Population Health Model - Physical Activity Dynamic Model

*Last updated May 24, 2011*

**Authors**

Claude Nadeau, *Senior Researcher* ¹
William Flanagan, *Chief Microsimulation, Health Analysis* ¹
Jillian Oderkirk, *Director, Health Analysis* ¹
Geoff Rowe, *Senior Researcher* ¹
Victoria Edge, *Sr Epidemiologist* ²
Daniel Gillis, *Researcher* ²
Ron Wall, *Senior Research Analyst* ²
Doug Manuel (Dr), *Senior Scientist* ¹,³,⁵
Mark Tremblay (Dr), *Director, Healthy Active Living and Obesity Research* ⁴,⁵
Jan Trumble Waddell, *Senior Advisor, Chief Public Health Officer* ²

**Affiliations:**

¹Statistics Canada;
²Public Health Agency of Canada;
³Ottawa Hospital Research Institute;
⁴Children’s Hospital of Eastern Ontario;
⁵University of Ottawa
Table of Contents

Abstract: ........................................................................................................................................... 3
Executive Summary.......................................................................................................................... 4
1. Introduction ................................................................................................................................... 5
2. Data ............................................................................................................................................... 6
   2.1. NPHS and CCHS overview .................................................................................................... 6
3. Microsimulation ............................................................................................................................ 10
   3.1. Introduction to microsimulation .......................................................................................... 10
   3.2. Physical Activity simulation ................................................................................................. 10
   3.3. Chronic conditions ............................................................................................................... 11
   3.4. HUI3 (Health Utility Index) ............................................................................................... 12
4. Baseline projections from the simulation ...................................................................................... 12
   4.1. Physical activity projection .................................................................................................. 12
       4.1.1. Leisure time physical activity (LTPA) .......................................................................... 13
       4.1.2. Overall Physical Activity ............................................................................................ 15
       4.1.3. Walking for errands ...................................................................................................... 17
   4.2. Chronic conditions projections ............................................................................................. 19
   4.3. Life Expectancy, HUI and HALE ..................................................................................... 20
5. Impact of increasing physical activity .......................................................................................... 21
   5.1. Life expectancy and Health-adjusted life expectancy .......................................................... 23
   5.2. Prevalence of chronic conditions ....................................................................................... 24
6. Limitations ................................................................................................................................... 26
   6.1. Limits of self-reported data .................................................................................................. 26
   6.2. Data gaps ............................................................................................................................... 28
   6.3. Limitations of simulation ...................................................................................................... 28
7. Conclusion .................................................................................................................................... 29
Acknowledgments ............................................................................................................................ 29
Glossary ............................................................................................................................................ 29
References .......................................................................................................................................... 30
Appendix A: Covariates involved in various models ........................................................................ 31
Appendix B: User guide (pocket version) ....................................................................................... 32
Abstract:

Physical activity is an important determinant of health and how active Canadians will be in the future will have important consequences. Lack of physical activity has been associated with elevated risk of a wide variety of chronic health conditions such as diabetes, hypertension, cardiovascular disease and certain cancers. We used the Population Health Model (POHEM), a dynamic microsimulation framework that projects a representative sample of the Canadian population, to build a decision making tool for the Public Health Agency of Canada in support of new policies and programs to address physical inactivity in Canada. Regression models for physical activity and chronic conditions were estimated from individual level, longitudinal, self-report data from large nationally representative Canadian surveys conducted every two years from 1994 to 2006, and included a wide range of explanatory variables, such as age, sex, income, education, ethnicity, smoking, obesity, and past levels of physical activity. To account for parameter uncertainty, we used bootstrapping techniques and projected confidence bands. We performed what-if scenarios to evaluate the potential impact on health if physical activity trends could be altered. Preliminary results suggest that if all Canadians walked 5 hours per week and engaged in at least one hour of leisure time activity, life expectancy could be increased by 2 years and health-adjusted life expectancy by 3 years, thereby living healthier and longer.
Executive Summary

Statistics Canada and the Public Health Agency of Canada have developed an empirically-based microsimulation tool, designed to evaluate the impact of physical activity on health outcomes in adults. The simulation projects the Canadian population forward 30 years to generate plausible future prevalence of major chronic conditions (heart disease, diabetes, hypertension, and cancer), life expectancy and health-adjusted life expectancy. Four forms of physical activity are modeled from recent societal trends: walking and biking for errands, active leisure time and overall level of activity.

The simulation is empirically grounded: the initial adult population is drawn from the Canadian Community Health Survey (2001); recent societal trends and the relationship between physical activity and health outcomes were modeled from the longitudinal National Population Health Survey (1994-2006). The simulation tool was validated by comparing projections to outcomes from subsequent cycles of the Canadian Community Health Survey.

Self-report survey data is subject to recall error or misinterpretation, and therefore introduces uncertainty. Ideally, measured physical activity would be collected longitudinally on the same Canadians to better understand (i) trends in physical activity, (ii) persistence of physical activity behaviours adopted in childhood into later years, (iii) the relationship between measured physical activity and health outcomes and (iv) the relationship between measured and self-reported physical activity. The Canadian Health Measures Survey will further our understanding since it collects measured and self-reported physical activity in the same individuals, although only at one point in time.

For illustration purposes, we used the simulation to show what might happen to future health outcomes if physical activity could be increased in the population, specifically if everyone has at least one hour of active leisure time per day and walks more than 45 minutes per day (on average). Such a change in physical activity patterns led to

- reduced incidence of heart disease by 20%, diabetes by 17%, hypertension by 13% and cancer by 3%
- life expectancy gains of 2.5 and 2.8 years for females and males, respectively
- improved health-adjusted life expectancy of 3.6 years for both males and females, indicating compression of morbidity in addition to living longer

The simulation tool is called the Population Health Model of Physical Activity (POHEM-PA) and could be used to explore the potential impact of broad societal scenarios of possible futures, from societies that are more community oriented versus cocooning or more technological versus unplugged, to the extent that these behaviours can be associated with differences in levels of physical activity.
1. Introduction

The impact of physical activity on population health has been identified as a priority area by the Public Health Agency of Canada (PHAC). Through its scenario development work, the Agency has identified a series of broad societal drivers bearing upon alternative future patterns of physical activity. PHAC and Statistics Canada (SC) agreed to collaborate to build simulation models of physical activity and its impact on health outcomes.

Statistics Canada has been building microsimulation models for over 20 years, including the mature Population Health Model (POHEM). POHEM is a dynamic microsimulation model of disease and health that projects a representative cross-section of the population (based on the large 2001 Canadian Community Health Survey) into future years. POHEM already includes many components such as an osteoarthritis module and a body mass index (obesity) module.

