canada, Canada, Tech. Rep. HP5-85/2009E-PDF, 2009.

[74] The Alberta Project Promoting active Living and healthy Eating. (n.d.). Retrieved January 5, 2015, from http://www.appleschools.ca/

[75] B. X. Tran *et al*., Life course impact of school-based promotion of healthy eating and active living to prevent childhood obesity. *PLoS One 9(7),* pp. e102242. 2014.

[76] R. Mendelson, "Dietary Interventions for the Treatment of Obesity in Adults," *Canadian Medical Association Journal,* vol. 176, pp. 57, 2007.

[77] E. K. Olander, *et al.*, What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: A systematic review and meta-analysis. *Int. J. Behav. Nutr. Phys. Act. 10*pp. 29-5868-10-29. 2013.

[78] V. Vance *et al.*, "Combined Diet and Exercise Therapy for the Treatment of Obesity in Adults," *Canadian Medical Association Journal,* vol. 176, pp. 60, 2007.

[79] Canadian Institute for Health Information, "Bariatric surgery in canada," Ottawa, 2014.

[80] C. F. Rueda-Clausen *et al.*, New pharmacological approaches for obesity management. *Nat. Rev. Endocrinol. 9(8),* pp. 467-478. 2013.

[81] Venables, W. N., and Smith, D. M. (2012). An introduction to R. Unpublished manuscript. Retrieved March 4, 2013, from http://cran.r-project.org/

[82] J. G. Ibrahim and G. Molenberghs. Missing data methods in longitudinal studies: A review. *Test. (Madr) 18(1),* pp. 1-43. 2009.

[83] Statistics Canada. Questionnaire(s) and reporting guide(s) - national population health survey. 1999. Available: http://www23.statcan.gc.ca/imdb-bmdi/document/3236_D7_T1_V3-eng.pdf.

[84] W. Cleveland and S. Devlin, "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," *Journal of the American Statistical Association,* vol. 83, pp. 596, 1988.

[85] K. Backholer *et al.*, Setting population targets for measuring successful obesity prevention. *Am. J. Public Health 100(11),* pp. 2033-2037. 2010.

[86] World Development Indicators. (2013). Birth rate, crude (per 1,000 people). Retrieved September 22, 2014, from http://data.worldbank.org/indicator/SP.DYN.CBRT.IN

[87] Statistics Canada, "Life tables, Canada, Provinces and territories, 2009 to 2011," Minister of Industry, Ottawa, ON, Tech. Rep. 84-537-X, 2013.

[88] H. M. Orpana *et al.*, BMI and mortality: Results from a national longitudinal study of Canadian adults. *Obesity (Silver Spring) 18(1),* pp. 214-218. 2010.

[89] C. Steensma *et al.*, Comparing life expectancy and health-adjusted life expectancy by body mass index category in adult Canadians: A descriptive study. *Popul. Health. Metr 11(1),* pp. 21-7954-11-21. 2013.

[90] M. C. Weinstein *et al.*, QALYs: The basics. *Value Health. 12 Suppl 1*pp. S5-9. 2009.

[91] National Institute for Health and Clinical Excellence, "Guide to the Methods of Technology Appraisal" 2008.

[92] A. Laupacis *et al.,* How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Cmaj 146(4),* pp. 473-481. 1992.

[93] M.D. Rossetti, *Simulation Modeling and Arena.* Hoboken, NJ: John Wiley, 2010.

[94] S. Karmali *et al.*, Bariatric surgery: A primer. *Can. Fam. Physician 56(9),* pp. 873-879. 2010.

[95] W. S. Richardson *et al.*, Long-term management of patients after weight loss surgery. *Ochsner J. 9(3),* pp. 154-159. 2009.

[96] A. Law and D. Kelton, *Simulation Modeling and Analysis.* McGraw-Hill Science/Engineering/Math, 2000.

[97] F. Sassi (2014). Obesity update. Retrieved December 2, 2014, from http://www.oecd.org/els/health-systems/Obesity-Update-2014.pdf

[98] B. X. Tran *et al.*, "Cost-Effectiveness of School-Based Health Promotion in Canada: A Life-Course Modeling Approach," *The International Society for Pharmacoeconomics and Outcomes Research,* vol. 16, pp. A391, 2013.

[99] D. Yeo *et al.*, "Boostrap Variance Estimation for the National Population Health Survey," 1999.

