

Medical Decision Making

<http://mdm.sagepub.com/>

Dynamic Microsimulation Models for Health Outcomes : A Review

Carolyn M. Rutter, Alan M. Zaslavsky and Eric J. Feuer

Med Decis Making 2011 31: 10 originally published online 18 May 2010

DOI: 10.1177/0272989X10369005

The online version of this article can be found at:

<http://mdm.sagepub.com/content/31/1/10>

Published by:



<http://www.sagepublications.com>

On behalf of:



<http://www.smdm.org>

Additional services and information for *Medical Decision Making* can be found at:

Email Alerts: <http://mdm.sagepub.com/cgi/alerts>

Subscriptions: <http://mdm.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Dynamic Microsimulation Models for Health Outcomes: A Review

Carolyn M. Rutter, PhD, Alan M. Zaslavsky, PhD, Eric J. Feuer, PhD

Background. Microsimulation models (MSMs) for health outcomes simulate individual event histories associated with key components of a disease process; these simulated life histories can be aggregated to estimate population-level effects of treatment on disease outcomes and the comparative effectiveness of treatments. Although MSMs are used to address a wide range of research questions, methodological improvements in MSM approaches have been slowed by the lack of communication among modelers. In addition, there are few resources to guide individuals who may wish to use MSM projections to inform decisions. **Methods.** This article presents an overview of microsimulation modeling, focusing on the development and application of MSMs for health

policy questions. The authors discuss MSM goals, overall components of MSMs, methods for selecting MSM parameters to reproduce observed or expected results (calibration), methods for MSM checking (validation), and issues related to reporting and interpreting MSM findings (sensitivity analyses, reporting of variability, and model transparency). **Conclusions.** MSMs are increasingly being used to provide information to guide health policy decisions. This increased use brings with it the need for both better understanding of MSMs by policy researchers, and continued improvement in methods for developing and applying MSMs. **Key words:** discrete event simulation; decision analytic models. (*Med Decis Making* 2011;31:10–18)

In the late 1950s, Guy Orcutt proposed microsimulation models (MSMs) as a method for exploring social policy questions by simulating the effect of policy on decision-making units (e.g., individuals, families, or corporations).¹ During the 1970s, MSMs were developed to guide US social policy decisions.² Models of traffic flow are used to plan transit projects.^{3,4} MSMs are used in operations research⁵ to describe queuing systems; where they are referred to as discrete event simulations.^{6–8} MSMs for the transmission of infectious diseases such as HIV,^{6,7} influenza,^{8–10} smallpox,¹¹ onchocerciasis,^{12,13} and schistosomiasis¹⁴ are used to examine the effects of intervention and policy change on disease

transmission, including the cost effectiveness of vaccination programs.¹⁵ MSMs for traffic and infectious diseases allow interactions among “agents” (e.g., cars or individuals). Such “agent-based models”^{16,17} are needed when agent interactions are critical to downstream outcomes, for example, when modeling epidemics. To simplify discussion, this article focuses on somewhat simpler MSMs that assume independence across individuals, although the issues raised here are equally relevant to agent-based models.

MSMs were applied to health policy questions as early as 1985, when Habbema and colleagues introduced the “Microsimulation Screening Analysis” (MISCAN) model¹⁸ to examine the impact of cancer screening on morbidity and mortality. MISCAN models have been developed to describe the effects of screening for cervical,¹⁹ breast,²⁰ colorectal,²¹ and prostate cancer.²² In 1994, the Population Health Model was introduced to evaluate costs of diagnosis and treatment of lung and breast cancer,^{23,24} as part of a broader microsimulation effort by Statistics Canada.²⁵ MSMs have also been developed for diabetes,^{26,27} cardiovascular disease,^{28,29} stroke,³⁰ reoperation rates after aortic valve replacement,³¹ organ transplant,^{32,33} osteoporosis,^{34–36} and end-stage liver disease.³⁷ Karnon and colleagues review

Received 29 June 2009 from Biostatistics Unit, Group Health Research Institute, Seattle, WA USA, and Department of Biostatistics, University of Washington School of Public Health and Community Medicine, Seattle, WA USA (CMR); Department of Health Care Policy Harvard Medical School, Boston, MA USA (AMZ); and Statistical Research and Applications Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda MD USA (EJF). Supported by NCI U01 CA97427. Revision accepted for publication 20 February 2010.

Address correspondence to Carolyn M. Rutter, PhD, 1730 Minor Ave, Ste 1600, Seattle, WA 98101; e-mail: rutter.c@ghc.org.

DOI: 10.1177/0272989X10369005

some of the models used to evaluate screening programs.³⁸ Several funding agencies now support the development and application of MSMs, including the National Cancer Institute's Cancer Intervention and Surveillance Modeling Network (CISNET),³⁹ the National Institute of General Medical Sciences' Models of Infectious Disease Agent Study,⁴⁰ and the Robert Wood Johnson Foundation and National Institutes of Health's Childhood Obesity Modeling Network.⁴¹

Many of the MSMs used to examine health policy questions are written by researchers for specific disease processes, an endeavor requiring a high level of programming expertise and a large time commitment. Software programs, such as TreeAge⁴² and Arena,⁴³ allow users to implement discrete event MSMs with relative ease,^{44–47} although there are few publications to guide this work. With the increasing use of MSMs, there is a growing need for both modelers and end users of MSMs to consider issues related to their application. This article provides an overview of MSMs, focusing on model development and applications to health policy questions.

