*Osteoarthritis and Cartilage* (2010) **18,** 303–311

doi:10.1016/j.joca.2009.10.010

# Osteoarthritis and Cartilage

## Development of a population-based microsimulation model of osteoarthritis in Canada

J. A. Kopec†‡*, E. C. Sayre†‡, W. M. Flanagan§, P. Fines§, J. Cibere†‡, Md. M. Rahman‡, N. J. Bansback‖, A. H. Anis†‡‖, J. M. Jordan¶, B. Sobolev†, J. Aghajanian‡, W. Kang‡, N. V. Greidanus†‡, D. S. Garbuz†‡, G. A. Hawker# and E. M. Badley#

† *University of British Columbia, Vancouver, BC, Canada*
‡ *Arthritis Research Centre of Canada, Vancouver, BC, Canada*
§ *Statistics Canada, Ottawa, ON, Canada*
‖ *Centre for Health Economics and Outcome Sciences, St. Paul's Hospital, Vancouver, BC, Canada*
¶ *University of North Carolina, Chapel Hill, NC, USA*
# *University of Toronto, Toronto, ON, Canada*

### Summary

*Objectives*: The purpose of the study was to develop a population-based simulation model of osteoarthritis (OA) in Canada that can be used to quantify the future health and economic burden of OA under a range of scenarios for changes in the OA risk factors and treatments. In this article we describe the overall structure of the model, sources of data, derivation of key input parameters for the epidemiological component of the model, and preliminary validation studies.

*Design*: We used the Population Health Model (POHEM) platform to develop a stochastic continuous-time microsimulation model of physician-diagnosed OA. Incidence rates were calibrated to agree with administrative data for the province of British Columbia, Canada. The effect of obesity on OA incidence and the impact of OA on health-related quality of life (HRQL) were modeled using Canadian national surveys.

*Results*: Incidence rates of OA in the model increase approximately linearly with age in both sexes between the ages of 50 and 80 and plateau in the very old. In those aged 50+, the rates are substantially higher in women. At baseline, the prevalence of OA is 11.5%, 13.6% in women and 9.3% in men. The OA hazard ratios for obesity are 2.0 in women and 1.7 in men. The effect of OA diagnosis on HRQL, as measured by the Health Utilities Index Mark 3 (HUI3), is to reduce it by 0.10 in women and 0.14 in men.

*Conclusions*: We describe the development of the first population-based microsimulation model of OA. Strengths of this model include the use of large population databases to derive the key parameters and the application of modern microsimulation technology. Limitations of the model reflect the limitations of administrative and survey data and gaps in the epidemiological and HRQL literature.


*Key words*: Osteoarthritis, Epidemiology, Microsimulation, Modeling, Population, Risk factors, Quality of life, Policy evaluation.

### Introduction

Osteoarthritis (OA) is the most common form of arthritis and a leading cause of disability[1,2]. As the population ages, the number of persons suffering from OA is expected to increase[3–6]. A number of strategies have been advocated to reduce the burden of OA[7–11]. However, quantification of the potential impact of such strategies on future disease burden is a complex undertaking. The effects of important risk factors for OA, such as age and overweight/obesity, are nonlinear and the distribution of both factors in the population is changing[2,12,13]. Similarly, extrapolating the benefits and side-effects of interventions from randomized trials to a population setting is not straightforward as one has to consider the existing treatment patterns, heterogeneity of effects, as well as changes in the demographic structure of the population[14]. Because of such complexities, standard economic and policy analyses of OA burden typically make simplifying assumptions about the population under study and the impact of interventions over time[15,16].

Simulation modeling of disease processes in a population can inform health policy and has consequently become increasingly accepted and adopted[17]. Recent developments in modeling technology include microsimulation, continuous-time models, and advances in model calibration[18]. An important distinction is between macrolevel (group-based) models and microsimulation. In macrolevel models, groups of individuals move through several stages based on transition probabilities and the summary measures of interest are calculated. In contrast, microsimulation models simulate individual life histories, with the timing of events determined stochastically[19]. The main advantages of microsimulation include a potentially infinite number of events and stages in disease process that

*Address correspondence and reprint requests to: J. A. Kopec, School of Population and Public Health, University of British Columbia, Research Scientist, Arthritis Research Centre of Canada, 895 West 10th Avenue, Vancouver, BC V5Z 1L7, Canada. Tel: 1-604-871-4588; Fax: 1-604-879-3791; E-mail: jkopec@arthritisresearch.ca
Received 4 April 2009; revision accepted 15 October 2009.

can be modeled and flexibility in defining input populations, merging multiple data sources, taking into account heterogeneity of effects, and implementing changes in the model[18–21]. Thanks to a dramatic progress in computer technology it is now possible to perform the calculations needed to simulate millions of individual life histories[18]. Complex population-level processes can be described by a series of relatively simple equations and probabilistic parameters at the level of the individual. The model integrates such individual-level data and translates them into projections of disease burden in the population. By changing the input parameters, such as the distribution of disease risk factors and comparing the likely outcomes, it is possible to evaluate various interventions in terms of their expected impact on disease incidence, prevalence, mortality, quality-adjusted life years (QALYs), and costs[17–19]. Although simulation models have been used in arthritis research for cost-effectiveness and decision analyses[22–27], none of these models has been based on generating individual life histories in a population or included a primary prevention component.

The purpose of the study was to develop a simulation model of OA in a Canadian population that can be used to quantify the future health and economic burden of OA under a range of plausible scenarios for changes in the risk factors and treatments. In this article we provide an overview of the model development process, with a focus on OA incidence and its impact on quality of life. Modeling of the impact of OA treatment (both surgical and non-surgical) and economic aspects of OA will be presented in separate publications. The following stages in model development are described: (1) conceptualization of the disease; (2) implementation of a computer simulation program; (3) derivation of the model parameters; and (4) model calibration and validation. The key assumptions and limitations of the model are discussed.

