Critical Review of Stochastic Simulation Literature and Applications for Health Actuaries

Appendix A

September, 2007

Article

Author	Babad H, Sanderson C, Naidoo B, White I, Wang D
Title	**The Development of a Simulation Model of Primary Prevention Strategies for Coronary Heart Disease**
Source	Health Care Management Science, 2002. 5:269-274.

Context

Description

The model is a discrete-event micro-simulation model designed to assess the impacts of various primary prevention strategies on use of health care resources. Hypothetical individuals are assigned risk factor profiles, sampled from a set of probability distributions based on coronary heart disease (CHD) data from the Framingham cohort study. A set of transition events are modeled, estimating time to a specified set of disease events. The simulation described in this paper is one component of a larger simulation effort that simulates not only the impact of primary prevention of CHD on health care resources, but also the impact of treatment once individuals are diagnosed with CHD. This paper addresses only the primary prevention component. The model described in this paper was still under development at the time of the publication.

The populations were derived from the Health Survey of England, and could be modeled as the entire English population or particular demographic sub-groups.

Outcome of Interest

The ultimate outcome of interest appears to be health care costs (from whose perspective costs are to be assessed is not specified, though the impression one gets from reading the article is that it would be the societal perspective that was of primary interest). Outputs from the model include annual counts and mean values of numbers in each of a set of predefined disease states, as well as prevalence of risk factors. These outputs will apparently be linked to models of health care costs that are apparently being developed in parallel with the work described here.

Model

Type

This is a state-transition model with random (stochastic) sampling of values for most model parameters. It appears to be a Monte Carlo, Markov Chain Microsimulation – MCMC Microsim – modeling the transition of events of hypothetical individuals within hypothetical

populations.

Software

Model was programmed using the PASCAL language, using a set of simulation routines called POST (Patient Oriented Simulation Techniques) developed specifically for modeling health care systems, apparently by collaborators of the paper authors. At least one paper and one book exist (cited in the paper) that describe the POST routines.

Model Quality

Data Sources

There are three main data sources used in this simulation. (1) The HSE – Health Survey of England, is used to characterize a population. (2) The Framingham Study is used to characterize the natural history of CHD ("…used to derive the time-to-disease-event distributions conditional on the attributes of the individual concerned…" p. 270). (3) The British Regional Heart Study is used to "calibrate" the Framingham data to the English situation, and to provide a "check" on the results derived from the Framingham study.

Parameters

The basic model parameters from the Health Survey of England were: Population size, demographic characteristics of population, simulation length in years.

The natural history of CHD parameters, from Framingham cohort study, were: onset of stable angina, onset of unstable angina, myocardial infarction, sudden cardiac death, stroke death, other CVD death, cancer death, other unknown death.

The baseline intervention parameters (source of these parameter values is not clear from the paper) were: smoking intervention, threshold blood pressure treatment, threshold cholesterol treatment, non-compliance blood pressure, non-compliance cholesterol.

The intervention parameters to be changed/evaluated (again, the source for setting these parameters is not specified in the text)were: type of program and program target – e.g. for reducing blood pressure or cholesterol, either drug therapy or advice on diet/lifestyle change can be specified; Stochastic parameters for treatment uptake rate, treatment delay, non-compliance rates, and treatment effectiveness.

Virtually all parameters in the model appear to be stochastic, in that their starting values, as well as change values at each annual cycle are drawn randomly from probability distributions.

Simplifying Assumptions	Effectiveness of primary prevention treatment is expressed in the model as a change in risk factors which assumes that the scale of the intervention effect can be estimated from the resulting change in risk factors. Limitations of the Framingham and HSE data mean that CHD risk factors in the model are limited to age, sex, systolic blood pressure, total cholesterol and smoking. Exclusions from the model create implicit assumptions that diet, physical activity, diastolic blood pressure, alcohol consumption, etc. do not impact CHD.
Duration / Time Perspective	Simulation length in years is one parameter that the model user specifies – so it can vary from run to run.
Iterations per Scenario	Number of model runs is a variable that is specified by the model user.
Validation	Model was still under development at time of publication but further testing, validation and further peer review were planned.
Sensitivity Analysis	Not applicable – model is not fully developed.