GOALS OF MICROSIMULATION MODELING

MSMs describe events and outcomes at the person-level with an ultimate goal of providing information that can guide policy decisions. (In contrast, mechanistic models, such as those that simulate the behavior of cells, are developed to provide insight into underlying processes. Biological models (discussed in the "Developing a New MSM" section) can incorporate mechanistic models into policy-focused MSMs). Examples of policy-relevant findings from MSMs include overdiagnosis of prostate cancer among prostate-specific antigen (PSA)-detected cases⁴⁸; identification of efficient cervical cancer screening policies⁴⁹; and the impact of modifiable risk factors, screening, and treatment on colorectal cancer (CRC) mortality rate.⁵⁰

MSMs provide policy-relevant information by generating predictions. For example, MSMs can be used to predict trends in disease incidence and mortality under alternative health policy scenarios,^{50,51} or to compare the effectiveness and cost-effectiveness of treatments.⁵² Using MSMs for prediction often requires integration of results across studies, and may require extension of results, for example, extending cross-sectional results to longitudinal-predictions, and extending results to more diverse patient populations or to comparisons not made in

available randomized control trials (RCTs). For example, RCTs demonstrate that fecal occult blood testing (FOBT) leads to reduced CRC mortality,^{53–55} and a case-control study found that flexible sigmoidoscopy was also associated with reduced CRC mortality.⁵⁶ Although there is no direct evidence that either optical colonoscopy or computed tomographic colonography reduce CRC mortality, several studies have estimated their sensitivity and specificity for detecting colorectal adenomas.^{57–59} MSMs for colorectal cancer have combined available information about CRC screening tests to compare the effect and cost-effectiveness of all 4 of these screening modalities.⁶⁰

Thus, the explicit goal of using MSMs for prediction is met by developing models that combine results from randomized controlled trials, epidemiologic studies (e.g., case-control and cohort studies), meta-analyses, and expert opinions. This combining of information becomes an implicit modeling goal and is conducted in conjunction with model calibration.

DEVELOPING A NEW MSM

MSMs have 2 components: a natural history model and an intervention model. Our focus is on the natural history model, which describes the disease process in the absence of intervention. The natural history model may be complex, but once developed it can be combined with different intervention models to answer policy questions. The appropriate level of model complexity depends on the disease process, the scientific questions to be addressed by the model, and data available to inform model parameters. There is a tension between the simplicity of a model and the complexity of the disease process. Simpler models include fewer parameters, making it more likely that model parameters can be estimated using observed data, and are easier to describe, making them more transparent. However, more complex MSMs can be used to address a wider range of scientific questions. One approach is to begin with relatively simple models, extending these as needed to address specific research questions.⁶¹

When considering model complexity, it is useful to distinguish between biological models, which describe the underlying disease processes, and epidemiological models, which focus attention on the observable portion of the disease process.⁶² For example, Luebeck and colleagues⁶³ predicted the effect of folate on colorectal cancer risk using a biological model that describes 3 phases

of carcinogenesis: stem cell initiation (acquisition of 1 or more mutations), cell proliferation, and malignant conversion. Their model allowed folate to both reduce initiation rates and increase proliferation rates, and predicted that early folate initiation (e.g., at age 2) decreases colorectal cancer risk whereas late folate initiation (e.g., at age 65) may increase risk. Although complex biological models can describe the disease process more completely than epidemiological models, they also include parameters that cannot be directly informed by data (such as the impact of folate on cell initiation and proliferation rates).

There are 3 essential steps in developing any MSM: 1) Identifying a fixed number of distinct states and characteristics associated with these states; 2) specifying stochastic rules for transition through states; and 3) setting values for model parameters.

Identify Distinct States

The specification of distinct states included in an MSM depends on the disease process, the research questions of interest, and the availability of data. The number of states included in a model may also depend on how characteristics are attributed to states. For example, an MSM could describe lesion size categorically with different size categories considered different states, or it could model a single lesion state with size treated as a characteristic of the lesion.

Specify Transition Rules

MSMs are either state-transition or continuous-time models. State-transition models allow individuals to move between states at fixed time intervals (possibly dependent on the state), with probabilities specific to this cycle length. In continuous-time MSMs, intervals between transitions have continuous distributions. Both types of MSMs allow transition rules to depend on individual characteristics. MSM transition probabilities may be specified as Markov,⁶⁴ meaning that the probabilities of the next transition depend only on the current state and not on the previous history.⁶⁵ Many models relax this assumption by carrying forward relevant past information into the current state.