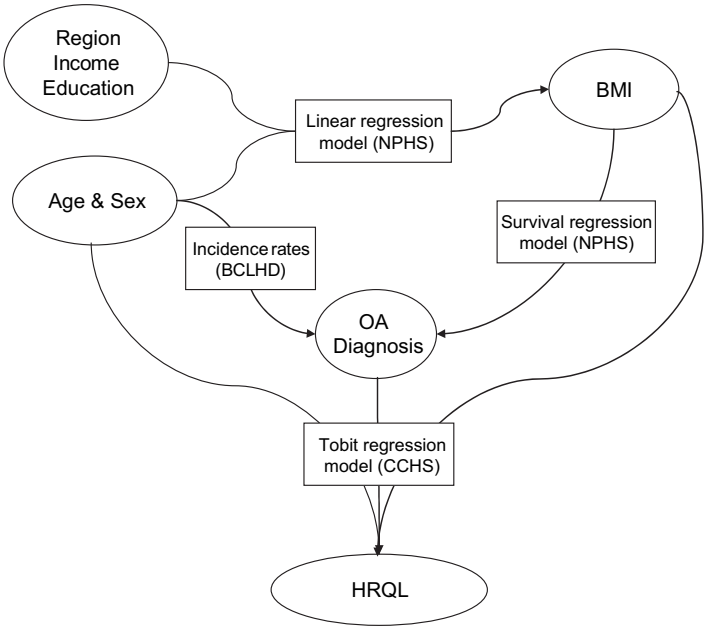
## Methods

### CONCEPTUAL FRAMEWORK

POHEM-OA is a population-based simulation model. The disease is defined as physician-diagnosed OA of any site, as opposed to radiographic or clinical OA. Population-level data on radiographically or clinically defined OA of specific joints are not available in Canada. Furthermore, this is an incidence-driven model – the key parameter is OA incidence (hazard) rate. The incidence rate is modeled as a function of established risk factors for OA, allowing for non-linear dose-response relationships and effect modification. Trends in OA incidence are determined by changes in the distribution of risk factors in the population over time. The burden of OA can be expressed in terms of the number of cases and impact on health-related quality of life (HRQL). The key components of the model and relationships between them are presented in Fig. 1. Main model parameters, methods of derivation, and data sources are listed in Table I. Baseline population data and technical details are provided in the Appendix.

### SIMULATION PLATFORM

Simulations are performed using the Population Health Model (POHEM) platform[28]. POHEM is a generic, continuous-time microsimulation environment developed at Statistics Canada[29]. The unit of simulation is the individual and the events can occur at arbitrarily small time intervals. This modeling approach differs from group-based methods of modeling proportions and




In the diagram, bubbles show the variables in the model, arrow connectors indicate the relationships between the variables, boxes depict the statistical models of changes in the variables and (in parentheses) sources of data. Baseline distributions of the variables were obtained from the 2001 cycle of the Canadian Community Health Survey (CCHS). Treatment effects and costs are not included in the diagram.
POHEM = Population Health Model; OA = Osteoarthritis; HRQL = Health-Related Quality of Life; BMI = Body Mass Index; BCLHD = British Columbia Linked Health Database; NPHS = National Population Health Survey; CCHS = Canadian Community Health Survey

Fig. 1. POHEM-OA: Schematic representation of the key relationships and sources of data pertaining to OA incidence and impact on HRQL.

Table I
*Key POHEM-OA parameters, data sources, and methods of derivation*

| Parameter | Source and method of derivation |
|---|---|
| Distribution of baseline population by age, sex, province, education, income, BMI and HUI3 | Observed in CCHS (2001) |
| Baseline prevalence of OA by age and sex | Obtained as the final stable prevalence from a simulation of the Canadian population over a 50-year time horizon, under constant age/sex-specific, BCLHD-based incidence rates. |
| Mortality rates by age and sex over time | Based on mortality data and using Statistics Canada projections of mortality for Canada. |
| Baseline incidence rates of OA by age and sex | Obtained from BCLHD. OA was defined as at least two visits to a health professional within 2 years or one hospitalization with the International Classification of Diseases, Ninth Revision (ICD-9) code 715. Incident cases were identified after excluding prevalent cases, using a 12-year run-in period. |
| Reference incidence rates of OA (for persons with normal BMI) by age and sex | Obtained numerically using an iterative algorithm (calibration). These rates are based on the incidence rates from BCLHD and the distribution of BMI levels in the CCHS population, within age/sex groups. |
| Effect of BMI on incident OA, by sex | Obtained from a survival regression model, separately for men and women, and adjusted for age. BMI was treated as a categorical variable ($<$18.5, 18.5–25, 25–30, 30+). The model is based on longitudinal data from the NPHS (2000–2002). |
| Effects of OA, age, sex, and BMI on HUI3 | Obtained from a tobit regression model including age, sex, BMI, and OA status (with interactions as required). The model is based on cross-sectional data from the CCHS (2001). |
| Effects of prior BMI, sex, education, income, and region on change in BMI | Obtained from a series of linear regression models for groups defined by BMI and age categories. The models are based on longitudinal data from the NPHS (1996–2004). |

HUI3 = Health Utilities Index Mark 3.

classic Markov models which assume constant transition rates between pre-specified states[30,31]. POHEM generates individual life trajectories within a large population representative of Canada, one individual at a time, until death. The stochastic nature of POHEM means that replicated lives with identical initial conditions have different life trajectories, because they have a different stream of random numbers used in their decision path. POHEM has been validated and used to assess the impact of prevention strategies, treatments, and cost of care in breast, lung, and colorectal cancer[32–38].

## PARAMETERS AND SOURCES OF DATA

### *Age and gender*

POHEM-OA uses the 2001 Canadian Community Health Survey (CCHS) sample as the baseline population[39]. The CCHS is a cross-sectional survey conducted every 2 years, with a total national sample of 130,000. The CCHS target population includes persons age 12+, living in households in the 10 Canadian provinces, with the exception of Indian reserves, military bases, and some remote northern areas. POHEM-OA includes subjects 18 years of age or older. Each subject from the CCHS sample is replicated at time of initialization to reflect its survey sample weight. Consequently, approximately 25 million individual lives, reflecting the non-institutionalized adult Canadian population, are simulated in a single run. Since POHEM-OA simulates a dynamic population, subjects are removed from the population as a result of death and new subjects are added to the population by aging into it (on their 18th birthday). To simulate these events, death and birth rates by age and sex were obtained from the national vital statistics and census databases routinely used by Statistics Canada for demographic projections[40].

### *Body mass index (BMI)*

At the population level, the most important modifiable risk factor for OA is BMI[2,12,41]. The distribution of BMI in the baseline population was obtained directly from the CCHS, based on the formula BMI = weight/height$^2$ (Table I). Individual BMI trajectories were simulated using a regression model based on longitudinal data from the National Population Health Survey (NPHS) in Canada[42]. The NPHS started in 1994 with a random household population sample of about 20,000, with a sampling frame similar to that of the CCHS. These subjects are followed every 2 years[43]. The model to simulate BMI used data from 1996 to 2004 and included age, sex, province of residence, education, income, and prior BMI.

### *OA incidence*

In POHEM-OA, OA incidence at the start of simulation is adjusted to reflect the incidence of physician-diagnosed OA in Canada. This adjustment, referred to as calibration, uses age/sex-specific incidence rates for 2003/4 (the last year for which data were available), derived from the British Columbia Linked Health Database (BCLHD) (Table I). The BCLHD is a well-established administrative database in the province of British Columbia, Canada that has been used extensively for research purposes[44]. It contains data on physician billings in the publicly funded health care system covering >95% of the population. The methodology for estimating incidence rates of OA from administrative data has been described elsewhere[45]. The definition of OA required either two physician visits within 2 years of each other or one hospitalization with the ICD-9 code 715 (Osteoathrosis and allied disorders) or the corresponding ICD-10 codes. To obtain incidence rates for Canada, age/sex-specific BC rates were applied to the age/sex distribution of the Canadian population.