This project led to an extension (called POHEM-PA) of POHEM. Additions include the change in physical activity over time as well as the relationship between physical activity level and the onset of key diseases, health-adjusted quality of life, and mortality. This tool allows for the evaluation of a series of counterfactual analysis (“what-if scenarios”) that inform the societal scenarios outlined by PHAC. This is done by projecting different levels of physical activity onto a synthetic population representative of the Canadian population and aging it through time to assess the potential impact on health outcomes, such as number of cases of certain diseases and health-adjusted life expectancy.

In addition to the development of the evidence-based simulation tool to study the impact of physical activity on population health outcomes, two other objectives were identified.

I. Identify data gaps in the available physical activity/determinants data holdings, including the relationships between physical activity and health outcomes.

II. Knowledge transfer of microsimulation modeling expertise from SC to PHAC.

This report gives an overview of the tool’s capability and the steps that led to its construction. A forthcoming technical document explains the tool, the statistical models under the hood and provides results, all in a more detailed fashion than done in this report. Together, they constitute an important component in the fulfillment of the knowledge transfer mentioned earlier.
2. Data

The simulation tool requires a large number of equations to propel the initial 2001 population forward. There are two ways to get those equations: (i) borrow from the literature or (ii) build equations from data we have access to (ideally from few comprehensive data bases). In the former case, we would have to borrow from many sources. For instance, relative risk of physical activity for heart disease could come from one study and relative risk of obesity for heart disease would be obtained from another study. Some pros and cons of each approach are listed in Table 2.1.

Table 2.1 Pros and Cons of borrowing versus building equations

<table>
<thead>
<tr>
<th>PROS</th>
<th>Borrow from the literature</th>
<th>Build equations from available data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Requires less work</td>
<td>Consistency of variables</td>
</tr>
<tr>
<td></td>
<td>Benefit from peer-reviewed work of experts</td>
<td>Relevency of equation to the Canadian population if we use Canadian data</td>
</tr>
<tr>
<td>CONS</td>
<td>Would risk factors (e.g. physical activity) be measured the same way in all studies?</td>
<td>Have consistency in the equations form and included covariates</td>
</tr>
<tr>
<td></td>
<td>Would we be able to generate those risk factors in an adequate and coherent manner?</td>
<td>Time consuming</td>
</tr>
<tr>
<td></td>
<td>If based on a foreign study, are equations relevant to the Canadian population?</td>
<td>May not find all the information we could wish for (e.g. measured PA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If many data bases are involved, then consistency of variables may still be an issue</td>
</tr>
</tbody>
</table>

We chose to build the required equations from survey data available at Statistics Canada, namely the National Population Health Survey (NPHS) and the Canadian Community Health Survey (CCHS). The main reason for this decision was data consistency. The two surveys have many pieces of information collected in the same manner (e.g. physical activity). It is then easier to load the relevant information from a cross-sectional snapshot of the population (CCHS cycle 1.1) and use equations built on NPHS and/or CCHS to simulate how individuals in the population will evolve over time.

2.1. NPHS and CCHS overview

The National Population Health Survey (NPHS) follows up a group of respondents that were randomly selected in 1994 to be representative of the Canadian population living in private households at that point in time. It excludes people on Indian reserves, in the territories, on Canadian Forces bases, and in some remote areas. The panel of 17276 initial respondents is followed up every other year. Swain et al., Tambay and Catlin provide further information on NPHS’s design, sample and interview procedures. We have used data from the first 7 cycles (1994 through 2006). This allows us to understand how some characteristics (such as physical activity, disease status, etc…) evolve over time at the individual level.

The Canadian Community Health Survey (CCHS) provides, every other year, a representative cross-sectional sample of the Canadian population. It covers the noninstitutionalized household population aged 12 or older in all provinces and territories, except regular members of the
Canadian Armed Forces and residents of Indian reserves, Canadian Forces bases, and some remote areas. Beland provides more details. The sample size for CCHS (around 130,000 respondents) is substantially larger than that of NPHS.

An important distinction between CCHS and NPHS is that fact that the former is cross-sectional and the latter is longitudinal. Therefore, respondents in CCHS cycle 1.1 differ from those of CCHS cycle 2.1. Hence CCHS is useful to understand population trend (large sample that is always representative) while NPHS is useful for analyses that requires information on the paths of individuals with respect to characteristics of interest such as physical activity or disease status.

A drawback of these two surveys is that they contain self-reported information where measured data may have been preferable (e.g. obesity, physical activity). This caveat however does not overshadow some important qualities of this data, namely (i) its relevancy to the Canadian population, (ii) the size of the samples (iii) the large amount of information collected for each respondent and (iv) the ability to relate physical activity and other factors to relevant health outcomes including chronic diseases, mortality and health-related quality of life. We discuss the issue of self-reported physical activity in the Section 6.1. Data gaps are identified in Section 6.2.

Figure 2.1.1 shows the trend of self-reported physical activity from CCHS and NPHS across measures of physical activity common to both. The four measures of physical activity included in both surveys are as follows:

1. **Leisure-time PA** (four categories): (i) none, (ii) 0-30 minutes/day, (iii) 30-60 minutes/day, (iv) 60+ minutes/day
2. **Walking for errands** (four categories): (i) none, (ii) some but no more than 5 hours/week, (iii) from 6 to 10 hours/week, (iv) more than 10 hours/week
3. **Biking for errands** (two categories): (i) none, (ii) some
4. **Overall activity level** (four categories): (i) Sit, (ii) stand/walk, (iii) lift, (iv) heavy work.

The inclusion of multiple measures of physical activity addresses a criticism of self-reported data when such data are limited only to individuals' leisure time.

As a whole, we see that the trends from NPHS (data used to build the physical activity model described in Section 3.2) agree with the trends from CCHS, which were not used to build the model.

- For overall activity, we see that there are fewer people in the second category (stand/walk) over time. This is partially compensated by a slight uptrend in the third category (lift).
- For walking for errands, we see that the two most active categories are not reported often. We observe that fewer people do not walk for errands over time, leading to an uptrend of people doing some, but no more than 5 hours/week of walking for errands.
- For LTPA, the trend from NPHS tends to be stronger than that from CCHS. NPHS shows a downtrend of people in the two least active categories and, therefore, an increased prevalence of people with more than 30 minutes per day of active leisure time. CCHS does not contradict this but the trends are faint.

---

1 Thinking back over the past 3 months, which of the following best describes your usual daily activities or work habits?