[100] M. S. Goel et *al.*, Obesity among US immigrant subgroups by duration of residence. *Jama 292(23),* pp. 2860-2867. 2004.

[101] C. Jackson, "Multi-State Models for Panel Data: The msm Package for R," *Journal of Statistical Software,* vol. 38, 2011.

[102] Chang, W. (2014). Web application framework for R. Retrieved June 1, 2014, from http://cran.r-project.org/web/packages/shiny/shiny.pdf

Appendix A

A1 - Health utilities for each weight status stratified by age group and gender

Age	Males				Females			
	Under	*Normal*	*Overweight*	*Obese*	*Under*	*Normal*	*Overweight*	*Obese*
18-24	0.84	0.91	0.92	0.88	0.88	0.92	0.90	0.85
25-29	0.82	0.92	0.92	0.90	0.92	0.92	0.90	0.89
30-34	0.86	0.92	0.93	0.90	0.89	0.91	0.89	0.85
35-39	0.81	0.90	0.91	0.90	0.92	0.91	0.91	0.85
40-44	0.83	0.90	0.91	0.89	0.86	0.90	0.88	0.84
45-49	0.70	0.89	0.90	0.89	0.85	0.89	0.86	0.82
50-54	0.73	0.87	0.89	0.87	0.79	0.88	0.86	0.80
55-59	0.72	0.89	0.87	0.84	0.82	0.88	0.87	0.78
60-64	0.81	0.86	0.87	0.83	0.80	0.88	0.84	0.79
65-69	0.74	0.88	0.86	0.81	0.84	0.87	0.86	0.77
70-74	0.74	0.86	0.85	0.78	0.76	0.84	0.83	0.73
75-79	0.73	0.80	0.79	0.80	0.75	0.79	0.78	0.66
80-84	0.54	0.74	0.73	0.69	0.62	0.74	0.69	0.61
>85	0.52	0.64	0.61	0.48	0.55	0.60	0.64	0.50

Obesity class I and II health utility for each gender and age group are reduced to one obese category with a weighted average of health utility and weight status prevalence for each age category. Additionally, 18 and 19 year olds in our research are assumed to share the same utility values as those in the 20-24 group.

An individual may reside in the same health utility age group over two years, or they may transition from one health utility age group to another during same simulation cycle, *t*. Therefore, QALYs for each individual and each simulation cycle are calculated by:

$$QALYs_{Individual,t} = \begin{Bmatrix} 2 * U_{M,A,G} & if in\ same\ age\ group\ over\ 2\ years \\ U_{M,A,G} + U_{M,A+1,G} & if\ in\ different\ age\ groups \end{Bmatrix} \forall\ individual, t$$

Appendix B

Simulation births of 18 year olds in 2012:

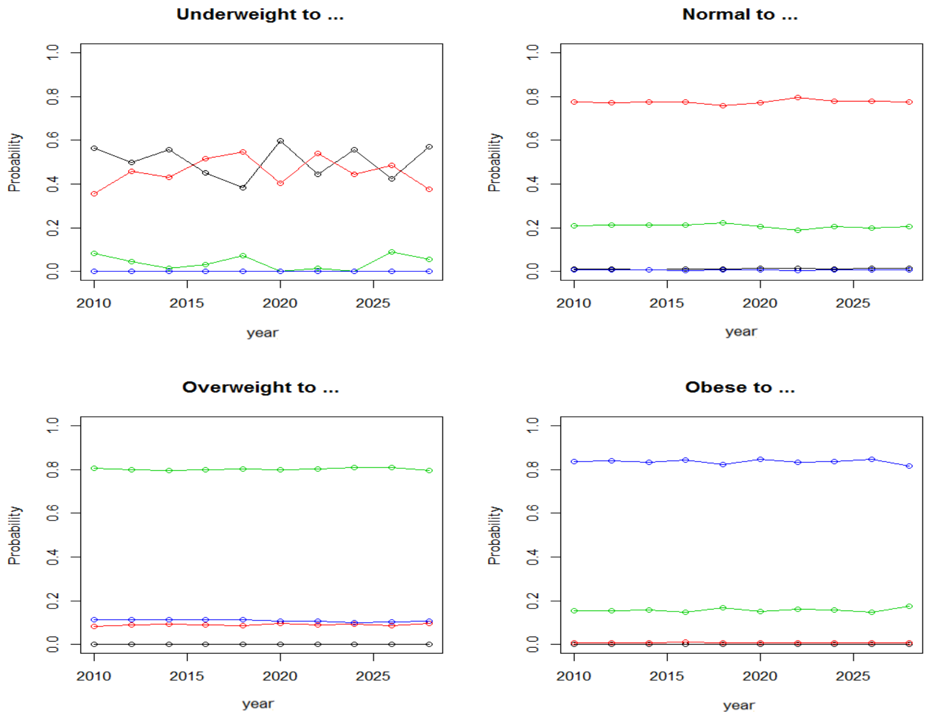
B1 - Expected versus sampled counts of simulated 2012 18 year old births