### *OA prevalence*

As not all individuals with diagnosed OA seek regular medical care for their condition[46], age/sex-specific prevalence proportions of OA from an administrative database may underestimate the true prevalence, particularly if the run-in time is too short[45]. To ensure consistency between incidence and prevalence data, we simulated OA prevalence in the Canadian population over a long period of time (up to 50 years) while keeping the incidence rates constant at the most recent level. The age/sex-specific prevalence proportions obtained from this simulation were applied to the baseline CCHS population to obtain the distribution of OA at baseline[47].

### *Effect of BMI*

When simulating the occurrence of OA in individuals, the age/sex-specific hazard rates are modified according to the person's BMI. The effect of BMI on OA risk has been estimated using longitudinal data from the two cycles of the NPHS (2000 and 2002) that asked the question on any self-reported physician-diagnosed OA and were available for analysis. To this end, we have fit a multivariable survival regression model. The effect of BMI is expressed as hazard ratios for underweight (BMI $<$ 18.5), overweight (BMI 25.0–29.9) and obesity (BMI 30+), compared to normal BMI (18.5–24.9). The effects were estimated separately for males and females and adjusted for age (Table I).

### *HRQL*

POHEM-OA uses the Health Utilities Index Mark 3 (HUI3) as a measure of HRQL. HUI3 is based on eight attributes (vision, hearing, speech, ambulation, dexterity, emotional function, cognitive function, and pain)[48]. The attributes are combined into an overall index, ranging from −0.36 to 1, using a multi-attribute utility theory model based on societal preferences. Therefore, HUI3 can be employed to estimate QALY. The initial distribution of HUI3 in POHEM-OA is obtained directly from the CCHS. Individual HUI3 trajectories were simulated with a tobit regression model (accounting for the skewed distribution of HUI3 scores) that included age, gender, BMI, and OA status[49]. The model was derived from the 2001 cycle of the CCHS (Table I). Model fit was evaluated by the likelihood ratio test and a comparison of the scale parameter, which estimates the standard deviation of the normal error term, between the reduced and saturated model.

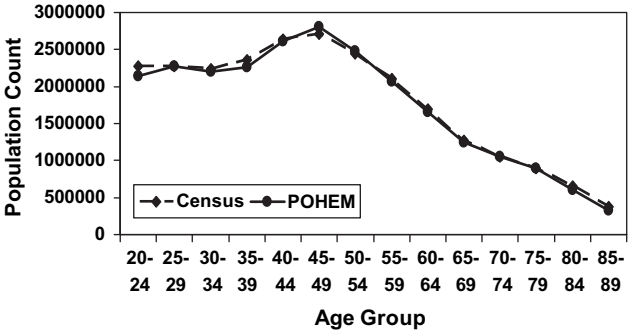
### MODEL CALIBRATION AND VALIDATION

The population modeled by POHEM-OA reflects the Canadian adult household population surveyed by the CCHS. As the model simulates individual life histories, the risk of OA is influenced by individual BMI values according to the relative risk derived from the NPHS. In order that the aggregated age/sex-specific incidence rates in the simulated population correspond to the observed population rates in administrative data, given the population distribution of BMI, the age/sex-specific rates for the reference BMI category need to be estimated. This was achieved through model calibration, whereby the reference rates were iteratively adjusted using numerical methods until the simulated and observed incidence rates (within age/sex categories) agreed[18,19].

The population distribution by age from 2001 to 2007 in POHEM-OA has been compared to the actual Canadian population-based on estimates from the 2001 and 2006 censuses. To validate the simulated OA frequency, we compared the incidence and prevalence of OA in POHEM-OA with self-reported incidence of arthritis and rheumatism in the NPHS, self-reported OA prevalence in the CCHS and BCLHD, and clinical and radiographic OA incidence and prevalence from the literature[45,47,50,51]. We also assessed the effects of OA definition and run-in time on the incidence and prevalence of OA in administrative data[45]. In addition, we have estimated the sensitivity and specificity of the administrative definition of OA by linking the BCLHD to a population-based cohort study of 171 subjects with extensive clinical and radiographic data[52]. Finally, our model for predicting HUI3 in the general population has been validated by comparing the predicted and observed HUI3 scores in a different cycle of the CCHS.

## Results

POHEM-OA has a large number of parameters, including descriptive parameters, such as age and sex distribution at baseline, birth and death rates by age, sex, and year, education and income distribution, age/sex-specific OA prevalence proportions at baseline, age/sex-specific OA incidence rates, and other parameters. In this article we discuss the key input parameters related to the frequency and health impact of OA (Table I).




POHEM-based counts use the 2001 Canadian Community Health Survey as baseline population. Census-based counts are estimated using the 2006 Canadian Census

Fig. 2. Comparisons of POHEM-based and Census-based population counts for Canada in 2007, by age group.

Table II
*Age/sex-specific incidence rates of physician-diagnosed OA per 1000 person-years in British Columbia, Canada*

| Age range | Men | Women |
|---|---|---|
| 20−24 | 0.54 | 0.63 |
| 25−29 | 0.91 | 0.82 |
| 30−34 | 1.54 | 1.55 |
| 35−39 | 2.55 | 2.59 |
| 40−44 | 4.05 | 4.09 |
| 45−49 | 6.22 | 7.32 |
| 50−54 | 8.10 | 12.04 |
| 55−59 | 11.36 | 18.21 |
| 60−64 | 14.66 | 22.48 |
| 65−69 | 17.59 | 26.93 |
| 70−74 | 20.30 | 31.00 |
| 75−79 | 23.26 | 34.47 |
| 80−84 | 24.22 | 33.77 |
| 85−89 | 25.77 | 33.42 |
| 90+ | 25.54 | 31.55 |

Incident OA was defined as at least two visits to a physician with ICD-9 code 715 within 1−730 days or one hospital diagnosis with a diagnostic code 715 in a person not diagnosed with OA in the previous 12 years. Rates are calculated for the period from April 1st 2003 to March 31st, 2004, based on data from the BC Linked Health Database.

In Fig. 2, we compare the simulated age distribution in POHEM-OA in 2007 with the distribution estimated from the 2006 census. The data show a nearly perfect correspondence between the two datasets. Minor discrepancies are mainly due to the exclusion of the institutionalized populations in the CCHS and differences between the weights developed for the CCHS in 2001 and the revised census estimates.

Age/sex-specific incidence rates of physician-diagnosed OA in POHEM-OA are presented in Table II. The rates increase approximately linearly with age in both sexes between the ages of 50 and 80 and plateau in the very old. In persons 50+, women have substantially higher rates than men, with the male-to-female ratio between 0.6 and 0.7 in most age groups. The baseline prevalence of OA generated from the POHEM-OA simulation is 11.5%, 13.6% in women and 9.3% in men.