1. Usually sit during the day and don’t walk around very much
2. Stand or walk quite a lot during the day but don’t have to carry or lift things very often
3. Usually lift or carry light loads, or have to climb stairs or hills often
4. Do heavy work or carry very heavy loads
One wishing to assess the plausibility of the above PA trends with measured PA will face a paucity of Canadian studies providing trends. According to the Canadian Fitness & Lifestyle Research Institute, measured PA on Canadian youths does not point to any strong trend: edging higher from 2005-2006 to 2007-2008 and then slipping down in 2008-2009. We are not aware of any counterpart for Canadian adults. The Canadian Health Measures Survey, conducted by Statistics Canada, will provide the first national baseline measure of physical activity captured through the use of accelerometers. It also collects self-reported physical activity and therefore will enable comparisons with accelerometers. First results from this cross-sectional survey are anticipated in the fall of 2010. More information on the comparison between self-reported and measured physical activity can be found in Section 6.
Figure 2.1.1 Physical activity trends from NPHS and CCHS for females and males

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LTPA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minutes/day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. 0-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. 30-60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. 60+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Overall PA**    |        |               |
|                   |        |               |
| 1. Sit            |        |               |
| 2. Stand/walk     |        |               |
| 3. Lift           |        |               |
| 4. Heavy          |        |               |

| **Walk errands**  |        |               |
|                   |        |               |
| Hours/week        |        |               |
| 1. None           |        |               |
| 2. 0-5            |        |               |
| 3. 6-10           |        |               |
| 4. 10+            |        |               |

Legend
Squares and triangles correspond to CCHS and NPHS respectively.
3. Microsimulation

After a general discussion of micro-simulation and POHEM in Section 3.1, we provide an overview of the methods used to simulate physical activity (3.2), chronic conditions (3.3) and health utility index (3.4). More details are provided in the technical document.

3.1. Introduction to microsimulation

The Population Health Model (POHEM) is a sophisticated empirically-grounded, policy-oriented, longitudinal microsimulation model of diseases and risk factors realistically representing the lifecycle dynamics of the Canadian population. POHEM is a discrete-event, continuous time, Monte Carlo microsimulation in which the basic unit of analysis is the individual person. The simulation creates and ages a large sample population representative of Canada, one individual at a time, until death. The life trajectory of individual simulated persons unfolds by exposure to a multitude of life-like events, such as smoking initiation and cessation, changes in body mass, and incidence and progression of diseases such as osteoarthritis, cancer, diabetes and heart disease. POHEM integrates data distributions and equations derived from a wide range of sources, including nationally representative cross-sectional and longitudinal surveys, cancer registries, hospitalization databases, vital statistics, Census, as well as parameters in the published literature.

We have drawn the starting sample population from the Canadian Community Health Survey (CCHS cycle 1.1 conducted in 2000-01, n>100,000, currently restricted to those aged 18 and older). This population is aged, one individual at a time, from 2001 until death in future years. The advantage of loading the starting population from the CCHS is that it provides a coherent set of variables (albeit, self-reported) at the level of individual at a common point in time (2001). These variables include (but are not limited to): sex, age group, province of residence, education level, income quartile, body mass index, smoking, various measures of physical activity, chronic conditions like diabetes, hypertension, heart disease, and cancer, and the Health Utility Index Mark 3 (a measure of health-related quality of life). The survey sample weights are used to weight the outcomes to the Canadian population. Due to missing data that could not be reasonably imputed, a number of records were excluded; the reduced dataset was re-weighted to represent the Canadian population for 2001.

In this project, POHEM was extended to model physical activity and its relationship to the onset of chronic conditions like diabetes, hypertension, heart disease, and cancer, as well as, the impact on health-adjusted life expectancy. The equations governing these dynamics were estimated from the dynamics observed in the NPHS, and include a large number of explanatory variables, including age, sex, and income amongst other. These are shown in Appendix A.

3.2. Physical Activity simulation

CCHS and NPHS provide self-reported information about four different aspects of physical activity: (i) Leisure time physical activity (LTPA), (ii) walking for errands, (iii) biking for errands and (iv) overall level of physical activity. Each aspect of physical activity is a categorical variable. Deciding what will be a person’s LTPA (for instance) in the future boils down to (i) getting the probability of each of the four possible LTPA categories the person may end up into in
the future and (ii) move the person according to those probabilities. Note that the “move” may leave the person in the category currently occupied.

The probabilities are modeled/estimated from NPHS, an individual-level longitudinal data base. We used generalized logit regression\(^2\) to model the probabilities of various categories of physical activity. The probabilities depend on the current and past values of all aspects of physical activity. The probability of future LTPA would depend on the history of LTPA as well as walking and biking for errands and overall level of physical activity. Consequently, the model captures the auto-correlation in LTPA as well as correlation between different aspects of physical activity. The probabilities also depend on other covariates listed in Appendix A. Since NPHS data is used, the estimated probabilities are for two-year transitions, and therefore the physical activity in the simulation is updated every other year for a given subject. Again, the “update” may leave the subject in the same physical activity category. More details are provided in the technical document.

### 3.3. Chronic conditions

Let us now look at the impact of physical activity as well as other risk factors on five outcomes of interest: heart disease, diabetes, hypertension, cancer and mortality. The onset of chronic conditions is modeled through the hazard function (see insert). This effectively allows us to model the timing of chronic conditions in a realistic manner, such that onset of chronic condition may occur at any time during the simulated life, rather than at discrete time intervals such as at the beginning or end of the year.

Early simulation results and survey data suggest that remission should not be ignored; chronic conditions are not always chronic. We therefore have hazard equations that may move a diseased individual into the disease-free state (e.g. go from cancer to no cancer). If such a transition occurs, the individual becomes eligible for the onset of the disease again. The covariates involved in those remission equations are a subset of those present in the onset equations.

The tool has a few controls that allow one to make adjustments to the hazard equations. Those have been used for calibration purposes. We have adjusted the cancer incidence rates so that they would agree with those of an external source (Canadian Cancer Statistics). Likewise, the rates of mortality have been adjusted so that they agree with those provided by Statistics Canada Demography Division.

---

\(^2\) proc logistic with link=glogit in SAS.
3.4. HUI3 (Health Utility Index)

Health utility index (HUI) measures overall health. It is derived from a set of numerous questions. Its maximum value is 1 which corresponds to perfect health and its minimal value is -0.36. A negative value indicates that one’s health is so bad that it is considered as worse than being dead (which has a value of 0). HUI is needed to produce health-adjusted life expectancy. The model updates the HUI based on other characteristics of the person such as age and chronic conditions.

HUI model:

HUI modeling first involved the following linear transformation so that possible values are restrained to the [0,1] interval:

HUI\_01 = (HUI+0.36)/1.36.