Evaluation

Strengths/ Weaknesses	This is a fairly detailed simulation that required substantially more than could have been produced by simple application of common sense assumptions. Its potential usefulness is limited primarily by the data sources used to construct the model. A substantial strength of this model is that it doesn't stop at the simulation of the natural history of CHD, but also allows the user to simulate the interaction of various primary CHD prevention interventions with the natural history of CHD and the risk factor profiles of individuals to alter the outcomes for a hypothetical population.
Presentation of Results	Few actual results were presented. However, flowcharts were included describing the population and primary CHD prevention model as well as a very high level description of the simulation logic structure. One graph was shown – displaying the simulated age structure at first CHD event (systolic blood pressure in the hypertensive range) for a hypothetical cohort of 5000 females age 45 at model start, with total cholesterol of 240, who smoked 10 cigarettes per day.
Interpretation of Results	No interpretation of results is provided – just description of the core components of the model itself and what it could produce in terms of output.

Value to Decision Making	To the extent that this model allows for simulation of the potential impact on both health care costs and health benefits of implementing hypothetical primary prevention interventions to prevent CHD, it has the potential to be a very valuable decision support tool for policy makers or others. It could certainly be used to support policy changes. The model itself doesn't appear to directly lend itself to identifying areas for future research, but might easily lead users to see important questions that future research could/should address.
Ease of Implementation	In its current state, this is not an "off the shelf" model that could be used by others than those who have been developing it. Discussion section of the paper suggests that it is ultimately the intention of the model developers to make this model available to an audience of users (possibly two – one of whom does not have experience working with simulation models and one of which does).
Further Reading	None.

Article

Author	Buchanan JL, Keeler EB, Rolph JE, Holmer MR
Title	**Simulating Health Expenditures Under Alternative Insurance Plans**
Source	Management Science, 1991. 37(9):1067-1090

Context

Description

This was a simulation of individual and family health care expenditures for a one year period. The article reported results for a sample of 970 families, 2297 individuals. Twenty-eight insurance benefit designs were modeled, a free plan and plans with coinsurance (25%, 50%, 100%) and maximum dollar expenditure (9 values from $50 to $3000 and no limit). Later simulations also included deductible plans. Individual and family expenditures were generated by occurrence of episode of medical care and care-seeking for the episode; four different types of treatment episodes were modeled: hospitalization, preventive services, chronic disease care, and acute care. Care-seeking was influenced by the current level cost-sharing (deductible, coinsurance, and out-of-pocket maximum), which changes throughout the year as expenditures were realized. Expenditures were measured as per capita (out-of-pocket and total) and population aggregate health care expenditures.

Outcome of Interest

The main outcomes were total family and individual health expenditures and related out-of-pocket expenditures; value of services and risk premiums were also estimated. All outcomes were compared across insurance benefit plans. The value of services was used to measure the wasted resources that result from moral hazard -- overuse of services because the services were obtained at less than full price. The risk premium was an estimate of additional risk faced by individuals and families because they were uncertain about their level of out-of-pocket expenditures.

Model

Type

The authors do not state the model type, but it is described as microsimulation with stochastic random variables for: coinsurance rate (a function of the expenditure up to time *t*), individual and family propensity for episodes of various types, an individual's number of episodes of various types under the free care plan, total expenditure for an episode under the free care plan, total expenditures for an individual and family at time *t*, and a censoring variable to identify free plan episodes also treated under cost-sharing plans. Statistical distributions

	used to represent model elements were selected to best represent the data generating process. For example, the predicted number of episodes per individual is based on a count regression model with parameters for individual characteristics and the predicted mean natural logarithm and standard deviation of cost per episode for each of the 4 types of episodes was estimated using a log cost regression model. The article provides detail on all formulas underlying the simulation.
Software	The type of software was not described.