In Table III we provide the age-adjusted OA hazard ratios for the different BMI categories, separately for men and women. Since the main purpose of this analysis was to provide the input parameters for the simulation model, confidence intervals are not reported. The effect of BMI is stronger in women, with the hazard ratios of 1.76 and 2.03 for BMI 25−29.9 and 30+, respectively, compared to normal weight (BMI between 18.5 and 24.9). The corresponding hazard ratios for men are 1.07 and 1.69. Rates are decreased for men and women with BMI $< 18.5$ (in our sample there were no cases of OA among men with BMI $< 18.5$).

Table III
*Effect of BMI on OA incidence in the NPHS in Canada (2000−2002)*

| | Underweight BMI $< 18.5$ | Normal weight $18.5 \leq$ BMI $< 25$ | Overweight $25 \leq$ BMI $< 30$ | Obese BMI $\geq 30$ |
|---|---|---|---|---|
| Females | 0.33 | 1.0 | 1.76 | 2.03 |
| Males | 0.00 | 1.0 | 1.07 | 1.69 |

The estimates are age-adjusted hazard ratios from a survival regression model. OA is defined as self-reported physician-diagnosed OA. BMI = weight/height$^2$.

Table IV
*Effects of age, sex, OA diagnosis and BMI category on HUI3 scores in the CCHS (2001)*

| | Regression coefficient |
|---|---|
| Intercept | 0.62 |
| Age 12–19 | 0.30 |
| Age 20–29 | 0.29 |
| Age 30–39 | 0.29 |
| Age 40–49 | 0.24 |
| Age 50–59 | 0.20 |
| Age 60–69 | 0.19 |
| Age 70–79 | 0.15 |
| Age 80+ (ref) | 0.00 |
| Sex – male | 0.06 |
| Sex – female (ref) | 0.00 |
| OA diagnosis – yes | −0.10 |
| OA diagnosis – no (ref) | 0.00 |
| BMI 0–18.4 | 0.05 |
| BMI 18.5–24.9 | 0.08 |
| BMI 25.0–29.9 | 0.06 |
| BMI 30+ (ref) | 0.00 |
| BMI (0–18.4) × Sex (M) | −0.07 |
| BMI (18.5–24.9) × Sex (M) | −0.06 |
| BMI (25.0–29.9) × Sex (M) | −0.03 |
| BMI (30+) × Sex (ref) | 0.00 |
| OA (yes) × Sex (M) | −0.04 |
| OA (no) × Sex (ref) | 0.00 |

HUI3 = Health Utilities Index Mark 3.

In POHEM-OA, HRQL (measured by HUI3) is predicted by OA, age, sex and BMI, using a tobit model[49] (Table IV). The effects of OA and BMI on HUI3 were different in men and women. This difference in effects is captured by the interactions between BMI and sex and between OA and sex that were highly significant. The average impact of OA diagnosis is to decrease HUI3 score by 0.10 in women and 0.14 in men. Age has a negative effect on HUI3 and women have, on average, lower HUI3 scores than men. Model fit was adequate, as suggested by the estimated error term which was very similar in the reduced and saturated models (data not shown). The likelihood ratio test was significant, indicating that additional interaction terms might improve the fit of the model, but this test is less informative when the sample size is very large.

## Discussion

This article describes the methodology and key parameters for a population-based simulation model of OA in Canada, with a focus on OA incidence and its impact on quality of life. Most of the input parameters have been derived from the Canadian Census and vital statistics databases, national surveys, administrative data in British Columbia, and the health literature. Advanced statistical methods have been used to derive time-to-event distributions and the relationships between the variables. Model calibration ensures that the simulated incidence rates agree with the observed rates.

The model uses the POHEM platform, a state-of-the-art microsimulation tool that can simulate a dynamic population as it changes over time, as opposed to modeling a fixed cohort or a stationary population. Events are modeled in continuous time and there is no limit on the number of events that can be simulated. An important feature of POHEM-OA is the transparency of the model. The structure of the model, parameter values, and sources of data are publicly available. Assumptions and limitations of the model are acknowledged and discussed.

OA incidence rates in POHEM-OA are higher than those estimated by Oliveria *et al.* for radiographic and symptomatic OA of the knee, hip and hand, derived from an administrative database in the US[50]. This is epidemiologically plausible since our definition included all OA sites and did not require radiographic confirmation. At the same time, our rates are lower than published incidence rates of self-reported "arthritis or rheumatism" in Canada[51]. Direct comparisons between the clinical and administrative diagnosis of OA are scarce. Harrold *et al.*[53] found a positive predictive value of 62% in the US using medical records as a gold standard. In our previous validation study, the administrative definition of OA had sensitivity around 30% and specificity over 90% against a diagnosis based on clinical and radiographic criteria[52].

The simulation-derived estimates of baseline OA prevalence in POHEM-OA are higher than those directly obtained from administrative data[45]. The most likely reason is the tendency for administrative data to underestimate OA prevalence due to insufficient run-in time. This tendency has been previously demonstrated in our database[45]. Our estimates of baseline prevalence are consistent with the most recent incidence rates observed in the data. Although our previous study using the BC database suggested an increase in age-standardized incidence rates of OA among women between 1996 and 2003, data on long-term trends in OA incidence are not available[3]. For this reason, and to simplify the analysis, we assumed that the age/sex-specific incidence rates in the past were constant. Baseline prevalence in our model is overestimated if the incidence rates were lower prior to baseline (i.e., if there was an increasing trend), and underestimated if the incidence rates were higher, but the difference is very unlikely to be substantial.

The risk of developing OA in POHEM-OA is determined by age, gender and BMI. The effect of BMI on the OA incidence rate was estimated from a longitudinal population survey in Canada using self-reported physician-diagnosed OA. Comparative data for a similar definition of OA are not available. In a study of radiographic knee OA by Felson *et al.*[41] in the Framingham cohort, the age-adjusted odds ratios in the highest and second highest quintiles of height-adjusted weight were 2.07 and 1.44, respectively, for women, and 1.51 and 1.00, respectively, for men. Given the differences in definitions and methodology between the two studies, the results are remarkably similar to the hazard ratios from our model (Table III).

The effects of age, sex and OA on HUI3 were estimated from the CCHS, a large national survey in Canada. Comparative population data for a similar definition of OA are not available. Schultz and Kopec[54] previously estimated the overall effect of "arthritis and rheumatism" on HUI3 at −0.09 using data from the 1996/7 NPHS and adjusting for co-morbidity. Similar to the current study, the effect was stronger in men. In a graphical analysis, our model predicted HUI3 scores observed in a different CCHS cycle with acceptable accuracy (data not shown).

While the results of preliminary validation are promising, there is a need for more validation studies. Future validation will include re-evaluation of the current input parameters, comparisons of model output with the actual trends observed in administrative and survey data from several Canadian provinces, and stochastic sensitivity analyses.