Specify Model Parameters

Given the structure of an MSM, it is necessary to specify the values of parameters that determine

transitions through disease states. Parameters associated with observable processes can be estimated directly from biological, clinical, or epidemiological evidence. For example, the distribution survival time following cancer diagnosis may be estimated from Surveillance, Epidemiology and End Reporting (SEER) data,⁶⁶ and other-cause mortality probabilities may be based directly on life tables derived by removing deaths attributable to the disease modeled from Berkeley mortality databases^A. For example, Rosenberg⁶⁷ demonstrates calculation of non-breast-cancer mortality using Berkeley mortality databases with breast cancer deaths estimated using data from the National Center for Health Statistics databases.⁶⁸

When such direct estimation is not possible, model parameters are selected so that the model reproduces observed results, a process called "calibration." Calibration is typically necessary for complex models involving transitions that are only observed through indirect consequences. For example, an observed rate of tumors detected is a function of 2 unobserved processes: the rate of tumor initiation and the growth of tumors to a detectable size. In this case, the functional relationships between MSM parameters and observable calibration data are complex, inasmuch as MSM parameters describe transitions between specific states, which might not all be observed, and data generally describe observed states resulting from a series of transitions. As a result, MSM parameters may be nonidentifiable.⁶⁹ Parameter identifiability is determined by the model in the context of available data. A nonidentifiable parameter cannot be estimated, even with an infinite amount of data. However, the parameter could become identifiable if a new type of data became available. When parameters are unidentifiable, it is still possible to find values that provide good fit to calibration data, but these parameter values are not unique. For example, (observed) rates of detected tumors at screening might depend on (unobservable) rates of initiation and growth to detectable size, with different combinations of these parameters producing the same rate of detection. Whether or not a parameter is identifiable may not be obvious. A parameter associated with an unobservable process can be identifiable if there is sufficient information about the entire process. Complete description of a disease process may necessitate specification of an MSM with nonidentifiable parameters. In this case,

^AThe Berkeley Mortality Database, <http://demog.berkeley.edu/~bmd/>. Additional data are available at this website.

parameters can be identified by adding information to the model, either by setting some parameters to fixed values and carrying out sensitivity analysis (described in the “Sensitivity Analysis” section) or by specifying prior distributions for parameters and carrying out Bayesian calibration (described below).

Early MSMs were calibrated by perturbing parameters one at a time and subjectively judging agreement with available data.⁷⁰ Subjective judgment has largely been replaced by the use of statistics that measure how well the model fits calibration data, and ad hoc parameter perturbation has been replaced by grid search methods.^{71–74} An undirected grid search selects parameter values based on evaluation of model fit at every node in a grid of parameter values or at a random set of values.^{73,75} Undirected searching is based only on MSM results, and therefore can be directly applied to virtually any MSM. However, grid searches are not computationally feasible for highly parameterized models, because the number of grid nodes grows exponentially with the number of model parameters. Furthermore, even a dense grid or a large random sample of parameters might miss regions of good fit.

Directed searches move through the parameter space by “hill climbing,” that is, moving in a direction of improving goodness of fit. (The Nelder Mead simplex algorithm⁷⁶ is an example of a commonly used directed search method.) The direction taken by a directed search is based on the derivative of the likelihood function, which describes the probability of the calibration data given the MSM and a set of parameter values. These derivatives provide information about the rate of change of objective functions, and so direct the algorithm to move in the direction of most rapid increase (“up the hill”). For MSMs, there is usually not a closed form expression for these derivatives, so that directional searches must rely on approximations.⁷⁷ In addition, directed searches may find parameter values that provide locally good fit, but not the best fit across all possible parameter sets (globally good fit). To avoid locally, but not globally, good solutions, directed searches should be initiated at widely dispersed points within the parameter space. Despite these difficulties, directional search methods for MSM calibration are generally more computationally efficient than grid search approaches, requiring fewer runs of the MSM for parameter estimation. Some modelers have focused on model simplification to allow likelihood-based estimation of MSM parameters that parallels usual frequentist or Bayesian estimation approaches.^{74,78–86}

Another challenge associated with calibration is the combination of data from different sources, collected at different times, from different populations. This problem, essentially one of meta-analysis, arises from the implicit MSM goal of integrating results from randomized controlled trials, observational studies, and expert opinion. If study data include information on participant characteristics, it can be incorporated into models to adjust for between-study differences, but such covariate adjustment is not always possible. In addition, some data may carry more weight than other data because of the sample size, the relevance of the population studied, or subjectively perceived data quality. Thus, for example, the modeler might have to assign relative weights to a study of moderate size conducted under current conditions and a much larger study conducted during an earlier period.

SIMULATING POPULATIONS

Once developed, MSMs are used to simulate a hypothetical population with specific characteristics, such as a specific age-sex distribution, or a specific risk factor profile. An MSM can be structured to simulate a population directly, taking distributions of population characteristics at baseline as inputs (e.g., age, gender, and relevant risk factors). Alternatively, cohort models simulate individuals for a relatively narrow age range and specify uniform age distributions within the age range, with hypothetical populations generated by combining multiple simulated cohorts.