Model Quality

Data Sources	The authors used data from the RAND Health Insurance Experiment to estimate distribution parameters. For example, the cost per episode assumed a lognormal distribution with mean and standard deviation estimated from the HIE data. Four episode types were modeled: hospital, preventive services, chronic care services, and acute services. For the simulation run described in the article, a population distribution was based on the March 1984 Current Population Survey.
Parameters	Input parameter estimates and regression model predictions were from the HIE data. Parameters included: number of episodes of various type for each individual, mean cost for episode of each type, standard error of episode cost, probability that individual with coinsurance seeks care for an episode that would have been treated under the free care plan, and gamma distribution shape and scale parameters for episode size and unmeasured individual and family propensity. The authors included a number of interdependencies in order to make the simulation as realistic as possible.
Simplifying Assumptions	The authors did not include dental services because most health insurance does not cover dental.
Duration / Time Perspective	Each simulation run modeled one year of health expenditure.
Iterations per Scenario	Each scenario was generated by a single iteration for 970 families, 2297 individuals.
Validation	The model was validated by comparing to the HIE results from the Seattle site.
Sensitivity Analysis	The authors did not comment on sensitivity analysis.

Evaluation

Strengths/ Weaknesses

The simulation model included behavioral assumptions to allow demand for health care services to be sensitive to the "price," where price was defined as the consumer portion of payment for the episode. Price was allowed to vary throughout the insurance year as the individual and family reached its maximum dollar expenditure. For example, at the beginning of the insurance year, a 25% coinsurance determined the price for a $1000 episode as $250. But if a $500 maximum dollar expenditure was reached mid-way through the year, the price became $0, regardless of episode cost.

The simulation model included relationships between changing price through the simulation year and the impact of price on care-seeking for a medical episode.

The article describes in great detail the variables in the model, sources of parameters, distributions assumed.

Presentation of Results

The authors selected a few insurance designs and presented per capita out-of-pocket and total spending for the cost-sharing plans compared to free care as a function of the maximum dollar expenditure. The authors presented a number of graphs, which were difficult to interpret without a careful description.

Interpretation of Results

In the authors' discussion section, they conclude that: 1) even small deductibles curb demand, 2) maximum dollar expenditure (MDE) limits or $1,000-$2,000 per person make good economic sense, 3) individual MDEs seem better than family MDEs, 4) actuaries and insurance companies do quite will in offering desirable policies by the criteria used here, 5) insurance experts are very good at generating premiums for plans with which they have experience, but don't seem as able to price new plans.

They recommend that the simulation model is most useful when estimating plan differences for insurance plans with various benefit designs and for comparing expenditures across subpopulations within plans.

Value to Decision Making

The simulation model captures many distributions underlying demand for medical: care-seeking for episodes, types of episodes, correlation among family members, and correlation among types of episodes. The simulation recognizes relationships between out-of-pocket "price" for medical care and therefore results in realistic simulation of demand.

Ease of Implementation

Many parameters and levels of detail were needed for model development. The data used to develop model is old and delivery and medical care treatments have changed considerably since the 1970's. However, underlying economic principles are constant and the relationship between family spending and decisions on seeking medical care are similar. The article discusses valuable methods for understanding demand which are highly applicable to current environment with shift to high deductible / health savings account forms of financing medical care.

Actuaries now have the advantage of using commercially available episode of care groupers which use diagnosis and procedure codes to clump claim data into meaningful treatment episodes. Using such data as a base, actuaries could develop distributions of episodes which resulted in care-seeking by individuals, and families if the family unit was covered under a single insurance carrier. Models developed with such data allow testing of various insurance benefit designs that do not currently have enough experience to be successfully modeled using cell-based aggregate approaches.

Further Reading

None.

Article

Author	Caro, J.J., Caro, G., Getsios, D., Raggio, G., Burrows, M., Black, L.
Title	**The Migraine ACE Model: Evaluating the Impact on Time Lost and Medical Resource Use**
Source	*Headache*, 2000. Volume 40:282-291

Context

Description

The migraine adaptive cost-effectiveness model (migraine ACE) is a simulation model developed to assess the impact of different migraine treatments on both the direct and indirect costs associated with migraine.