The definition of OA adopted in our model is useful from a healthcare utilization perspective but it has clear limitations. Our model is unable to distinguish between different OA sites, which limits its ability to model heterogeneous effect of risk factors on OA incidence. In terms of modeling the impact of OA on quality of life, many cases of symptomatic or even disabling OA remain undiagnosed. The effect of our OA definition on the estimates of disease burden in

POHEM-OA will be evaluated in a sensitivity analysis. Modeling joint pain and other symptoms in the population, as opposed to physician-diagnosed OA, might be a potentially useful alternative in future studies.

We acknowledge that the values of the specific parameters in POHEM-OA are potentially subject to criticism and may need to be modified as new and better data become available. Different studies may produce different estimates of the incidence, prevalence, relative risk, or health impact of OA. A limitation of the current version of POHEM-OA is that it does not include other factors that may potentially affect the incidence of OA, such as geographic region, race, socio-economic status, injury, physical activity, or family history of OA[2]. Data on the distribution of these factors in the population, changes over time, and causal effects on the risk of OA are insufficient at this time. Projections of disease incidence and prevalence from our model may be inaccurate if there are temporal changes in the distribution of these factors, as well as other and unknown risk factors for OA. This limitation can be minimized by introducing a correction for trend in incidence rates based on historical data. Furthermore, the effect of BMI on OA incidence in the model is derived from data on self-reported physician-diagnosed OA of any joint. While this effect may be attenuated due to unavoidable misclassification of the disease, it is based on the best Canadian data currently available.

At present, POHEM-OA does not contain any parameters describing the effects of specific interventions on changes in health behaviors, such as diet or healthcare utilization, in different population groups. Implementing such parameters would allow a more realistic simulation of the impact of health policies across various segments of the population. Although the available data are limited, we intend to incorporate behavioral effects into future versions of the model.

Despite its limitations, the current model can inform policy by providing information not directly available from epidemiological studies. Its main purpose is to study the effects of changes in the distribution of risk factors for OA on the future burden of this disease in a dynamically changing population. Current policy analyses typically make simplifying assumptions about the population under study and the effects of health interventions at the population level[15,16]. In particular, standard methods of calculating attributable fraction usually assume that the risk factor is eliminated and that the underlying population is static. In addition, such methods generally do not consider the impact of disease on quality of life[14].

The current model can be used to project the future burden of OA in Canada while taking into account trends in population aging and obesity prevalence. For example, we have recently applied the model to compare different scenarios for future OA incidence rates in terms of their impact on health-adjusted life expectancy (HALE) in Canada[55]. Preliminary results show that if new cases of OA due to excess weight could be avoided, average HALE would improve by 2–4 months and the gain would be somewhat greater in women. Another example is the use of the model to analyze the relative contributions of aging and obesity to the projected increase in the number of persons with OA[56].

In conclusion, we have developed the first population-based microsimulation model of OA. We hope the model will provide a benchmark for the synthesis of quantitative data on the epidemiology and population health impact of OA. We expect to be able to improve the model as gaps in knowledge are addressed in future studies. While the initial results from this model will have to be taken with caution, the confidence in the results will increase as the validity of the model is further confirmed under a variety of conditions.

## Conflict of interest

There are no conflicts of interest to be disclosed.

## Acknowledgements

The authors thank Claude Nadeau for help in revising the manuscript.
Source of support: The study was funded by a grant from the Canadian Institutes of Health Research.

## Appendix

*Population parameters*

Tables A1–A7 shows the distribution of key POHEM-OA variables in the CCHS. These tables include all CCHS respondents. Weighted data reflect the distribution in the Canadian household population. Data in Tables A8–A10 are shown for illustration purposes. The estimates are rounded and shown for selected ages and/or years only. When running POHEM, more precise estimates for all ages and years are used.

Table A1
*Distribution of the CCHS (2001) population by province of residence*

| Province | Sample | | Population (weighted) | |
|---|---|---|---|---|
| | *N* | % | *N* | % |
| Newfoundland | 3,870 | 3.0 | 461,104 | 1.8 |
| Prince Edward Island | 3,651 | 2.8 | 116,327 | 0.5 |
| Nova Scotia | 5,319 | 4.1 | 787,972 | 3.1 |
| New Brunswick | 4,996 | 3.8 | 634,264 | 2.5 |
| Quebec | 22,012 | 16.8 | 6,216,722 | 24.1 |
| Ontario | 39,278 | 30.0 | 9,877,292 | 38.3 |
| Manitoba | 8,470 | 6.5 | 907,494 | 3.5 |
| Saskatchewan | 8,009 | 6.1 | 805,993 | 3.1 |
| Alberta | 14,456 | 11.0 | 2,481,568 | 9.6 |
| British Columbia | 18,302 | 14.0 | 3,421,671 | 13.3 |
| Yukon and NWT | 2,517 | 1.9 | 76,928 | 0.3 |
| Total | 130,880 | 100.0 | 25,787,334 | 100.0 |

NWT = Northwest Territories.

Table A2
*Distribution of the CCHS (2001) population by age group*

| Age group | Sample | | Population (weighted) | |
|---|---|---|---|---|
| | *N* | % | *N* | % |
| 12–14 | 6,476 | 4.9 | 1,186,119 | 4.6 |
| 15–19 | 11,081 | 8.5 | 2,131,999 | 8.3 |
| 20–24 | 7,584 | 5.8 | 2,112,568 | 8.2 |
| 25–29 | 8,742 | 6.7 | 2,006,021 | 7.8 |
| 30–34 | 10,281 | 7.9 | 2,158,989 | 8.4 |
| 35–39 | 12,447 | 9.5 | 2,587,642 | 10.0 |
| 40–44 | 12,886 | 9.8 | 2,707,970 | 10.5 |
| 45–49 | 11,388 | 8.7 | 2,369,433 | 9.2 |
| 50–54 | 10,255 | 7.8 | 2,051,946 | 8.0 |
| 55–59 | 8,355 | 6.4 | 1,585,225 | 6.1 |
| 60–64 | 7,152 | 5.5 | 1,244,611 | 4.8 |
| 65–69 | 6,842 | 5.2 | 1,151,556 | 4.5 |
| 70–74 | 6,360 | 4.9 | 1,003,709 | 3.9 |
| 75–79 | 5,237 | 4.0 | 740,459 | 2.9 |
| 80+ | 5,794 | 4.4 | 749,088 | 2.9 |
| Total | 130,880 | 100.0 | 25,787,335 | 100.0 |

Table A3
*Distribution of the CCHS (2001) population by sex*

| Sex | Sample | | Population (weighted) | |
|---|---|---|---|---|
| | *N* | % | *N* | % |
| Male | 60,514 | 46.2 | 12,714,150 | 49.3 |
| Female | 70,366 | 53.8 | 13,073,184 | 50.7 |
| Total | 130,880 | 100.0 | 25,787,335 | 100.0 |