Some investigation suggested that the beta distribution would be adequate to model HUI\_01. We then modeled the mean as E[HUI\_01] = exp(β’x)/(1 + exp(β’x)), where β is a vector of coefficients and x contains the covariates listed in Appendix A. This is analogous to the usual logistic regression set up except that the response variable is not binary; it may take any value between 0 and 1.

We use CCHS 1.1 data to estimate the parameters of the model. Since this is a cross-sectional survey, we are connecting one’s HUI to current value of the covariates. NPHS data was then used to figure out how current HUI correlates with previous HUI values. The HUI module in POHEM-PA incorporates the auto-correlation found in this analysis. The reader is referred to the technical document for further details.

4. Baseline projections from the simulation

In this section we show baseline simulation projections for physical activity, chronic conditions including diabetes, hypertension, heart disease and cancer, and health-adjusted life expectancy for selected sex and age group combinations. A full set of graphs for all sex-age group combinations are provided in companion documents (pa_level.doc, pa_level_cohort.doc, walk.doc, walk_cohort.doc, pa_level and pa_level_cohort.doc).

4.1. Physical activity projection

We used the simulation tool to project physical activity from 2001 to 2040. Two different graphical views of the projections are provided: “cross-sectional” and “cohort”. The cross-sectional view is useful to compare how people in their forties in 2010 will compare to people in their forties in 2030, for instance. The “cohort” view starts with people in their forties (for instance) in 2001 and follow them in the future and thus is useful if you want to see what happens to a group of people as they age. In addition, the cross-sectional view also provides the opportunity to validate the simulation projections with data from multiple waves of NPHS and CCHS.
4.1.1. Leisure time physical activity (LTPA)

There are four categories for this variable.

1. none (shown in black in the graphs)
2. 0-30 minutes/day (shown in red in the graphs)
3. 30-60 minutes/day (shown in green in the graphs)
4. 60+ minutes/day (shown in blue in the graphs)

Projections of the proportion of men and women engaged in leisure time physical activity aligned reasonably well with both NPHS and CCHS estimates, across all age groups. For instance, Figure 4.1.1.1 (left panel) shows the projections for females aged 50 compared to survey results.

The least active category (none) is negligible for persons less than 70 years old. In most sex age group combination, the most popular category is “0-30 minutes/day”. The two least active categories (black and red) are projected to go down and the two most active categories (green and blue) are projected to go up. For instance, females in their fifties in the future will be more active than females in their fifties now (see left panel of Figure 4.1.1.1). Overall, the projected trends agree with data from the Canadian Community Health Survey that were not used to develop the model. For older age categories, the projected trends tend to be more pronounced and also tend to persist longer before levelling off (plateau).

Consider a cohort of females in their fifties in 2001 as shown in the right panel of Figure 4.1.1.1. Overall, women in their fifties and early sixties tend to increase their LTPA, but once they reach 75 they reduce it. More specifically, we see that the proportion of women exercising “0-30 minutes/day” goes down until about 2015 (i.e. when they are aged 65-75), a move that is compensated by an increase in the two higher activity categories (30-60 minutes/day and 60+ minutes/day). After this activity peak, the reverse happens, i.e. activity flows from higher activity categories to lower activity categories. The “none” category is fairly negligible for a long time. However, it begins to seriously ramp up after 2030 (when the cohort of females is in their eighties) and the most active category (“60+ minutes/day”) vanishes. The projections for the “30-60 minutes/day” category remains fairly flat for old age; the inflow in this category from the “60+ minutes/day” category possibly balances out with the outflow from “30-60 minutes/day” to “0-30 minutes/day”, say.

In summary, an increase in physical activity is projected for all age groups. However, this does not mean that people from a given cohort will be more active as they age. As seen in the right panel of Figure 4.1.1.1, eventually people reduce activity at older ages.

Additional graphs for each sex and age group combination are found in LTPA.doc and LTPA_cohort.doc, which depict projections according to the “cross-sectional” and “cohort” views, respectively (same legend applies).
Figure 4.1.1.1 Projections of LTPA for females in their fifties

<table>
<thead>
<tr>
<th>Cross-sectional view</th>
<th>Cohort view</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key message:</strong> Women in their fifties in the future will be more active than women in their fifties in the recent past. This increased activity plateaus around 2020.</td>
<td><strong>Key message:</strong> Women in their fifties and early sixties tend to increase their LTPA, but once they reach 75 they reduce it.</td>
</tr>
</tbody>
</table>

Legend:
Squares, triangles and stars correspond to CCHS, NPHS and POHEM-PA respectively. The legend on the graph contains the names of the data source (NPHS, CCHS or POHEM) coupled with a number/color that indicates the level of leisure time physical activity:
1. none (shown in black in the graphs)
2. 0-30 minutes/day (shown in red in the graphs)
3. 30-60 minutes/day (shown in green in the graphs)
4. 60+ minutes/day (shown in blue in the graphs)
4.1.2. **Overall Physical Activity**

People are asked the following question on the survey questionnaire:

“Thinking back over the past 3 months, which of the following best describes your usual daily activities or work habits?

1. Usually sit during the day and don’t walk around very much
2. Stand or walk quite a lot during the day but don’t have to carry or lift things very often
3. Usually lift or carry light loads, or have to climb stairs or hills often
4. Do heavy work or carry very heavy loads”

There are therefore four categories for this variable;

1. Sit (shown in black in the graphs)
2. Stand/walk (shown in red in the graphs)
3. Lift (shown in green in the graphs)
4. Heavy (shown in blue in the graphs)

The most active category (Heavy) is negligible for women and older men. In most sex age group combination, the most popular category is “stand/walk”. For some combination there is a trend for higher activity in the future, in some other combination there is no trend. If we consider males in their forties (left panel of Figure 4.1.2.1) for instance, there will be fewer people in the “stand/walk” in the future compared to now. In counterpart there will be an increase in the “Heavy” category. Trends are however less pronounced than those observed with LTPA. Projected trends typically agree with CCHS. Note that CCHS provides information up to only 2005 because the question has been dropped afterwards.

The right panel of Figure 4.1.2.1 shows the cohort of males in their forties in 2001 (malesex=1 age_group=4). The most active category becomes more negligible as the cohort ages. Prior to 2025, the increase in the “Lift” is likely coming from males leaving the “Heavy” category. The “Lift” category itself experiences an exodus past 2025. The “Stand/walk” category increases modestly and then reaches a plateau. Though not visible in Figure 4.1.2.1, this category is pulled down by the ascent of the “Sit” category for higher ages (past 2040). As observed for LTPA, the least active category seriously ramps up when people get to advanced ages (past 2030 in Figure 4.1.2.1). Before 2025, the decrease in the “Sit” category comes from the fact that relatively few males in their sixties are expected to be in this category in the future. The cross-sectional black curve for males in their sixties (shown in pa_level.doc) is lower than that for males in their forties (left panel of Figure 4.1.2.1).