Direct costs are assessed in terms of migraine medication use, as well as physician visits, emergency room visits and hospitalizations for treatment of migraine headaches. Indirect costs are assessed in terms of time lost from both work and non-work activities.

The model is developed using Canadian data and costs are expressed in terms of Canadian dollars.

Ostensibly, this model was developed to compare the economic impact of new migraine therapies (e.g. Sumatriptan® – Imitrex®) to "customary" therapies (acetaminophen, acetaminophen with codeine, Fiorinal, Fiorinal with codeine). However, in this paper, only a single model is run, using drug costs associated only with "customary" therapies prior to the introduction of triptan therapies. No results are presented that would allow for comparison of these therapies with newer therapies. The paper is presented "to illustrate the results that the model can produce…" (p. 283)

Outcome of Interest

Duration of migraine symptoms; indirect costs (work time lost; unpaid work time lost); and direct medical costs (physician visits; ER visits; hospitalizations; and drug therapy).

The authors indicate that the model allows for analyses from number of different perspectives (e.g. Employers; healthcare payers; patients; societal). In the simulations presented in this paper, the authors attempt to approximate a societal perspective, but acknowledge that they are not able to measure all costs that would be required for a true societal perspective model.

Model

Type

This model is best described as a static microsimulation model. The only detail given about the model itself is that Monte Carlo techniques are used – presumably to sample from known frequency distributions for various individual characteristics in assigning characteristics to the simulated population.

Software

No information is given in the paper about software used to develop or implement this model.

Model Quality

Data Sources

A primary data source used in developing the migraine-specific aspects of the model is something the authors refer to as the Q24 study. This was a five country study of 749 adults with moderate to severe migraines, treated initially for a 12 week period with customary treatments, followed by 24 weeks of treatment with sumatriptan. This study provided several quality of life and clinical outcomes, but apparently did not include cost outcomes. (See reference in further reading below). These data were used to specify sex-specific migraine frequency, severity, duration of symptoms, time lost from work and non-work activities, medical services use, as well as pharmaceutical use in migraine patients (both prescription & non-prescription).

Age and sex-specific migraine prevalence rates are specified in the model based on data from the Canadian Migraine Prevalence Study, combined with Canadian demographic statistics from Statistics Canada.

Proportions employed full-time, part-time, and not employed were specified for age and sex-specific strata, and hourly wages rates for full-time employees were specified – all based on data from Statistics Canada.

Parameters

A population of 10,000 simulated persons is defined in terms of age, sex, employment status (full-time, part-time, not employed) and average hourly wage, based on the demographic profile of patients with migraine (population from which this demographic profile is derived is not defined).

Sex-specific, daily probability of migraine symptoms is calculated (calculations not specified). The number of migraine days per year is set for each simulated individual through Monte Carlo sampling from data-based probability distributions for migraine. Severity of migraine symptoms is established for individuals on days with symptoms by sampling from a sex-specific probability distribution (categories of

	mild, moderate, or severe). These model settings are then used to compute daily parameters of symptom duration, time losses with specified costs assigned (depending on assigned employment status – lost time from work and non-work activities), costs of pharmaceutical treatment, physician visit, ER visit, and hospitalization. Again, the paper is not crystal clear about this, but it appears that most of these parameters were established through developing probability distributions based on real-life data, on which Monte Carlo samples were then run. Daily experience is then aggregated for all individuals across a year of simulated experience. Although not demonstrated in this paper, the authors indicate that aggregation of simulated time can also be conducted over different population sub-groups if that is of interest (e.g. employed women ages 25-45 years). Unpaid work activities (domestic work, caregiving to household members, shopping, transportation, etc.) was assigned a value of $11.41/hr, based on estimate from Statistics Canada.
Simplifying Assumptions	Wage rates for part-time employees were simply assumed to be 20% lower than those for the full-time employed. Leisure time was not assigned a monetary value in the model.
Duration / Time Perspective	The model projects the migraine experience of a specified population over the course of one year.
Iterations per Scenario	The authors do not describe the number of iterations per scenario, and only one scenario is presented – presumably representing a single model iteration.
Validation	Validation was not discussed in the paper.
Sensitivity Analysis	The authors indicate that extensive sensitivity analyses were performed on the valuation of each type of time lost. But the results of any such analyses are not included in the paper.