	Expected Counts	Sampled Counts
Male Underweight	157	166
Male Normal	1523	1509
Male Overweight	807	785
Male Obese	409	407
Female Underweight	120	119
Female Normal	2234	2281
Female Overweight	485	468
Female Obese	227	233

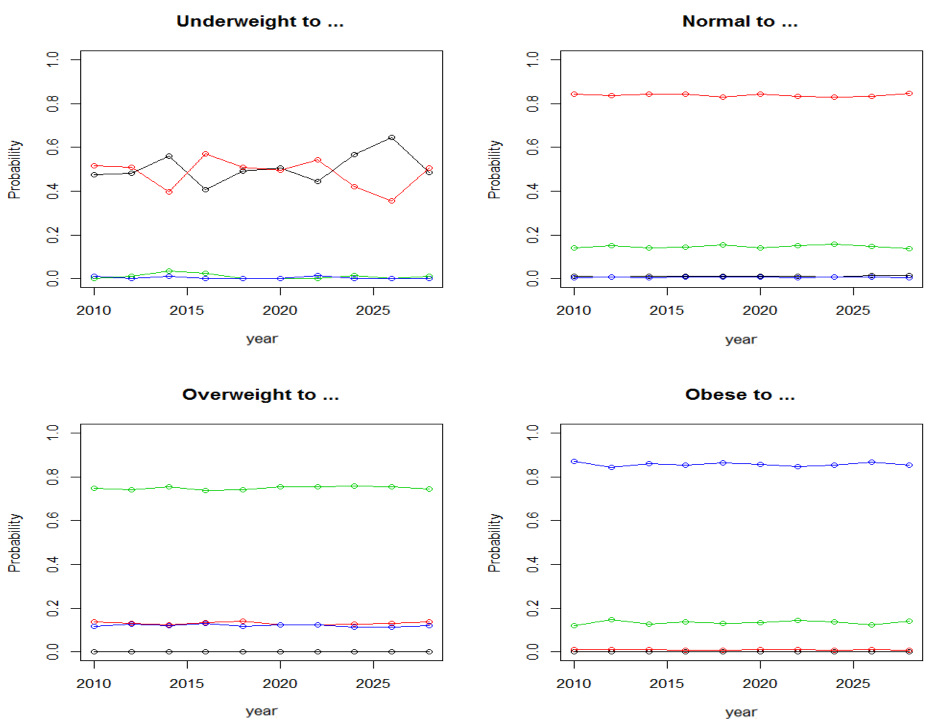
```
  Pearson's Chi-squared test

data:  sampled_counts and expected_counts
X-squared = 56, df = 49, p-value = 0.2289
```