MSMs are sometimes used to simulate cohorts rather than representative populations. The choice of the target set (population or cohort) can affect the conclusions drawn from the MSM when costs, effectiveness, or cost effectiveness vary across cohorts.⁸⁷ For example, because cervical cancer screening is less cost-effective for older women, simulations targeted to a cohort of 15–20-year-olds will find that shortening the screening interval from 3 to 2 years is more cost effective than simulations targeted to a population that includes more older women.⁸⁸

MODEL ASSESSMENT

Assessment of a calibrated MSM includes model validation, examination of sensitivity to untestable assumptions, and incorporation of variability.

Model Validation

Model validation is the process of assessing whether a model is consistent with data not used for calibration, a process also called “external validation.” Because validation requires that data be held out of the calibration process, MSMs may not be validated, or validation may be based on simulation of randomized trials that are not directly related to the processes of interest. A related problem is that the models generally cannot be validated against the outcomes of greatest interest, because the models focus on unobserved or unobservable phenomena. In some cases, important outcomes, such as survival, may be set aside and used as validation points.⁸⁹

Comparison of model outputs to calibration data is sometimes called “internal validation.” There is a gray area between internal and external validation that involves using detailed calibration data for model assessment. For example, suppose that calibration data include incidence rates by decade of life, and these rates can be used to internally validate the MSM. If incidence rates were also available by gender and year of life, then further model validation could compare these more detailed rates that were not directly used for calibration.

Sensitivity Analysis

Sensitivity analysis refers to estimation and presentation of model results under various scenarios, often corresponding to varying values of model parameters that are inestimable or poorly estimable. Sensitivity analyses can also provide insight into the impact of specific model assumptions. For example, sensitivity analysis can be used to explore whether adenoma regression, which cannot be directly observed, is plausible by comparing predictions under specific scenarios, for example, “no regression” and “10% of lesions regress.”⁷² Probabilistic sensitivity analysis places distributions on unknown parameters, providing a range of possible results. Parameters are sampled from specified distributions, and multiple MSM runs are used to infer variability in model results that result from variability in model parameters.^{83,90–93} Sensitivity analyses are common, largely because most models include unobservable components.

Between-Model Comparisons

Another type of model assessment compares predictions from different models that are calibrated

using the same data. Such comparative modeling studies allow exploration of uncertainty due to model structure. Examples of this approach include estimation of the combined effects of screening and treatment on breast cancer mortality based on 7 CISNET models for breast cancer⁸⁹ and the Mt. Hood Challenge comparing diabetes models.^{26,27} Each of these groups compared models only after standardizing the calibration data. Without such cooperation to simulate and present results, it can be difficult to directly compare model results. For example, 4 different simulation studies examining the cost-effectiveness of spiral CT for lung cancer screening based on results from the Early Lung Cancer Action Project (ELCAP)^{94,95} reported a wide range of cost-effectiveness, from \$2,500 per life-year gained to \$154,000 per quality adjusted life-year gained.^{96–99} The reasons for these large differences includes differences in the screening frequency examined, lung cancer risk in the hypothetical population, and costs attributed to screening, diagnosis, and treatment. Differences in the underlying assumptions of the microsimulation models is another important source of variability, but the impact of these differences is difficult to determined without a systematic comparison of both model assumptions and model predictions.

Although comparison of results across independently developed models provides an important avenue for addressing variability due to model structure, these comparisons are very time consuming and are only practical for major policy questions. In addition, it can be difficult to determine which models provide the best fit to available data. Bayesian model averaging provides a formal framework for model comparison, which could, in theory, be used to compare different parameterization of a single model; but this approach requires a maximized likelihood for each competing model.¹⁰⁰

Sources of Variability

Sources of MSM variability and uncertainty include inherent variability in the population of interest, variability due to estimation of unknown parameters, selection of calibration data, sampling variability of the selected calibration data, simulation (Monte Carlo) variability, and variability due to model structure assumptions.^{83,101} Bayesian calibration methods parallel Bayesian estimation, and provide interval estimates that describe the variability in both parameter estimates and model predictions due to parameter estimation, sampling variability of

the selected calibration data, and simulation variability.⁸⁵ Repeated model runs that systematically vary parameter values can be used to assess the relationship between parameter variability and variability in model predictions, with findings used to direct model improvement toward reducing variability of those parameters that have the greatest impact on prediction variability (e.g., additional data collection or modifications to the model structure).^{102,103}

To illustrate the use of replication in error analysis for calibration, suppose the target population includes N individuals. Rather than simulating a single population of $1000N$ individuals, more information can be gained by simulating results independently in 1000 samples of size N ; whereas a precise point estimate is obtained as the mean of the 1000 simulations, an interval estimate can be obtained by calculating the variability in model results across simulations.

Although these methods can be used to summarize and explain uncertainty in model predictions, because of computational and conceptual limitations, MSM results are routinely provided without measures of precision. Further development of computationally efficient methods to estimate the uncertainty in MSM model estimates and predictions is an important area of research.