Evaluation

Strengths/ Weaknesses	This paper is lacking much detail about the development of the model. The authors also indicate that extensive sensitivity analyses were performed with respect to each type of time loss included a model, yet no results of the sensitivity analyses are included in the paper. There is also no evidence presented of any validation work having been done on the model.

Financial support for the development of this paper, and presumably for the development of the model came from a commercial entity that, as the maker and distributor of Sumatriptan® (Imitrex®) has a vested interest in the results of the model (GlaxoWellcome at the time of writing – now GlaxoSmithKline).

Because the principal aim of the model is to assess the economic consequences of migraine, there is no attempt to assess outcomes such as quality of life. The authors also admit to the controversial assumptions that are involved in the costing of time lost, particularly in the category of non work activity.

Presentation of Results

Model results for the primary outcomes of interest are presented in both tabular and graphical form. Tabled outcome data includes presentation of the mean and median annual hours of migraine symptoms and time lost from work and non work time. Tabled outcome data also include mean and median annual costs (standardized to 1997 Canadian dollars) associated with various time lost categories, as well as direct medical care costs associated with migraine. Graphical presentation of simulation outcomes show the (continuous) distribution of a simulated population of 10,000 individuals by their annual migraine related costs (direct, indirect, and total costs).

Interpretation of Results

The authors indicate that the model presents clear evidence that the societal costs of migraine are considerable, with the majority of costs being indirect, in terms of time lost from work, and non work activities. They conclude: "Given the indirect costs associated with migraine approach $2000 per patient and make up almost 90% of total costs, the potential for savings through more effective migraine therapies that reduce the amount of time lost from both work and nonwork activities due to migraine symptoms is enormous and could well offset the higher costs of newer therapies."(p. 290)

Value to Decision Making

The value to decision makers of utilizing this model is not clear from what is presented in this paper. The authors claim that the model can be adapted to other populations with specification of most model parameters according to user needs, and that it provides for the performance of sensitivity analyses on key parameters and assumptions. If these claims are true, the obviously proprietary nature of this model might still preclude it from being of much use in decision making.

Ease of Implementation

The authors do not provide enough information to develop a similar model, neither do they suggest that this model is available to the public. Parameterization of several migraine-specific aspects of the model would require detailed source data not typically available other than

from the sources used in this paper.

Further Reading

The model presented in this paper was developed by pharmaco-economic researchers at the consulting firm – Caro Research. The company's web site can be found here: http://www.caroresearch.com/

The migraine study representing a primary source of data for this model development is described further here: Heywood, J., Bouchard, J., Cortelli, P., et al., A multinational investigation of the impact of subcutaneous sumatriptan. I: Design, methods and clinical findings, *Pharmacoeconomics*, 1997; 11 Suppl(1): 11-23.

Article

Author Cooper, NJ, Sutton, AJ, and Abrams, KR

Title / Source **Decision analytical economic modeling within a Bayesian framework: application to prophylactic antibiotics use for caesarean section**, Statistical Methods in Medical Research, 2002. 11: 491-512

Context

Description

The primary purpose of this article is to review Bayesian models for economic decision making to demonstrate the benefits of this over the traditional "frequentist" approach to implementing simulation modeling. Based on this purpose, the first three sections of the paper are expositional and not descriptive of a particular modeling exercise. The fourth and fifth sections of the paper do, however, describe an illustrative example of a Markov Chain Monte Carlo (MCMC) simulation of the impact on cost outcomes of prophylactic antibiotic use in conjunction with caesarean section. It is these sections of the paper that are summarized here.