Table A4
*Distribution of the CCHS (2001) population by level of education*

| Education | Sample | | Population (weighted) | |
|---|---|---|---|---|
| | *N* | % | *N* | % |
| Less than secondary | 44,338 | 33.9 | 7,594,745 | 29.5 |
| Secondary | 22,860 | 17.5 | 4,758,801 | 18.5 |
| Some post-secondary | 9,832 | 7.5 | 2,107,601 | 8.2 |
| Post-secondary grad. | 52,586 | 40.2 | 11,108,414 | 43.1 |
| Not stated | 1,264 | 1.0 | 217,774 | 0.8 |
| Total | 130,880 | 100.0 | 25,787,335 | 100.0 |

Table A5
*Distribution of the CCHS (2001) population by income*

| Income | Sample | | Population (weighted) | |
|---|---|---|---|---|
| | *N* | % | *N* | % |
| None | 5,704 | 4.4 | 1,327,650 | 5.1 |
| <$15,000 | 34,229 | 26.2 | 6,141,039 | 23.8 |
| $15,000–29,999 | 28,061 | 21.4 | 5,281,922 | 20.5 |
| $30,000–49,999 | 24,220 | 18.5 | 5,191,844 | 20.1 |
| $50,000–79,999 | 13,578 | 10.4 | 2,945,634 | 11.4 |
| $80,000 or more | 4,524 | 3.5 | 1,081,049 | 4.2 |
| Not applicable | 6,476 | 4.9 | 1,186,119 | 4.6 |
| Not stated | 14,088 | 10.8 | 2,632,077 | 10.2 |
| Total | 130,880 | 100.0 | 25,787,334 | 100.0 |

Table A6
*Distribution of the CCHS (2001) population by BMI*

| BMI | Sample | | Population (weighted) | |
|---|---|---|---|---|
| | *N* | % | *N* | % |
| Underweight (BMI $<20$) | 6,040 | 4.6 | 1,473,150 | 5.7 |
| Acceptable (BMI 20–24.9) | 35,404 | 27.1 | 7,944,017 | 30.8 |
| Overweight (BMI $\geq 25$) | 44,730 | 34.2 | 8,882,610 | 34.4 |
| Not applicable | 42,866 | 32.8 | 7,189,595 | 27.9 |
| Not stated | 1,840 | 1.4 | 279,962 | 1.2 |
| Total | 130,880 | 100.0 | 25,787,334 | 100.0 |

Table A7
*Distribution of the CCHS (2001) population by the Health Utilities Index Mark 3 (HUI3)*

| HUI3 | Sample | | Population (weighted) | |
|---|---|---|---|---|
| | *N* | % | *N* | % |
| <0.0 | 799 | 0.6 | 138,751 | 0.5 |
| 0.0–0.099 | 888 | 0.7 | 151,656 | 0.6 |
| 0.1–0.199 | 1,082 | 0.8 | 180,531 | 0.7 |
| 0.2–0.299 | 1,916 | 1.5 | 308,631 | 1.2 |
| 0.3–0.399 | 2,979 | 2.3 | 501,022 | 1.9 |
| 0.4–0.499 | 2,428 | 1.9 | 427,027 | 1.7 |
| 0.5–0.599 | 2,438 | 1.9 | 425,220 | 1.7 |
| 0.6–0.699 | 6,031 | 4.6 | 1,096,935 | 4.3 |
| 0.7–0.799 | 8,545 | 6.5 | 1,578,011 | 6.1 |
| 0.8–0.899 | 12,469 | 9.5 | 2,388,627 | 9.3 |
| 0.9–0.999 | 57,344 | 43.8 | 11,504,065 | 44.6 |
| 1.0 | 32,272 | 24.7 | 6,855,037 | 26.6 |
| Missing | 1,689 | 1.3 | 231,822 | 0.9 |
| Total | 130,880 | 100.0 | 25,787,334 | 100.0 |

Table A8
*Baseline prevalence (%) of OA for selected ages*

| Age | Men | Women |
|---|---|---|
| 20 | 0.1 | 0.1 |
| 30 | 0.9 | 0.9 |
| 40 | 2.9 | 3.0 |
| 50 | 7.9 | 8.8 |
| 60 | 16.8 | 22.2 |
| 70 | 29.2 | 39.8 |
| 80 | 43.0 | 56.1 |
| 90 | 54.8 | 68.2 |

Table A9
*Reference incidence rates of OA per 1000 person-years for selected ages*

| Age | Men | Women |
|---|---|---|
| 20 | 0.5 | 0.6 |
| 30 | 1.4 | 1.2 |
| 40 | 3.5 | 3.1 |
| 50 | 6.8 | 8.2 |
| 60 | 13.1 | 16.4 |
| 70 | 18.3 | 23.2 |
| 80 | 23.8 | 29.9 |
| 90 | 28.1 | 31.1 |

Table A10
*Mortality rates in Canada per 1000 population for selected ages and years (including projected rates)*

| Age | 2001 | 2006 | 2011 | 2016 | 2021 |
|---|---|---|---|---|---|
| *Women* | | | | | |
| 20 | 0.26 | 0.25 | 0.24 | 0.24 | 0.23 |
| 30 | 0.35 | 0.34 | 0.33 | 0.32 | 0.31 |
| 40 | 0.69 | 0.68 | 0.66 | 0.64 | 0.63 |
| 50 | 1.78 | 1.74 | 1.67 | 1.65 | 1.61 |
| 60 | 4.70 | 4.58 | 4.47 | 4.36 | 4.25 |
| 70 | 12.28 | 11.98 | 11.69 | 11.40 | 11.12 |
| 80 | 31.03 | 30.26 | 29.52 | 28.79 | 28.08 |
| 90 | 73.83 | 72.01 | 70.23 | 68.50 | 66.81 |
| *Men* | | | | | |
| 20 | 0.59 | 0.52 | 0.49 | 0.47 | 0.45 |
| 30 | 0.73 | 0.70 | 0.66 | 0.63 | 0.60 |
| 40 | 0.84 | 0.80 | 0.76 | 0.72 | 0.69 |
| 50 | 2.07 | 2.00 | 1.87 | 1.78 | 1.69 |
| 60 | 5.73 | 5.45 | 5.19 | 4.94 | 4.70 |
| 70 | 14.43 | 13.73 | 13.07 | 12.44 | 11.84 |
| 80 | 34.41 | 32.75 | 31.17 | 29.67 | 28.24 |
| 90 | 81.33 | 77.40 | 73.67 | 70.11 | 66.73 |

INCIDENCE MODEL

The following steps describe the basic algorithm for determining when OA occurs in the microsimulation:

(1) POHEM selects a record from the CCHS database in the simulation start year (2001).
(2) At each birthday in the person's simulated life, annualized hazard h is calculated as: $h = h_0 \times \mathrm{RR}$. In this

equation $h_0$ is the age/sex-specific OA incidence rate for persons with normal BMI estimated from BCLHD rates using calibration methods (Table A9 in the Appendix), and RR is the relative risk (hazard ratio) of OA based on person's BMI category (Table III in the manuscript).