Additional graphs for each sex and age group combination are found in pa_level.doc and pa_level_cohort.doc, which depict projections according to the “cross-sectional” and “cohort” views, respectively (same legend applies).
Figure 4.1.2 Projections of overall physical activity for males in their forties

**Cross-sectional view**

**Key message:**
Men in their forties in the future will be slightly more active than those of present days.

**Cohort view**

**Key message:**
Men in their forties in 2001 reduce their high intensity activities as they age but remain fairly active. When they reach 70, they reduce their overall activities.

Legend:
Squares, triangles and stars correspond to CCHS, NPHS and POHEM-PA respectively. The legend on the graph contains the names of the data source (NPHS, CCHS or POHEM) coupled with a number/color that indicates the level of overall physical activity:

1. Sit (shown in black in the graphs)
2. Stand/walk (shown in red in the graphs)
3. Lift (shown in green in the graphs)
4. Heavy (shown in blue in the graphs)
4.1.3. Walking for errands

There are four categories for this variable.

1. none (shown in black in the graphs)
2. some walking for errands but no more than 5 hours/week (shown in red in the graphs)
3. walks from 6 to 10 hours/week for errands (shown in green in the graphs)
4. walks more than 10 hours/week for errands (shown in blue in the graphs)

In most age groups, the most popular category is “some walking for errands but no more than 5 hours/week” with the first category (none) often not far behind. The two most active categories are not frequently reported (approximately 10% each). The left panel of Figure 4.1.3.1 says that the fewer males in their sixties will report no walking for errands in the future compared to males in their sixties in 2001. This projected decrease in the least active category is restricted in the first few years prior to 2010. The decrease in this “none” category is compensated by a similar increase in the “some walking for errands but no more than 5 hours/week” category. The most active category is expected to be more infrequent in the future, which is partially compensated by a slight increase in the “walks from 6 to 10 hours/week for errands” category.

The right panel of Figure 4.1.3.1 shows the cohort of males in their sixties in 2001 (malesex=1 age_group=6). For males in their sixties (right panel of Figure 4.1.3.1), the two most active categories become more negligible as people age. The “no more than 5 hours/week” initially goes up to mirror what we see in the left panel (cross-sectional view), reaches a plateau and then starts to decline sharply past 2020 (when the cohort is in their eighties). The least active category shows the opposite pattern and, as seen for previous form of physical activity, becomes very prevalent when people are old (more obvious here because we start with an older cohort, namely males in their sixties in 2001).

Additional graphs for each sex and age group combination are found in walk.doc and walk_cohort.doc, which depict projections according to the “cross-sectional” and “cohort” views, respectively (same legend applies).
Figure 4.1.3.1 Projections of walking for errands for males in their sixties

**Cross-sectional view**

**Key message:**
Men in their sixties in the future will not be drastically different from those of present days.

**Legend:**
Squares, triangles and stars correspond to CCHS, NPHS and POHEM-PA respectively. The legend on the graph contains the names of the data source (NPHS, CCHS or POHEM) coupled with a number/color that indicates the level of walking for errands:

1. None (shown in black in the graphs)
2. Some walking for errands but no more than 5 hours/week (shown in red in the graphs)
3. Walks from 6 to 10 hours/week for errands (shown in green in the graphs)
4. Walks more than 10 hours/week for errands (shown in blue in the graphs)

**Cohort view**

**Key message:**
Men in their sixties in 2001 are more likely to at least do some walking for errands as they age but at the same time shy away from excessive walking (> 5 hr/week). When they reach 80, walking becomes far less prevalent.
4.2. Chronic conditions projections

As mentioned in Section 3.3, we have models for the onset of chronic conditions. These are used to generate projections for the prevalence of those chronic conditions. A summary is provided in Figure 4.2.1 for males and females of all ages (plots by age groups are found in file CC_prevalence.doc). We see that hypertension and diabetes are projected to be markedly more prevalent in the future. The prevalence of hypertension could reach the greatest absolute increase but it is for diabetes that we may see the greatest relative increase: diabetes is projected to more than triple over the time period shown (2001-2039). Heart disease prevalence would also increase: from 5.1% to 7.6% for females and from 6.1% to 10.4% for males. Though not easily perceived in Figure 4.2.1, cancer prevalence would increase as well according to the projections: from 2.0% to 2.2% for females and from 1.8% to 3.2% for males.

Figure 4.2.1 Projections of the prevalence of chronic conditions

<table>
<thead>
<tr>
<th>Females</th>
<th>Males</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="#">Graph showing prevalence of chronic conditions for females and males over time. Prevalence of cancer, diabetes, heart disease and hypertension are shown in black, red, green and blue respectively.</a></td>
<td><a href="#">Graph showing prevalence of chronic conditions for females and males over time. Prevalence of cancer, diabetes, heart disease and hypertension are shown in black, red, green and blue respectively.</a></td>
</tr>
</tbody>
</table>
4.3. Life Expectancy, HUI and HALE

As mentioned in Section 3.4, the Health Utility Index (HUI) measures overall health and takes values between -0.36 and 1; high values correspond to better health. It is particularly useful to compute the health-adjusted life expectancy (HALE). HALE is similar to LE (life expectancy). Each year lived contributes 1 to LE, but that same year’s contribution to HALE is the HUI during that year lived and thus the HALE is incremented by 1 for someone in perfect health (HUI=1) and by an amount less than 1 otherwise.

Figure 4.3.1 shows the projection of HUI by age groups and sex. It shows that older people (especially those in their eighties) have a lower HUI according to both CCHS and the projections. For the first few years, there is an uptrend in the projected HUI which is in agreement with survey for most age groups and sex combinations. The uptrend then plateaus with a slight downturn in younger age groups.

Figure 4.3.1 Projected mean HUI by sex and age groups

<table>
<thead>
<tr>
<th>Females</th>
<th>Males</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Average HUI By Age Group and Sex" /></td>
<td><img src="image" alt="Average HUI By Age Group and Sex" /></td>
</tr>
</tbody>
</table>

Squares and stars correspond to CCHS and POHEM-PA respectively.

Age groups definitions: 3 (30-39), 4 (40-49), 5 (50-59), 6 (60-69), 7 (70-79), 8 (80-89).

According to the simulation tool, the life expectancy of females and males aged 18 in 2001 is approximately 85.6 and 82.7, respectively. The corresponding projected HALEs are 76.5 (females) and 74.9 (males). In the HALE calculation, we have assumed that HUI over the first 18 years of life is about 94% (i.e., the first 18 years of life contribute 17 to HALE). The projected HALE is less than LE as its definition would suggest.