The illustrative example examines the cost implications of prophylactic antibiotic use to reduce the incidence of wound infections following caesarean sections in one hypothetical maternity hospital in the UK. The model describes a fairly simple two by two table of transition probabilities and their associated medical care costs that are the primary model outcomes of interest:

	Prophylactic antibiotic used?	
Wound infection?	No	Yes
No	(1-p1)	(1-p2)
Yes	(p1)	(p2)

Where:
(1 – p2) → Medical care cost with antibiotics
(p2) → Cost with antibiotics + wound treatment costs
(1-p1) → Cost with no antibiotics
(p1) → Cost of wound treatment

Outcome of Interest

Probabilities of wound infection following caesarean section with and without prophylactic antibiotic use, and the medical care costs associated with the caesarean section.

Model

Type

The model is a Markov Chain Monte Carlo simulation. Software code for the model is provided as an appendix to the paper.

Software

The model was implemented using the Bayesian statistics software package, WinBUGS. (See Further Reading section at end of review)

Model Quality

Data Sources

- Cost estimates used in the model were obtained from Netten, A., Dennett, J., and Knight, J., *Unit costs of Health and social care*. University of Kent: PSSRU, 1999.
- Effectiveness of prophylactic antibiotic use in preventing wound infection following caesarean section was derived through a Bayesian meta-analysis of 61 studies that were published in a systematic review as a Cochran Review – Smaill, F., and Hofmeyr, G.J. Antibiotic prophylaxis for cesarean section. *Cochrane Review* 2001; 3.

Parameters

In the hypothetical example, a single maternity hospital in the UK is assumed to have performed 750 caesarean sections in one year (1997).

- 8% (60 of these) assumed to have experienced postoperative wound infection.
- Costs of prophylactic antibiotics assumed to be fixed at £10.00.
- Consultant time to administer the antibiotic assumed to take 4 to 7 minutes at £1.00 per minute.
- Mean hospital length of stay for caesarean section without wound infection = 6.7 days at cost of £173 per day (1998/99).
- Mean hospital length of stay for caesarean section WITH wound infection = 8.8 days at cost of £262 per day (1998/99).

The model is built by first conducting a Bayesian meta-analysis of the pertinent literature in order to derive the key parameter of interest needed for the simulation – i.e. the relative risk of wound infection with and without the use of prophylactic antibiotics. The expected probability of wound infection in absence of antibiotic treatment is p_1. A very simple formula is used to derive the expected probability of wound infection if the new intervention is applied (e.g. prophylactic use of antibiotics): $p_2 = p_1 \times RR_{antibiotics}$. The Bayesian meta-analysis is the source from which the $RR_{antibiotics}$ is derived, and is specified as follows:

$$\gamma_i^c \sim Binomial(n_i^c, p_i^c) \quad \gamma_i^t \sim Binomial(n_i^t, p_i^t) \quad i = 1,\ldots, 61$$

$\mu = \log(p_i^c)$ $\log(p_i^t) = \mu_i + \min(\delta_{i,} - \log(p_i^c))$

$\delta_i \sim Normal(\Delta, \tau^2)$ $RR_{antibiotics} = \exp(\Delta)$

$p_i^c \sim Beta(\alpha, \beta)$ $\alpha \sim Uniform(1,100)$ $\beta \sim Uniform(1,100)$

$\Delta \sim$ Normal(0,0.1) $\tau^2 \sim$ InverseGamma(0.001,0.001)

"where for the *i*th trial, γ_i^c out of n_i^c have a wound infection in the placebo group and γ_i^t out of n_i^t have a wound infection in the prophylactic antibiotics group; P_i^c and P_i^t are the estimated infection rates in the placebo and prophylactic groups, respectively; μ_i the natural logarithm of the event rate in the placebo group; δ_i is the estimated $\log_e(RR_{antibiotics})$; Δ is the pooled $\log_e(RR_{antibiotics})$ and τ^2 is the between-study variance parameter often referred to as a heterogeneity parameter." (p. 500)

Simplifying Assumptions

Cesarean sections are not distinguished by type, though this could be investigated in sensitivity analyses.

The costs attached to resource use are also simplified, not reflecting a variety of real-world variability in terms of fee schedules, regional variability in reimbursement, contractual arrangements that lead to reduced charges per procedure, etc.

Duration / Time Perspective

The timeframe for the model is not perfectly clear from the paper, but appears to be short-term – i.e. one year's experience.