MSM TRANSPARENCY

Model transparency refers to the ability to convey model assumptions. There are divergent views on what transparency means and what level of transparency is optimal.^{104,105} Different levels of disclosure can be consistent with model transparency, such as describing MSM assumptions (including relative costs), providing algorithms used in the MSM, providing equations used to program the MSM, and, ultimately, releasing the computer code underlying the model. Although release of computer code is seemingly the most transparent approach, this strategy is time consuming and ultimately uninformative to the vast majority of end users so that code release may obscure rather than clarify the model. Although transparency often refers to model structure (assumptions, algorithms, equations), it also should incorporate the data used to calibrate or estimate model parameters, goodness of fit to calibration data, including fit to subgroups of interest, and validation results.

METHODOLOGICAL REVIEW

MSM transparency can be difficult to achieve. In experimental or observational research, transparency is often understood as the hypothetical ability to repeat an experiment or analysis. Yet many published articles use complex data analytic models that cannot be completely described within page limits, so that even these analyses are not completely transparent. Furthermore, there are very few venues for complete model description and review, which is necessary for complete evaluation of the model. Technical appendices, including supplemental material published online, can be critical. For example, the CISNET modeling group provides online model profiles.¹⁰⁶ Finally, there are disincentives to fully transparent models. Models take years to develop, resulting in hesitation to fully disclose this “intellectual property.” There are explicit financial disincentives for making proprietary models fully transparent. Despite these difficulties, model transparency is necessary if MSMs are used to inform policy decisions.

CONCLUSIONS

Microsimulation models combine expert opinion with observational and experimental results, providing a relatively inexpensive way to estimate population-level effects (including costs and benefits) of interventions, policy changes, or shifts in risk factors. Microsimulation modeling is beginning to coalesce as a defined area of expertise, as evidenced by publication of modeling guidelines,^{107,108} funding of microsimulation modeling efforts by NCI and NIH, microsimulation-focused conferences such as the Conference of the International Microsimulation Association, sponsored by Statistics Canada¹⁰⁹ and the General Conference of the International Microsimulation Association,¹¹⁰ and the recently launched *International Journal of Microsimulation*. Given this interest, we anticipate both an increased use of microsimulation models to address research questions and parallel improvements in methods used to develop and apply microsimulation models.

REFERENCES

1. Orcutt GH. A new type of socio-economic system. *Review of Economics and Statistics*. 1957;80:1081–100.
2. National Research Council (US). Panel to Evaluate Microsimulation Models for Social Welfare Programs. Improving information for social policy decisions: the uses of microsimulation modeling. In: Citro CF, Hanushek EA, eds. *Review and*