Table 4.3.1 Life expectancy and HALE of an individual aged 18 in 2001

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life Expectancy (LE)</td>
<td>85.6</td>
<td>82.7</td>
</tr>
<tr>
<td>Health-adjusted life expectancy (HALE)</td>
<td>74.5</td>
<td>72.9</td>
</tr>
</tbody>
</table>
5. Impact of increasing physical activity

Besides generating baseline projections for the population, the main purpose of the tool is to consider “what if” scenarios. We consider a simple scenario where (i) everybody is forced to have at least one hour of active leisure time per day and (ii) everybody walks more than 5 hours/week for errands. Therefore, those that would normally not walk for errands or would walk no more than 5 hours/week are promoted to the “6 to 10 hours/week for errands” category. Those in the highest category “walks more than 10 hours/week for errands” are not demoted. Appendix B explains how to set up the simulation in POHEM-PA to implement this scenario.

About 80% of people increase their walking for errands and a similar proportion will need to augment their active leisure time to meet the LTPA threshold imposed. The proportion of individuals already meeting the threshold depends on the age group and calendar time considered. For instance, the proportion of people in their eighties meeting the LTPA threshold would be markedly less than 20%. A scenario that is age-specific could be more appealing but, for illustration purpose, we stick to this simple scenario. The increased PA scenario is phased in gradually over 2002, 2003 and 2004. Once an individual adheres to the increased PA, he/she never reverts back to lower activity. Thus, from 2004 onward, every one is complying with the new PA regime.

Figure 5.1 illustrates the increased PA scenario. It shows the baseline projection of activity for males in their fifties, as well as the projections under the increased PA scenario. We used the cohort view for illustration. For walking for errands, we increased PA in people in the two least active categories to the “6 to 10 hours/week for errands”, so that there is no one walking less than 6 hours/week in 2004 and onward (black and red curves down to 0). The proportion of people who “walk more than 10 hours/week for errands” is about the same under the two scenarios. For LTPA (second row of Figure 5.1), we move everyone to the highest intensity so that past 2003, every one has at least one hour of active leisure time.

We also show the impact of the scenario on the overall PA (as defined in Section 4.1.2). We see that increased walking and LTPA leads to higher overall PA (third row of Figure 5.1): the “lift” category becomes more prevalent and the “sit” category less prevalent. This occurs even though the scenario did not directly call for such an increase. As mentioned in Section 3.2, the modeling captures the “correlation between different aspects of physical activity”. The increase in overall PA thus reflects its correlation to LTPA and walking for errands built into the baseline equations. The user may optionally choose to ignore this correlation, so that the overall PA would behave the same way under the baseline and increased PA scenario; see the technical document for further details.
Figure 5.1 PA for males in their fifties under the baseline and increased PA scenarios

<table>
<thead>
<tr>
<th>Walk errands</th>
<th>Baseline scenario</th>
<th>Increased PA scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. 0-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. 6-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. 10+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LTPA</th>
<th>Baseline scenario</th>
<th>Increased PA scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes/day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. 0-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. 30-60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. 60+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall PA</th>
<th>Baseline scenario</th>
<th>Increased PA scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Sit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Stand/walk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Lift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Heavy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.1. Life expectancy and Health-adjusted life expectancy

We first look at the impact of the increased PA scenario on the life expectancy (LE) and the health-adjusted life expectancy (HALE) of males and females aged 18 in 2001. The increased PA leads to a life expectancy gain of 2.5 and 2.8 years for females and males respectively. The gain in healthy-adjusted life expectancy is even greater reaching 3.6 years for both females and males.

Figure 5.1.1 Impact of increased PA on life expectancy and HALE

Though the increased PA has a benefit on both HALE and LE, it is more pronounced for HALE. This can only happen if the increased PA has a positive impact on HUI, i.e. there is a compression of morbidity, which is a very desirable outcome from a public-health perspective. We verify this in Figure 5.1.2. We see that for each age group, the average HUI under the increased PA (shown as curves with “+”) exceeds that of the baseline scenario (shown as curves with triangles) i.e. curves of “+” lay above curves of triangles.

Figure 5.1.2 Impact of increased PA on average HUI

Triangles and “+” signs refer to the baseline and increased PA scenarios, respectively. Age groups are color coded as suggested in the legend on the graphs, namely: [30,40], [40,50], [50,60], [60,70], [70,80], [80,90].
5.2. Prevalence of chronic conditions

When we increased physical activity over baseline assumptions, we projected a reduction in the prevalence of hypertension, heart disease and diabetes (Figure 5.2.1 right column). In 2040, the reduction was:

- 2.1% (35.4%-33.3%) for hypertension;
- 1.6% (9.1%-7.5%) for heart disease; and
- 2.3% (19.0%-16.7%) for diabetes.

For cancer, prevalence increased by 0.2% (from 2.7% to 2.9%) in 2040.

To understand the increase in cancer prevalence, we need to understand that there are two opposing forces at play when we increase PA. First, as may be seen in the first column of Figure 5.2.1, increased PA typically reduces the incidence rate of chronic conditions and therefore contributes to lower prevalence. On the other hand, increased PA also has a beneficial impact on longevity, which contributes to higher prevalence. This tug of war may be won by either side and is strongly influenced by the average age of onset and death of each cause. In Figure 5.2.1 we see that lower incidence wins over increased longevity for all chronic conditions except for cancer. For cancer, the two opposing forces tend to cancel each other out for females, while increased longevity wins over lowered incidence rate for males. This nicely illustrates the power of POHEM to take into account competing events in a complex web of “causality”, and generate results, that at first blush appear counter-intuitive, but do in fact represent plausible dynamics of a real population.

Figure 5.2.1 Impact of increased PA on the incidence and prevalence of chronic conditions
Triangles and “+” signs correspond to the baseline and increased PA scenarios respectively. Sex is color coded (black for females, red for males).
6. Limitations

6.1. Limits of self-reported data

In this section, we discuss some pros and cons of self-reported and measured physical activity. Overall:

- Measured and self-reported physical activity (PA) differ
- Self-reported PA is useful, and so is measured PA, and
- The use of self-reported PA in this project was motivated by availability of relevant (Canadian) data and data coherence.

Measured and self-reported PA differ

Measures of physical activity (PA) may be classified in two groups: (i) subjective/self-reported and (ii) objective i.e. measured by an apparatus, such as a pedometer or accelerometer. There are differences between those two groups. There are also differences within groups. Indeed, different apparatus will yield different pictures of PA and likewise different questions will lead to different information about PA. No gold standard has emerged yet as acknowledged by Prince et al. and other authors. All tools (objective or subjective) seem to come with some caveats. As a consequence, it is the availability of data that drives the use or non-use of certain form of PA information. The reality is that, for the time being, self-reported PA is much more readily available than measured PA. This is particularly the case for databases that contain other pieces of information such as disease status, the kind of database one needs to correlate PA to variables of interest, such as disease onset, mortality or health-related quality of life. Thus the use of self-reported PA is often inevitable.