Iterations per Scenario

The simulation model was "burned in" (to achieve model convergence) using an initial run of 5000 iterations. These iterations were discarded and a further run of 20,000 iterations was used to derive the inferences for the primary model results.

Validation

No secondary validation is conducted here. However, in the Bayesian framework in which the model is developed out of an initial systematic review, one could argue that the formal meta analysis at the front end of model development serves as an implicit validation.

Sensitivity Analysis

The authors note that standard Cochrane reviews assess study quality in terms of allocation concealment on an "A" through "D" quality scale. They use this information to assess sensitivity of model results to changes in the set of studies included in the meta analysis under different quality thresholds. None of the studies used in their meta analysis was graded as "D" and only three were graded as "C." They ran the simulation excluding the three "C" grade studies and found that this exclusion had a minimal effect on the overall outcome. Specifically,

where the initial model results indicated a mean cost reduction of £49.53 per caesarean section with the use of prophylactic antibiotics (see Presentation of results below), in the model with a tighter study quality inclusion threshold, the mean cost reduction was £49.53, with 95% "Credible Interval" of £24.57 to £72.27.

A second type of sensitivity analysis conducted was to assess the influence of which "prior distributions" were used to specify the variance component parameters, τ^2. Three "non-informative" prior distributions were chosen: (1) the Inverse-Gamma (0.001, 0.001) distribution on τ^2, the Normal (0, 1.0^{-6}) distribution truncated at zero on τ, and the Uniform (0, 20) distribution on τ. The authors report that model results were robust over this array of distributions for the τ parameter.

Evaluation

Strengths/ Weaknesses

The authors describe a coherent, integrated, and systematic process of developing, estimating, and evaluating simulation models in a Bayesian framework. The four primary steps of this process are: (1) a systematic literature review of the substantive area to be modeled, incorporating meta-analyses; (2) estimating the model transition probabilities; (3) conducting sensitivity analyses for both the data and the model specification; and (4) evaluating the model.

The key characteristic of this approach is the explicit integration of information external to the model in its development and estimation (e.g. expert opinion; prior distributions). Whether this is viewed as a strength of a weakness will depend on how one evaluates Bayesian analytic methods compared to frequentist methods.

Presentation of Results

The model results are presented simply and coherently in a set of six graphs, each of which shows the posterior distribution of one of the primary outcome parameters of the model across 20,000 sample iterations.

Model results indicate a mean cost reduction of £49.53 per caesarean section with the use of prophylactic antibiotics. The 95% "Credible Interval" – a Bayesian analog to the confidence interval – for this mean reduction is -£77.09 to -£26.79.

Relative risk of infection with antibiotic use, $RR_{antibiotics,}$ was estimated to have mean of 0.30 with credible interval of 0.21 to 0.40.

Probability of a wound infection in the placebo group, p1, had mean of 0.08, CI of 0.06 to 0.10.

	Probability of a wound infection in the group receiving antibiotics, p2, had mean of 0.02 with CI of 0.02 to 0.03. Number of wound infections avoided using antibiotics was estimated at 42.55 with CI of 31.59 to 55.04. Between-study variance, τ^2, was estimated as 0.30 with CI of 0.05 to 0.74. Cost-effectiveness results were presented graphically on a quadrant-based graph that indicated that for all 20,000 model iterations the cost difference of using prophylactic antibiotics were negative (favorable) and the number of wound infections avoided were positive (favorable) suggesting that prophylactic antibiotic treatment was the dominant strategy.
Interpretation of Results	The results are interpreted in terms of incremental cost savings (in British £), and reduction in the number of infections between a treatment and a placebo condition.
Value to Decision Making	The specific example of prophylactic antibiotic use in cesarean sections may be useful for decision making around that specific issue, but this is secondary to the main purpose of the paper. The Bayesian MCMC modeling process described in this paper would be most useful to analysts already familiar with more traditional "frequentist" forms of simulation modeling for decision support.
Ease of Implementation	For analytic topics where a systematic literature review already exists (e.g. a Cochrane review), the Bayesian MCMC modeling method would appear to be fairly straightforward for those previously familiar with the Bayesian analytical framework. Actuaries not already familiar with this framework would need to educate themselves sufficiently to use the WinBUGS sofrware program to apply the approach.
Further Reading	WinBUGS software is freely available at: http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml

Article

Author	Cooper NJ, Sutton AJ, Abrams KR, Turner D, and Wailoo A
Title	**Comprehensive decision analytic modeling in economic evaluation: a Bayesian approach**
Source	Health Economics, 2004. 13:203-226.