B2 - Pearson's Chi-square test R output

Appendix C

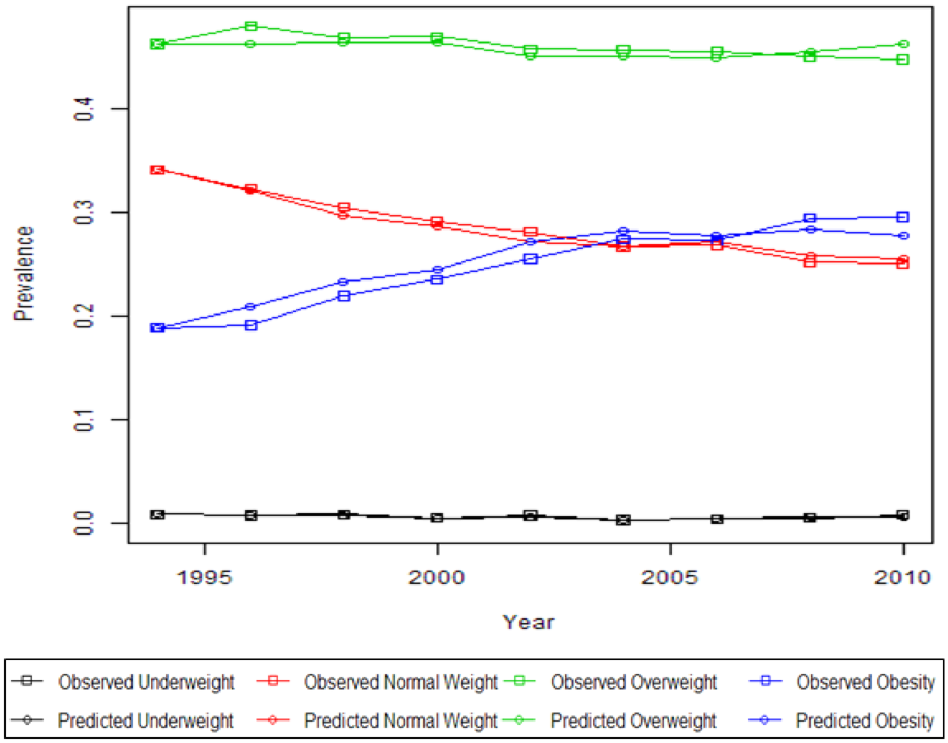


C1 - Transition probabilities by year for males calculated from the 2010-2030 forecast using the multinomial transition matrix

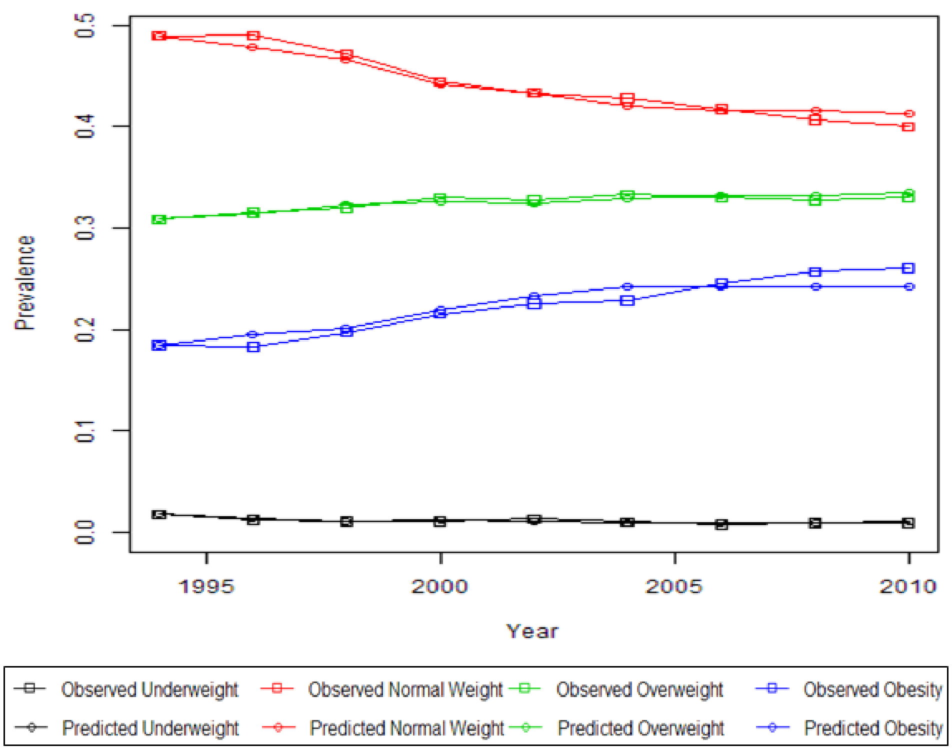


C2 - Transition probabilities by year for females calculated from the 2010-2030 forecast using the multinomial transition matrix