- Recommendations. Vol I. Washington, DC: National Academy Press; 1991.
3. Lemp JD, McWethy LB, Kockelman KM. From aggregate methods to microsimulation—assessing benefits of microscopic activity-based models of travel demand. *Transportation Research Record*. 2007;1994:80–8.
 4. Hoogendoorn SP, Bovy PHL. State-of-the-art of vehicular traffic flow modelling. *Proc Inst Mech Eng*. 2001;215(14):283–303.
 5. Hollocks BW. Forty years of discrete event simulation—a personal reflection. *Journal of the Operational Research Society*. 2006;57:1383–99.
 6. Cassels S, Clark SJ, Morris M. Mathematical models for HIV transmission dynamics: tools for social and behavioral science research. *J Acquir Immune Defic Syndr*. 2008;47(suppl 1):S34–9.
 7. Castiglione F, Pappalardo F, Bernaschi M, Motta S. Optimization of HAART with genetic algorithms and agent-based models of HIV infection. *Bioinformatics*. 2007;23(24):3350–5.
 8. Ackerman E, Longini IM Jr, Seaholm SK, Hedin AS. Simulation of mechanisms of viral interference in influenza. *Int J Epidemiol*. 1990;19(2):444–54.
 9. Ackerman E, Zhuo Z, Altmann M, et al. Simulation of stochastic micropopulation models—I. The SUMMERS simulation shell. *Comput Biol Med*. 1993;23(3):177–98.
 10. Peterson D, Gatewood L, Zhuo Z, Yang JJ, Seaholm S, Ackerman E. Simulation of stochastic micropopulation models—II. VESPERs: epidemiological model implementations for spread of viral infections. *Comput Biol Med*. 1993;23(3):199–213.
 11. Longini IM Jr, Halloran ME, Nizam A, et al. Containing a large bioterrorist smallpox attack: a computer simulation approach. *Int J Infect Dis*. 2007;11(2):98–108.
 12. Plaisier AP, van Oortmarssen GJ, Habbema JD, Remme J, Alley ES. ONCHOSIM: a model and computer simulation program for the transmission and control of onchocerciasis. *Comput Methods Programs Biomed*. 1990;31(1):43–56.
 13. Habbema JD, Alley ES, Plaisier AP, van Oortmarssen GJ, Remme JH. Epidemiological modelling for onchocerciasis control. *Parasitol Today*. 1992;8(3):99–103.
 14. Habbema JD, De Vlas SJ, Plaisier AP, Van Oortmarssen GJ. The microsimulation approach to epidemiologic modeling of helminthic infections, with special reference to schistosomiasis. *Am J Trop Med Hyg*. 1996;55(5 suppl):165–9.
 15. Kim SY, Goldie SJ. Cost-effectiveness analyses of vaccination programmes: a focused review of modelling approaches. *Pharmacoeconomics*. 2008;26(3):191–215.
 16. Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. *Proc Natl Acad Sci U S A*. 14 2002;99(suppl 3):7280–7.
 17. Tesfatsion L, Judd KL. *Handbook of Computational Economics*. Vol 2. North-Holland (Amsterdam): Elsevier; 2006.
 18. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Comput Methods Programs Biomed*. 1985;20(1):79–93.
 19. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. Model building on the basis of Dutch cervical cancer screening data. *Maturitas*. 1985;7(1):11–20.
 20. van Oortmarssen GJ, Habbema JD, van der Maas PJ, et al. A model for breast cancer screening. *Cancer*. 1990;66(7):1601–12.
 21. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. *Comput Biomed Res*. 1999;32(1):13–33.
 22. Draisma G, Boer R, Otto SJ, et al. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst*. 2003;95(12):868–78.
 23. Wolfson MC. POHEM—a framework for understanding and modelling the health of human populations. *World Health Stat Q*. 1994;47(3–4):157–76.
 24. Will BP, Berthelot JM, 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.
 25. Microsimulation. 2009. Available from: <http://www.statcan.ca/english/spsd/>. Accessed June 12, 2009.
 26. Brown JB, Palmer AJ, Bisgaard P. The Mount Hood Challenge: cross-testing two diabetes simulation models. *Diabetes Care*. 2000;50(suppl 3):S57–64.
 27. Group TMHM. Computer modeling of diabetes and its complication: a report on the 4th Mount Hood Challenge Meeting. *Diabetes Care*. 2007;30:1638–46.
 28. Kuntz KM, Tsevat J, Goldman L, Weinstein MC. Cost-effectiveness of routine coronary angiography after acute myocardial infarction. *Circulation*. 1996;94(5):957–65.
 29. Smolen HJ, Cohen DJ, Samsa GP, et al. Development, validation, and application of a microsimulation model to predict stroke and mortality in medically managed asymptomatic patients with significant carotid artery stenosis. *Value Health*. 2007;10(6):489–97.
 30. Matchar DB, Samsa GP, Matthews JR, et al. The Stroke Prevention Policy Model: linking evidence and clinical decisions. *Ann Intern Med*. 1997;127(8 pt 2):704–11.
 31. Puvimanasinghe JP, Takkenberg JJ, Edwards MB, et al. Comparison of outcomes after aortic valve replacement with a mechanical valve or a bioprosthesis using microsimulation. *Heart*. 2004;90(10):1172–8.
 32. Ouwens JP, Groen H, TenVergert EM, Koeter GH, de Boer WJ, van der Bij W. Simulated waiting list prioritization for equitable allocation of donor lungs. *J Heart Lung Transplant*. 2002;21(7):797–803.
 33. Groen H, van der Bij W, Koeter GH, TenVergert EM. Cost-effectiveness of lung transplantation in relation to type of end-stage pulmonary disease. *Am J Transplant*. 2004;4(7):1155–62.
 34. Vanness DJ, Tosteson AN, Gabriel SE, Melton LJ 3rd. The need for microsimulation to evaluate osteoporosis interventions. *Osteoporos Int*. 2005;16(4):353–8.
 35. Liu H, Michaud K, Nayak S, Karpf DB, Owens DK, Garber AM. The cost-effectiveness of therapy with teriparatide and alendronate in women with severe osteoporosis. *Arch Intern Med*. 2006;166(11):1209–17.
 36. Schousboe JT, Taylor BC, Fink HA, et al. Cost-effectiveness of bone densitometry followed by treatment of osteoporosis in older men. *JAMA*. 2007;298(6):629–37.