In their review of studies where both self-reported and measured PA were used, Prince et al. found that self-reported PA tends to exceed the measured counterpart. As can be seen in their paper, this bias depends on the way the measured PA was obtained (accelerometer, doubly labelled water, heart rate monitor, pedometer, indirect calorimetry), suggesting that objective measures of PA also disagree among themselves.

If different instruments (subjective/questionnaire or objective) lead to different level of physical activity, it could be due to the fact that some instruments are more stringent than others. Those instruments may still be well correlated, i.e. rank people in a similar way. Correlations between self-reported and measured PA gathered by Prince et al. are on average around 40%, with a few negative correlations. This is certainly not high correlation.

An interesting instance of how measured and self-reported PA differ is provided by Ham et al. Based on measured PA in NHANES, they found that Mexican Americans are more active than non-Hispanic Blacks and Whites, contrary to what self-reported PA data from the US National Health Interview Survey suggested. Thus in this case, the two form of PA do not rank groups of people the same way, let alone individuals. The authors conclude that measured PA is better. But we must recognize that, in this particular case, self-reported-PA was restricted to leisure time and did not cover other forms of PA such as active commuting or work. It is important to underline that criticism of self-reported data sometimes revolves around the subjective nature of responses and sometime around the scope of the questions asked (e.g. only leisure-time activities).
Self-reported PA is useful

Differences between measured and self-reported PA generally do not say much about which is bad and which is good (or whether any of those is bad … or good). Whichever way the PA is assessed, its relevancy should be judged by how it correlates with health outcomes. To this point, it may be worthwhile to quote Troiano.

“It is important to recognize that the current recommendation to accumulate 30 min of physical activity on most days (17,31) is based on epidemiological associations between self-reported physical activity and health outcomes. Epidemiological relationships based on objective measures might result in different recommendations for physical activity levels. Less than 30 min/d of physical activity as measured by an accelerometer may provide significant health benefits, because lower levels of objectively measured physical activity correspond to higher levels of self-reported physical activity.”

However, there is currently no data source that allows the likelihood of health outcomes to be estimated by amount of measured physical activity (e.g. accelerometry), and so no precise recommendation can be given on the amount of such objectively measured physical activity that would be needed to reduce risk of adverse health outcomes. By contrast, a relationship between self-report PA and health outcomes has been established (in many studies) and has been used to inform the current recommendations.

Use of self-reported PA in this project

For this project we needed equations to describe how PA connects with (i) the health utility index (to estimate health related quality of life), (ii) the onset of certain chronic conditions and (iii) mortality. Equations describing how PA evolves over time were also required. Literature could have provided those pieces of information but they would have been scattered across different studies (and around the world). Those would have involved PA defined in various ways. For instance, what is meant by inactive in one study may be defined differently in another. PA as defined in NPHS/CCHS (i) uses consistent definitions, (ii) is relevant to the Canadian population and (iii) allows one to correlate PA to prevalence (NPHS and CCHS) and incidence (NPHS) of chronic conditions and health outcomes.

As mentioned earlier, criticisms about self-reported PA often revolve around its limited scope. In order to broaden the scope, we used 4 forms of PA, all present on NPHS and CCHS:

1. Leisure-time PA (four categories): (i) none, (ii) 0-30 minutes/day, (iii) 30-60 minutes/day, (iv) 60+ minutes/day
2. Walking for errands (four categories): (i) none, (ii) some but no more than 5 hours/week, (iii) from 6 to 10 hours/week, (iv) more than 10 hours/week
3. Biking for errands (two categories): (i) none, (ii) some
4. Overall activity level (four categories): (i) Sit, (ii) stand/walk, (iii) lift, (iv) heavy work.

Thinking back over the past 3 months, which of the following best describes your usual daily activities or work habits?

1. Usually sit during the day and don’t walk around very much
2. Stand or walk quite a lot during the day but don’t have to carry or lift things very often
3. Usually lift or carry light loads, or have to climb stairs or hills often
4. Do heavy work or carry very heavy loads
6.2. Data gaps

While the current exercise used the best available evidence in the form of self-reported physical activity and self-reported health outcomes, future work would benefit from other types of data:

a. measured physical activity collected longitudinally on the same Canadian individuals for a sufficiently long period of time (20 or more years), across various ages, in order to better understand trends in physical activity and the relationship between measured PA and health outcomes
b. data to understand potential cohort differences, including any potential impacts on physiology resulting from less physically active and more overweight and obese children today than their historical counterparts, and what if any, long term consequences arise, such as earlier onset of type II diabetes
c. data that describes the persistence of physical activity behaviours adopted in childhood into later years, and any explanatory factors (such as, community, cocooning, technology etc) that play a role in persistence and opportunity for PA
d. data that could provide more insight into sufficient levels of regular physical activity required to maintain good health (eg, impact of low-intensity/high duration versus frequent high-intensity/low duration activities)
e. physical activity data related to more specific health outcomes, such as, specific types of cancer, injuries, and osteoarthritis
f. costing data, both in terms of costs required to implement programs to increase physical activity, and in terms of outcomes
g. data on the demographics and physical activity in children (our model starts at age 18)
h. data linkages between survey and administrative data that help us better quantify the extent to which self-reported chronic conditions are aligned to registry data, or allow us to more directly build equations between physical activity and outcomes
i. data on sedentarism (e.g. TV, computer, reading) which is not always available, especially for NPHS

It may be possible to overcome some of those gaps by introducing some additional flexibility in the tool. For instance, to explore potential physiological differences across cohorts (item b), we could provide a parameter that would be indexed by year the person was born and a value for the assumed cohort effect, like a relative risk that multiply through our equations to increase of risk of disease. The parameter could be varied by experts to see the potential impact of assumptions on cohort differences on outcomes.

6.3. Limitations of simulation

Beyond the data limitations already stated, the simulation has some additional limitations.

a. It is restricted to adults.

b. It does not reflect the dynamic nature of some characteristics such as alcohol consumption, education and income. For instance, as people age, income remains constant in POHEM-PA while in reality people typically have increasing income up to retirement and then decreasing income.

c. POHEM-PA does not produce projections that systematically capture the uncertainty in the input parameters (e.g. coefficients in equations); we rely on sensitivity analyses to generate ranges of plausible outcomes.
7. Conclusion

Microsimulation is a powerful technique for exploring the complex dynamics of physical activity and health outcomes as a population ages. What-if scenarios that change baseline assumptions about physical activity in a population provide insight into the possible future health of that population and can be a useful means to compare different types of interventions against each other.