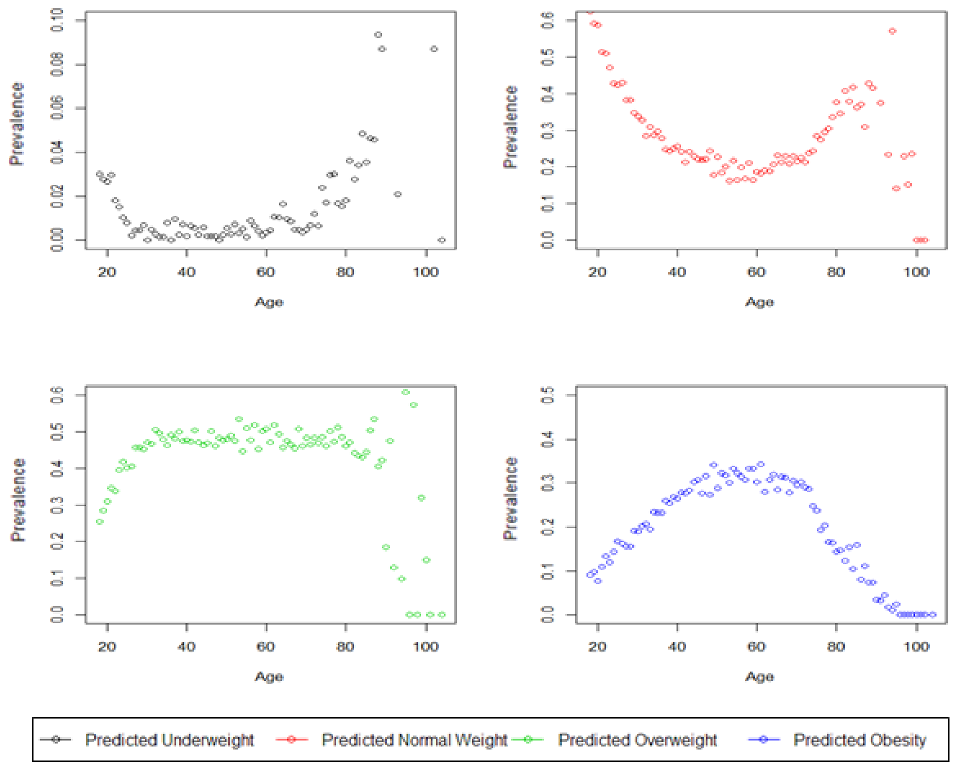
Appendix D



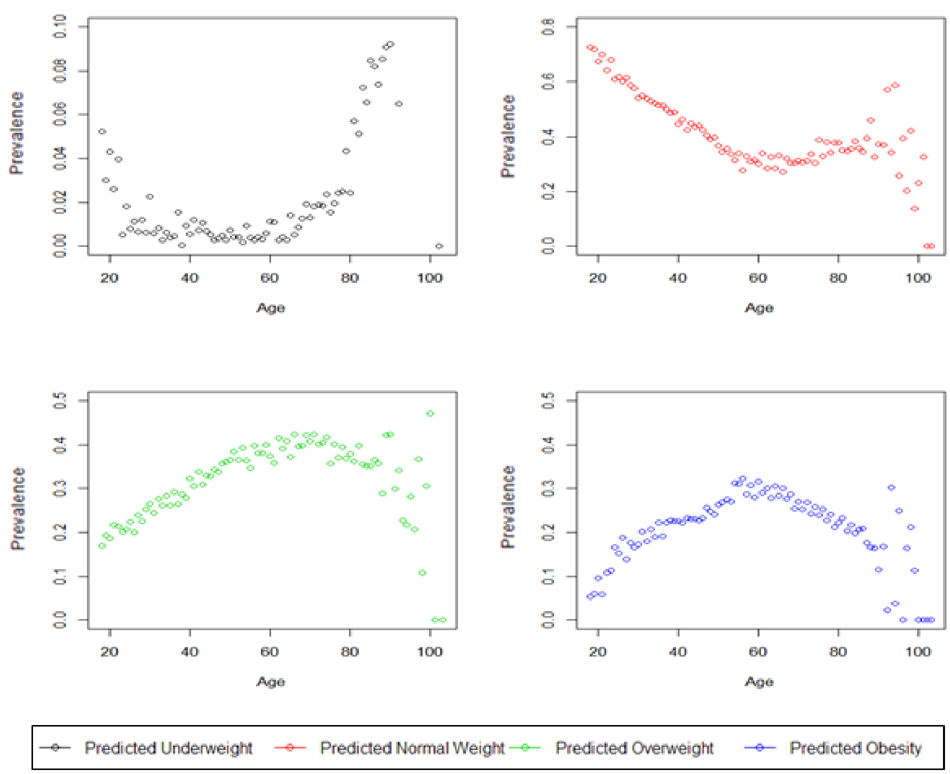
D1 - Predicted versus observed weight status prevalence among males from 1994-2010 using empirically derived transition probabilities and a maximum age of 104