37. Alagoz O, Bryce CL, Shechter S, et al. Incorporating biological natural history in simulation models: empirical estimates of the progression of end-stage liver disease. *Med Decis Making*. 2005;25(6):620–32.
38. Karnon J, Goyder E, Tappenden P, et al. A review and critique of modelling in prioritising and designing screening programmes. *Health Technol Assess*. 2007;11(52):iii–iv, ix–xi, 1–145.
39. National Cancer Institute. Cancer Intervention and Surveillance Modeling Network (CISNET). 2008. Available from: <http://cisnet.cancer.gov/>. Accessed on 9 April 2010.
40. Models of Infectious Disease Agent Study, MIDAS. 2009. Available from: <http://www.nigms.nih.gov/Initiatives/MIDAS/>. Accessed June 12, 2009.
41. COMnet. 2009; <http://obesitymodeling.net/beta/>. Accessed June 12, 2009.
42. TreeAge [computer program]. Williamstown (MA): TreeAge Software, Inc.; 2008.
43. Kelton DW, Sadowski RP, Sturrock DT. *Simulation with Arena*. 4th ed. Boston: McGraw-Hill; 2007.
44. McEwan P, Williams JE, Griffiths JD, et al. Evaluating the performance of the Framingham risk equations in a population with diabetes. *Diabet Med*. 2004;21(4):318–23.
45. Bellsey J, Sheldon AM. Healthcare Applications of Decision Analytic Modeling. *Health Economics in Prevention and Care*. 2000;1:37–43.
46. Hernandez R, Vale L. The value of myocardial perfusion scintigraphy in the diagnosis and management of angina and myocardial infarction: a probabilistic economic analysis. *Med Decis Making*. 2007;27(6):772–88.
47. Delaney G, Jacob S, Featherstone C, Barton M. The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer*. 2005;104(6):1129–37.
48. Etzioni R, Penson DF, Legler JM, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst*. 2002;94(13):981–90.
49. van der Akker-van Marle ME, van Ballegooijen M, van Ootmarssen GJ, Boer R, Habbema JDF. Cost-effectiveness of cervical cancer screening: comparison of screening policies. *J Natl Cancer Inst*. 2002;94:193–204.
50. Vogelaar I, Van Ballegooijen M, Schrag D, et al. How much can current interventions reduce colorectal cancer mortality in the U.S.? *Cancer*. 2006;107:1623–33.
51. National Cancer Institute. Colorectal Cancer Mortality Projections. Dec 2007. Available from: <http://cisnet.cancer.gov/projections/colorectal/>. Accessed 9 April 2010.
52. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM. Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2008;149(9):659–69.
53. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet*. 1996;348(9040):1472–7.
54. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet*. 1996;348(9040):1467–71.
55. Towler B, Irwig L, Glasziou P, Kewenter J, Weller D, Silagy C. A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, hemoccult. *BMJ*. 1998;317:559–65.
56. Selby JV, Friedman GD, Quesenberry CP Jr, Weiss NS. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med*. 1992;326(10):653–7.
57. Hixson LJ, Fennerty MB, Sampliner RE, McGee D, Garewal H. Prospective study of the frequency and size distribution of polyps missed by colonoscopy. *J Natl Cancer Inst*. 1990;82(22):1769–72.
58. Rex DK, Cutler CS, Lemmel GT, et al. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology*. 1997;112(1):24–8.
59. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. *N Engl J Med*. 2008;359(12):1207–17.
60. HHS.gov. Centers for Medicare & Medicaid Services (CMS). 2008. Available from: <http://www.cms.hhs.gov/mcd/viewmccac.asp?from2=viewmccac.asp&where=index&mid=45&> Strike 2008. Accessed 9 April 2010.
61. Lee RC, Donaldson C, Cook LS. The need for evolution in healthcare decision modeling. *Med Care*. 2003;41(9):1024–33.
62. 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.
63. Luebeck EG, Moolgavkar SH, Liu AY, Boynton A, Ulrich CM. Does folic acid supplementation prevent or promote colorectal cancer? Results from model-based predictions. *Cancer Epidemiol Biomarkers Prev*. 2008;17(6):1360–7.
64. Cox DR, Miller HD. *The Theory of Stochastic Processes*. London: Chapman and Hall Ltd.; 1965.
65. Beck J, Pauker S. The Markov process in medical prognosis. *Med Decis Making*. 1983;3:419–58.
66. National Cancer Institute. Surveillance, Epidemiology and End Results—SEER 9 Regs Public Use. 2004. Available from: Strike 2004. Accessed 9 April 2010.
67. Rosenberg MA. Competing risks to breast cancer mortality. *J Natl Cancer Inst Monogr*. 2006(36):15–9.
68. National Center for Health Statistics. US Life Tables. 2000. Available from: www.cdc.gov/nchs/products/life_tables.htm, Strike 1960–2000. Accessed 9 April 2010.
69. Bickel PJ, Doksum KA. *Mathematical Statistics: Basic Ideas and Selected Topics*. New York: Olden-Day; 1977.
70. Ramsey SD, McIntosh M, Etzioni R, Urban N. Simulation modeling of outcomes and cost effectiveness. *Hematol Oncol Clin North Am*. 2000;14(4):925–38.
71. Ness RM, Holmes AM, Klein R, Dittus R. Cost-utility of one-time colonoscopic screening for colorectal cancer at various ages. *Am J Gastroenterol*. 2000;95(7):1800–11.
72. Loeve F, Boer R, Zauber AG, et al. National Polyp Study data: evidence for regression of adenomas. *Int J Cancer*. 2004;111(4):633–9.
73. Salomon JA, Weinstein MC, Hammitt JK, Goldie SJ. Empirically calibrated model of hepatitis C virus infection in the United States. *Am J Epidemiol*. 2002;156(8):761–73.