(3) A random number u between 0 and 1 is generated and the time of event (in years) is estimated as $t = -\ln(1-u)/h$. If $t \geq 1$ no event occurs during the year. Note that competing events, such as death, could censor this event.
(4) Steps 2–3 are repeated at every subsequent birthday until the person develops OA or dies.
(5) Steps 1–4 are repeated for every individual record in our start-up database.

### BMI MODEL

The BMI model is part of the POHEM software and was developed by Statistics Canada using biannual data from the longitudinal NPHS 1996–2004. The model predicts current BMI within 14 groups defined by BMI category and age, separately for men and women, based on BMI history and other covariates, such as region of residence, income and education. The response variable is change in BMI compared to 2 years prior, treated as a continuous, normally distributed variable. The model includes the most recent BMI value and up to three prior changes in BMI. The model can be presented in a simplified form as:

$$\begin{aligned}\Delta\mathrm{BMI}_{t,t+2} = {} & \alpha + \beta_1\mathrm{BMI}_t + \beta_2\Delta\mathrm{BMI}_{t-2,t} + \beta_3\Delta\mathrm{BMI}_{t-4,t-2} \\ & + \beta_4\Delta\mathrm{BMI}_{t-6,t-4} + \beta_5\mathrm{Income}_{t-6} \\ & + \beta_6\mathrm{Education}_{t-6} + \beta_7\mathrm{Region}_{t-6}\end{aligned}$$

In this equation, $\Delta\mathrm{BMI}_{i,j} = \mathrm{BMI}_j - \mathrm{BMI}_i$ is the difference in BMI between time *i* and time *j*.

There are 112 regression equations arising from 28 strata (modeling groups) generated by the 14 age × BMI categories and the sex variable (male and female). Within each stratum, there are four models, depending on the number of prior BMI values used.

## References


1. Lawrence RC, Helmick CG, Arnett FC, Deyo RA, Felson DT, Giannini EH, *et al.* Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. Arthritis Rheum 1998;41(5):778–99.
2. Felson DT, Lawrence RC, Hochberg MC, McAlindon T, Dieppe PA, Minor MA, *et al.* Osteoarthritis: new insights. Part 1: the disease and its risk factors. Ann Intern Med 2000;33(8):635–46.
3. Kopec JA, Rahman MM, Sayre EC, Cibere J, Flanagan WM, Aghajanian J, *et al.* Trends in physician-diagnosed osteoarthritis incidence in an administrative database in British Columbia, Canada, 1996–1997 through 2003–2004. Arthritis Rheum 2008;59(7):929–34.
4. Badley EM, Wang PP. Arthritis and the aging population: projections of arthritis prevalence in Canada 1991–2031. J Rheumatol 1998;25(1):138–44.
5. Hootman JM, Helmick CG. Projections of US prevalence of arthritis and associated activity limitations. Arthritis Rheum 2006;54(1):226–9.
6. Perruccio AV, Power JD, Badley EM. Revisiting arthritis prevalence projections–it's more than just the aging of the population. J Rheumatol 2006;33(9):1856–62.
7. Woolf AD. The bone and joint decade. Strategies to reduce the burden of disease: the Bone and Joint Monitor Project. J Rheumatol Suppl 2003;67:6–9 [Review].
8. Bijlsma JW, Knahr K. Strategies for the prevention and management of osteoarthritis of the hip and knee. Best Pract Res Clin Rheumatol 2007;21(1):59–76.
9. Vignon E, Valat JP, Rossignol M, Avouac B, Rozenberg S, Thoumie P, *et al.* Osteoarthritis of the knee and hip and activity: a systematic international review and synthesis (OASIS). Joint Bone Spine 2006;73(4):442–55.
10. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, *et al.* OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;16(2):137–62.
11. Arthritis Foundation, Association of State and Territorial Health Officials, Centers for Disease Control and Prevention. National Arthritis Action Plan. A Public Health Strategy. 1999. Internet address, http://ww2.arthritis.org/resources/ActionPlanInterior.pdf; 2008 [accessed 20.01.09].
12. Lohmander LS, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Engström G. Incidence of severe knee and hip osteoarthritis in relation to different measures of body mass. A population-based prospective cohort study. Ann Rheum Dis 2009;68(4):490–6.
13. Luo W, Morrison M, de Groh M, Waters C, DesMeules M, Jones-McLean E, *et al.* The burden of adult obesity in Canada. Chronic Dis Can 2007;27(4):135–83.
14. Bartlett C, Doyal L, Ebrahim S, Davey P, Bachmann M, Egger M, *et al.* The causes and effects of socio-demographic exclusions from clinical trials. iii–iv, ix–x. Health Technol Assess 2005;9(38):1–152.
15. Andrews G, Simonella L, Lapsley H, Sanderson K, March L. Evidence-based medicine is affordable: the cost-effectiveness of current compared with optimal treatment in rheumatoid and osteoarthritis. J Rheumatol 2006;33(4):671–80.
16. Segal L, Day SE, Chapman AB, Osborne RH. Can we reduce disease burden from osteoarthritis? An evidence-based priority-setting model. Med J Aust 2004;180(Suppl 5):S11–7.
17. Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, *et al.* Systematic review of the use and value of computer simulation modelling in population health and health care delivery. J Public Health Med 2003;25(4):325–35.
18. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. Pharmacoeconomics 2006;24(11):1043–53.
19. Feuer EJ, Etzioni R, Cronin KA, Mariotto A. The use of modeling to understand the impact of screening on U.S. mortality: examples from mammography and PSA testing. Stat Methods Med Res 2004;13(6):421–42.
20. Brown JB, Russell A, Chan W, Pedula K, Aickin M. The global diabetes model: user friendly version 3.0. Diabetes Res Clin Pract 2000;50(Suppl 3):S15–46.
21. Le Lay A, Despiegel N, Francois C, Duru G. Related Articles. Can discrete event simulation be of use in modelling major depression? Cost Eff Resour Alloc 2006;5(4):19–30.
22. Bansback NJ, Regier DA, Ara R, Brennan A, Shojania K, Esdaile JM, *et al.* An overview of economic evaluations for drugs used in rheumatoid arthritis: focus on tumour necrosis factor-alpha antagonists. Drugs 2005;65(4):473–96.
23. Marshall JK, Pellissier JM, Attard CL, Kong SX, Marentette MA. Incremental cost-effectiveness analysis comparing rofecoxib with nonselective NSAIDs in osteoarthritis: Ontario ministry of health perspective. Pharmacoeconomics 2001;19(10):1039–49.
24. Zabinski RA, Burke TA, Johnson J, Lavoie F, Fitzsimon C, Tretiak R, *et al.* An economic model for determining the costs and consequences of using various treatment alternatives for the management of arthritis in Canada. Pharmacoeconomics 2001;19(Suppl 1):49–58.
25. Haglund U, Svarvar P. The Swedish ACCES model: predicting the health economic impact of celecoxib in patients with osteoarthritis or rheumatoid arthritis. Rheumatology (Oxford) 2000;39(Suppl 2):51–6.
26. Chancellor JV, Hunsche E, de Cruz E, Sarasin FP. Economic evaluation of celecoxib, a new cyclo-oxygenase 2 specific inhibitor. Pharmacoeconomics 2001;19(Suppl 1):59–75.
27. Maetzel A, Krahn M, Naglie G. The cost effectiveness of rofecoxib and celecoxib in patients with osteoarthritis or rheumatoid arthritis. Arthritis Rheum 2003;49(3):283–92.
28. Will BP, Berthelot J-M, Nobrega KM, Flanagan W, Evans WK. Canada's Population Health Model (POHEM): a tool for performing economic evaluations of cancer control interventions. Eur J Cancer 2001;37(14):1797–804.
29. Wolfson MC. POHEM—a framework for understanding and modelling the health of human populations. World Health Stat Q 1994;47(3–4):157–76.
30. Law A, Kelton WD. Simulation Modeling and Analysis. 3rd edn. McGraw-Hill; 1999.
31. Gupta A, Harding A, Eds. Modeling Our Future. Population Aging, Health and Aged Care. Amsterdam: Elsevier; 2007.
32. Will BP, Le Petit C, Berthelot J-M, Tomiak EM, Verma S, Evans WK. Diagnostic and therapeutic approaches for nonmetastatic breast cancer in Canada, and their associated costs. Br J Cancer 1999;79(9–10):1428–36.
33. Will BP, Nobrega KM, Berthelot J-M, Flanagan W, Wolfson MC, Logan DM, *et al.* First do no harm: extending the debate on the provision of preventive tamoxifen. Br J Cancer 2001;85(9):1280–8.