D2 - Predicted versus observed weight status prevalence among females from 1994-2010 using empirically derived transition probabilities and a maximum age of 104



D3 - Simulated weight status prevalence by age among males for the 2010-2030 forecasting period



D4 - Simulated weight status prevalence by age among females for the 2010-2030 forecasting period

Appendix E

E1 - Confidence interval data for weight status prevalence among males from 2012-2030 based on 30 replications

Year	Underweight	Normal Weight	Overweight	Obese
2012	0.0065	0.2506	0.4575	0.2854
	[0.0065,0.0066]	[0.2503, 0.2509]	[0.4572, 0.4578]	[0.2851, 0.2856]
2014	0.007	0.2447	0.4589	0.2895
	[0.0069, 0.007]	[0.2444, 0.2450]	[0.4585, 0.4592]	[0.2891, 0.2898]
2016	0.0068	0.2401	0.4602	0.2928
	[0.0068, 0.0069]	[0.2399, 0.2404]	[0.4598, 0.4606]	[0.2924, 0.2931]
2018	0.0068	0.2365	0.4611	0.2956
	[0.0067, 0.0068]	[0.2362, 0.2368]	[0.4607, 0.4615]	[0.2952, 0.2959]
2020	0.0068	0.2341	0.4615	0.2976
	[0.0067, 0.0068]	[0.2338, 0.2344]	[0.4611, 0.4620]	[0.2972, 0.2980]
2022	0.0069	0.2329	0.4618	0.2984
	[0.0069, 0.0070]	[0.2326, 0.2333]	[0.4612, 0.4623]	[0.2979, 0.2988]
2024	0.0071	0.2320	0.4619	0.2990
	[0.0070, 0.0071]	[0.2317, 0.2324]	[0.4615, 0.4624]	[0.2986, 0.2993]
2026	0.0073	0.2325	0.461	0.2991
	[0.0072, 0.0074]	[0.2322, 0.2329]	[0.4606, 0.4615]	[0.2988, 0.2995]
2028	0.0075	0.2325	0.4603	0.2996
	[0.0074, 0.0076]	[0.2322, 0.2329]	[0.4599, 0.4607]	[0.2992, 0.3000]
2030	0.0077	0.2332	0.4597	0.2994
	[0.0076, 0.0078]	[0.2332, 0.2335]	[0.4594, 0.4601]	[0.2991, 0.2997]

E2 - Confidence interval data for weight status prevalence among females from 2012-2030 based on 30 replications

Year	Underweight	Normal Weight	Overweight	Obese
2012	0.0106	0.4125	0.3261	0.2508
	[0.0106, 0.0107]	[0.4122, 0.4128]	[0.3257, 0.3264]	[0.2506, 0.2509]
2014	0.0101	0.4059	0.3304	0.2536
	[0.0101, 0.0102]	[0.4056, 0.4062]	[0.3300, 0.3308]	[0.2534, 0.2538]
2016	0.0100	0.4000	0.3327	0.2573
	[0.0099, 0.0102]	[0.3997, 0.4004]	[0.3323. 0.3331]	[0.2570, 0.2576]
2018	0.0098	0.3944	0.3349	0.2609
	[0.0097, 0.0099]	[0.3941, 0.3948]	[0.3345, 0.3352]	[0.2606, 0.2612]
2020	0.0096	0.3899	0.3369	0.2636
	[0.0095, 0.0097]	[0.3895, 0.3903]	[0.3365, 0.3373]	[0.2633, 0.2639]
2022	0.0096	0.3862	0.3381	0.2661
	[0.0096, 0.0097]	[0.3858, 0.3866]	[0.3377, 0.3385]	[0.2658, 0.2664]
2024	0.0097	0.3836	0.3391	0.2676
	[0.0096, 0.0097]	[0.3833, 0.3840]	[0.3388, 0.3394]	[0.2673, 0.2680]
2026	0.0097	0.3820	0.3395	0.2687
	[0.0096, 0.0098]	[0.3817, 0.3824]	[0.3392, 0.3399]	[0.2684, 0.2691]
2028	0.0098	0.3815	0.3395	0.2691
	[0.0098, 0.0099]	[0.3811, 0.3819]	[0.3392, 0.3398]	[0.2689, 0.2694]
2030	0.0101	0.3812	0.3396	0.2690
	[0.0100, 0.0102]	[0.3809, 0.3815]	[0.3393, 0.3399]	[0.2687, 0.2694]