74. Chia YL, Salzman P, Plevritis SK, Glynn PW. Simulation-based parameter estimation for complex models: a breast cancer natural history modelling illustration. *Stat Methods Med Res.* 2004;13(6):507–24.
75. Kim JJ, Kuntz KM, Stout NK, et al. Multiparameter calibration of a natural history model of cervical cancer. *Am J Epidemiol.* 2007;166(2):137–50.
76. Press BP, Teukolsky SA, Vetterby WT, Flannery BP. *Numerous Recipes in Fortran—The Art of Scientific Computing.* New York: Cambridge University Press; 1992.
77. Tan SY, van Oortmarrsen GJ. Estimating parameters of a microsimulation model for breast cancer screening using the score function method. *Annals of Operations Research.* 2003;119:43–61.
78. Lee SJ, Zelen M. Scheduling periodic examinations for the early detection of disease: applications to breast cancer. *J Am Stat Assoc.* 1998;93:1271–81.
79. Shen Y, Zelen M. Parametric estimation procedures for screening programmes: stable and nonstable disease models for multimodality case finding. *Biometrika.* 1999;86(3):503–15.
80. Shen Y, Zelen M. Robust modeling in screening studies: estimation of sensitivity and preclinical sojourn time distribution. *Biostatistics.* 2005;6(4):604–14.
81. Tsodikov A, Szabo A, Wegelin J. A population model of prostate cancer incidence. *Stat Med.* 2006;25(16):2846–66.
82. Flehinger BJ, Kimmel M. The natural history of lung cancer in a periodically screened population. *Biometrics.* 1987;43(1):127–44.
83. Parmigiani G. Measuring uncertainty in complex decision analysis models. *Stat Methods Med Res.* 2002;11(6):513–37.
84. Berry DA, Inoue L, Shen Y, et al. Modeling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. *J Natl Cancer Inst Monogr.* 2006(36):30–6.
85. Rutter CM, Miglioretti DL, Savarino JE. Bayesian calibration of microsimulation models. *J Am Stat Assoc.* 2009;104(488):1338–50.
86. Craig BA, Fryback DG, Klein R, Klein BE. A Bayesian approach to modelling the natural history of a chronic condition from observations with intervention. *Stat Med.* 1999;18(11):1355–71.
87. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine.* New York: Oxford University Press; 1996.
88. Dewilde S, Anderson R. The cost-effectiveness of screening programs using single and multiple birth cohort simulations: a comparison using a model of cervical cancer. *Med Decis Making.* 2004;24(5):486–92.
89. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med.* 2005;353(17):1784–92.
90. Cronin KA, Legler JM, Etzioni RD. Assessing uncertainty in microsimulation modelling with application to cancer screening interventions. *Stat Med.* 1998;17(21):2509–23.
91. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. *Med Decis Making.* 1985;5(2):157–77.
92. Critchfield GC, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. *Med Decis Making.* 1986;6(2):85–92.
93. Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: necessary and required? *Value Health.* 2006;9(4):244–52.
94. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet.* 1999;354(9173):99–105.
95. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early lung cancer action project: a summary of the findings on baseline screening. *Oncologist.* 2001;6(2):147–52.
96. Wisnivesky JP, Mushlin AI, Sichertman N, Henschke C. The cost-effectiveness of low-dose CT screening for lung cancer: preliminary results of baseline screening. *Chest.* 2003;124(2):614–21.
97. Marshall D, Simpson KN, Earle CC, Chu CW. Economic decision analysis model of screening for lung cancer. *Eur J Cancer.* 2001;37(14):1759–67.
98. Chirikos TN, Hazelton T, Tockman M, Clark R. Screening for lung cancer with CT. *Chest.* 2002;121:1507–14.
99. Mahadevia PJ, Fleisher LA, Frick KD, Eng J, Goodman SN, Powe NR. Lung cancer screening with helical computed tomography in older adult smokers: a decision and cost-effectiveness analysis. *JAMA.* 2003;289(3):313–22.
100. Jackson CH, Thompson SG, Sharples LD. Accounting for uncertainty in health economic decision models by using model averaging. *Journal of the Royal Statistical Society, Series A: Statistics in Society.* 2009;172:282–404.
101. Parmigiani G, Ancukiewicz M, Matchar DB. Decision models in clinical recommendations development: the Stroke Prevention Policy Model. In: Berry DA, Stangl DK, eds. *Bayesian Biostatistics.* Vol 151. New York: Marcel Dekker; 1996. p 207–33.
102. Thurston SW, Zaslavsky AM. Variance estimation in microsimulation models of the Food Stamp Program, 1996 Proceedings of the Social Statistics Section, American Statistical Association, pp 4–9, August 4–8, Chicago, IL.
103. Zaslavsky AM, Thurston SW. Error analysis of food stamp microsimulation models: Further results, 1995 Proceedings of the Social Statistics Section, American Statistical Association, pp 151–156, August 13–17, Orlando, FL.
104. Eddy DM. Accuracy versus transparency in pharmacoeconomic modelling: finding the right balance. *Pharmacoeconomics.* 2006;24(9):837–44.
105. Fendrick AM. The future of health economic modeling: have we gone too far or not far enough? *Value Health.* 2006;9(3):179–80.
106. CISNET. 2009. Available from: <http://cisnet.cancer.gov>. Accessed June 12, 2009.
107. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—Modeling Studies. *Value Health.* 2003;6(1):9–17.
108. American Diabetes Association Consensus Panel. Guidelines for computer modeling of diabetes and its complication. *Diabetes Care.* 2004;27(9):2262–5.
109. Statistics Canada. 2009. Available from: <http://www.statcan.gc.ca/start-debut-eng.html>. Accessed June 12, 2009.
110. International Microsimulation Association. 2009. Available from: www.microsimulation.org. Accessed June 12, 2009.