The scenario considered in this report for illustration purposes consisted of encouraging everyone to (i) have at least one hour of active leisure time per day and (ii) walk more than 5 hours per week for errands. A 100% compliance rate was assumed. In the simulation, increased PA leads to lower incidence of diabetes, hypertension, heart disease and cancer, and lower prevalence in future years for all but cancer due to the complex interaction of competing risks of mortality. The increased PA leads to a life expectancy gain of 2.5 and 2.8 years for females and males respectively. The gain in HALE is 3.6 for both females and males.

As with any analysis, the strength of results and conclusions are only as strong as the inputs used to build the model. Self-report physical activity as currently captured on surveys have well-documented biases from their measured counterparts. Nevertheless, a consistent and plausible set of equations has been developed in this exercise using self-report data. These assumptions should be challenged with a broader set of sensitivity analyses that, for instance, make different assumptions about baseline projections. For instance, what-if overall physical activity was not truly going up over the period captured by the survey data, but was declining, what impact would we then expect to see in the Canadian population in future years.

Acknowledgments

I thank Jillian Oderkirk, Geoff Rowe, William Flanagan, Didier Garriguet, Doug Manual, Mark Tremblay, Jan Trumble Waddell, Victoria Edge and Ron Wall for their contribution to this project via discussions and/or redactional help.

Glossary

Acronyms
- **CCHS**: Canadian Community Health Survey
- **HALE**: Health Adjusted Life Expectancy
- **HUI, HUI3**: Health Utility Index
- **LE**: Life Expectancy
- **LTPA**: Leisure Time Physical Activity
- **NHANES**: National Health And Nutrition Examination Survey (an American survey)
- **NPHS**: National Population Health Survey
- **PA**: Physical Activity
- **PHAC**: Public Health Agency of Canada
- **POHEM**: POpulation HEalth Model
- **POHEM-PA**: POpulation HEalth Model (Physical Activity version)
- **SC**: Statistics Canada
Statistical terminology

- **Relative Risk:** ??

**References**

### Appendix A: Covariates involved in various models

Table A.1 provides information about covariates involved in the modeling of chronic conditions, physical activity and HUI.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Chronic Conditions</th>
<th>Physical Activity</th>
<th>HUI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time varying?</td>
<td>Blood Pressure</td>
<td>Diabetes</td>
</tr>
<tr>
<td>Calendar time</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Region (Atlantic, Quebec, Ontario, Prairies, BC)</td>
<td>No</td>
<td>√</td>
<td>NA</td>
</tr>
<tr>
<td>Sex</td>
<td>No</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>age (five year groups)</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>age (ten year groups)</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>age (linear term past 85 years)</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>No</td>
<td>√</td>
<td>NA</td>
</tr>
<tr>
<td>Race (non-white)</td>
<td>No</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Smoke (never, former, current)</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Drink (never, former, light, moderate, heavy)</td>
<td>No</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Education</td>
<td>No</td>
<td>√</td>
<td>NA</td>
</tr>
<tr>
<td>Income quartile (1\textsuperscript{st}, 2\textsuperscript{nd}, 3\textsuperscript{rd}, 4\textsuperscript{th})</td>
<td>No</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>BMI (&lt;18.5, 18.5-25, 25-30, 30-35, &gt;35)</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>BMI (&lt;18.5, 18.5-25, 25-30, &gt;30)</td>
<td>Yes</td>
<td>√</td>
<td>NA</td>
</tr>
<tr>
<td>Arthritis</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Yes</td>
<td>NA</td>
<td>√</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Yes</td>
<td>√</td>
<td>NA</td>
</tr>
<tr>
<td>Heart disease</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Cancer</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>CC_COUNT_EX4_CAP3_T</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>CC_COUNT_EX5_CAP3_T</td>
<td>Yes</td>
<td>√</td>
<td>NA</td>
</tr>
<tr>
<td>Overall physical activity</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Walk for errands</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Bike for errands</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>LTPA</td>
<td>Yes</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

For chronic conditions, √√ means that the variable was used for both the onset and the remission equations, √ means that the variable was used for the onset equation only.
Appendix B: User guide (pocket version)

Prior to running the simulation, one may need to go in the parameter groups to tell the POHEM-PA what needs to be done. For instance, to set up the increased PA scenario presented in Section 5, we need to go in the parameter group pgPA_All and then into pgPA_CONTROL as illustrated in Figure B.1. Then, double-clicking on LTPA_4CAT_transition allows one to edit the transition matrix for LTPA. In this particular case, we want 100% of those with no LTPA to have 60+ minutes/day and therefore the first line of LTPA_4CAT_transition should be 0 0 0 1. The second line of LTPA_4CAT_transition says that we want 100% of those in the “0-30 minutes/day” category to go in the “60+ minutes/day”. The same logic applies for the third line. Other transition matrices (shown in Figure B.1) may be edited in the same manner. One must not forget to tick the box for the parameter “PA_intervention”; otherwise, the transition matrices will not be used.

Once the simulation is completed, you may look at the results. Simply double-click “Table Groups”, locate the table you are interested in and double-click it. Here are the main tables of possible interest:

- tPA contains the distribution (i.e. proportion of people in each category) for each of the four forms of physical activity. This is provided by age groups, sex and calendar time.
- tPA_by_age_start is the “cohort view” counterpart of tPA.
- tCC contains the prevalence of the four chronic conditions (heart disease, hypertension, diabetes and cancer). This is provided by age groups, sex and calendar time.
- tCC_by_age_start is the “cohort view” counterpart of tCC.
- tHUI contains the average HUI by age groups, sex and calendar time.
- tHUI_by_age_start is the “cohort view” counterpart of tHUI.
- tLife_Expectancy contains the life expectancy and HALE. These provided by sex and initial age (in 2001).

As may be seen in Figure B.2, the physical activity tables are under the table group tgPA_tables while the others tables listed above are (currently) found under “other tables”.

Tables may be exported to Excel by following the next five easy steps.
1. Click on Scenario on the top left corner of the simulation tool
2. Select “Export…”
3. Select “MS Excel” when prompted for the output format
4. Click OK in the “Output contents” pop-up window (default settings should be OK; however, Pivot table format is useful for working with pivot tables in Excel).
5. Select the name and location of the file and click on “Save”

The resulting Excel file may be used to create tables and/or graphs within Excel. It can also be exported to some other software (e.g. SAS).
Figure B.1 Setting up the increased PA scenario in POHEM-PA
Figure B.2 Some output tables from the simulation