34. Evans WK, Will BP, Berthelot J-M, Logan DM, Mirsky DJ, Kelly N. Breast cancer: better care for less cost: is it possible? Int J Technol Assess Health Care 2000;16(4):1168–78.
35. Will BP, Berthelot J-M, Le Petit C, Tomiak E, Verma S, Evans WK. Estimates of the lifetime costs of breast cancer treatment in Canada. Eur J Cancer 2000;36(6):724–35.
36. Berthelot J-M, Will BP, Evans WK, Coyle D, Earle CC, Bordeleau L. Decision framework for chemotherapeutic interventions for metastatic non-small cell lung cancer. J Natl Cancer Inst 2000;92(16): 1321–9.
37. Flanagan W, Le Petit C, Berthelot J-M, White KJ, Coombs BA, Jones-McLean E. Potential impact of population-based colorectal cancer screening in Canada. Chronic Dis Can 2003;24(4):81–8.
38. Maroun J, Ng E, Berthelot J-M, Le Petit C, Dahrouge S, Flanagan WM, *et al*. Lifetime costs of colon and rectal cancer management in Canada. Chronic Dis Can 2003;24(4):91–101.
39. Statistics Canada. Canadian Community Health Survey (CCHS) Cycle 1.1 Internet address, http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SurvId=3226&SurvVer=0&InstaId=15282&InstaVer=1&SDDS=3226&lang=en&db=imdb&adm=8&dis=2. [accessed: 20.01.09]
40. George MV, Loh S, Verma RBP, Shin YE. Population Projections for Canada, Provinces and Territories 2000–2026. Statistics Canada, Demography Division. Catalogue no. 91-520-XPB ISBN 0-660-60634-8. Ministry of Industry; 2001.
41. Felson DT, Anderson JJ, Naimark A, Walker AM, Meenan RF. Obesity and knee osteoarthritis. The Framingham study. Ann Intern Med 1988;109(1):18–24.
42. Statistics Canada. National Population Health Survey (NPHS) Internet address, http://www.statcan.ca/english/concepts/nphs/; 1988. Last updated: 2004-08-20.
43. National Population Health Survey, Cycle 6 Synthetic Master File, 2007. Statistics Canada. Health Statistics Division. Internet address, http://data.library.ubc.ca/java/jsp/database/production/detail.jsp?id = 1092; 1988.
44. Chamberlayne R, Green B, Barer ML, Hertzman C, Lawrence WJ, Sheps SB. Creating a population-based linked health database: a new resource for health services research. Can J Public Health 1998;89(4):270–3.
45. Kopec JA, Rahman MM, Berthelot JM, Le Petit C, Aghajanian J, Sayre EC, *et al*. Descriptive epidemiology of osteoarthritis in British Columbia, Canada. J Rheumatol 2007;34(2):386–93.
46. Bedson J, Mottram S, Thomas E, Peat G. Knee pain and osteoarthritis in the general population: what influences patients to consult? Fam Pract 2007;24(5):443–53.
47. Flanagan WM, Kopec JA, Finès P, Rahman MM, Sayre EC, Cibere J, *et al*. Estimating prevalence of physician-diagnosed osteoarthritis in Canada using microsimulation (Abstract). Osteoarthritis Cartilage 2008;16(Suppl 4):S153.
48. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, *et al*. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002;40(2):113–28.
49. Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. Qual Life Res 2000;9(8):901–10.
50. Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. Arthritis Rheum 1995;38(8): 1134–41.
51. Wilkins K. Incident arthritis in relation to excess weight. Health Rep 2004;15:39–49.
52. Rahman MM, Aghanjanian J, Kopec JA, Cibere J. Validation of osteoarthritis diagnosis in administrative data using a clinically and radiologically defined population-based cohort of osteoarthritis (Abstract). Osteoarthritis Cartilage 2008;16(Suppl 4):S150.
53. Harrold LR, Yood RA, Andrade SE, Reed JI, Cernieux J, Straus W, *et al*. Evaluating the predictive value of osteoarthritis diagnoses in an administrative database. Arthritis Rheum 2000;43(8):1881–5.
54. Schultz SE, Kopec JA. Impact of chronic conditions. Health Rep 2003; 14(4):41–53.
55. Kopec JA, Finès P, Flanagan WM, Sayre EC, Rahman M, Bansback N, *et al*. Projecting the burden of osteoarthritis in Canada using microsimulation (Abstract). Ann Rheum Dis 2009;68(Suppl 3):99.
56. Finès P, Flanagan W, Kopec JA, Rahman MM, Sayre EC, Cibere J, *et al*. Projecting number of osteoarthritis patients from 2001 to 2021 in Canada using microsimulation (Abstract). Arthritis Rheum 2007; 56(Suppl 9):S341.