Context

Description

The authors used two illustrative examples (the prophylactic use of neuraminidase inhibitors (NI) to reduce the incidence of influenza and the use of taxanes for the second-line treatment of advanced breast cancer compared to conventional treatment) to demonstrate the use of Bayesian Markov Chain Monte Carlo (MCMC) simulation. The authors claim that Bayesian MCMC simulation has advantages over the conventional approach for evaluation of decision analytical models. The conventional approach has two stages: 1) systematic reviews and/or meta-analyses are used to estimate the decision model parameters, 2) a decision model is developed with parameter estimates defining the variable distributions and the model is evaluated using Monte Carlo simulation. The conventional approach should be familiar to actuaries and is often carried out using excel or other spreadsheet software.

The Bayesian MCMC approach, in comparison, allows the analyst to incorporate information external to that included directly in the model (such as professional judgment of clinicians). The approach has four stages: 1) a systematic review of the relevant data incorporating meta-analyses, 2) estimation of all inputs into the model (including effectiveness, transition probabilities, and costs), 3) sensitivity analysis for data and model specifications,, and 4) evaluation of the pre-developed model.

Outcome of Interest

For the first illustration (prophylaxis use of NIs), the outcomes of interest were: odds ratio of contracting influenza, probability of contracting influenza with prophylaxis NIs, standard probability of contracting influenza with no prophylaxis, cost difference (NIs – standard), and days avoided (NIs – standard). For the second illustration (taxane use), the outcomes of interest were: incremental cost and incremental utility.

Results were given for the Bayesian MCMC approach and the conventional two stage approach: the 95% confidence limits were narrower for the conventional approach in both illustrations.

Model

Type	Bayesian Markov chain Monte Carlo simulation compared to a conventional Monte Carlo simulation.
Software	Win-BUGS software for Bayesian MCMC. Code was provided by the authors.

Model Quality

Data Sources	The authors used meta-analysis to develop model parameters; the source publications were provided by the authors.
Parameters	For the NI illustration, the random parameters were the rates of influenza, the cost for treatment, and the number of days until symptoms were alleviated. For the taxane illustration, the several random parameters were developed from the meta-analyses. They include time to event (from response to stable state, from progressive to response state, overall survival time) and several response rates.
Simplifying Assumptions	It is unclear what simplifying assumptions were made.
Duration / Time Perspective	For illustration one, the model is run for an influenza epidemic period. For illustration two, the model is run for about two years, which represents 35 3-week cycles of chemotherapy treatment intervals and by which time the majority of the individuals have reached the absorbing state (death).
Iterations per Scenario	For the NI illustration, 20,000 iterations. For the taxane illustration, 10,000 iterations.
Validation	The authors did not discuss validation.
Sensitivity Analysis	The authors discuss sensitivity analyses for both the data and the model specification, for both of the illustrations. The authors test different prior distributions for their impact on model results.

Evaluation

Strengths/ Weaknesses	The authors gave a high level overview of the advantage of Bayesian methods for decision analytic modeling. Then they used two examples to illustrate the use of Bayesian MCMC methods and compared results

	to the conventional Monte Carlo method. This article is valuable as an introduction to Bayesian methods for effectiveness analysis of various treatments within health care, and the authors provided WinBUGS code and references for further reading. However, the interested reader will need to pursue the topic further in order to use the techniques described.
Presentation of Results	Results were presented as aggregate outcomes over the duration of the model run, with 95% confidence intervals.
Interpretation of Results	Findings are interpreted as incremental gains (losses) between two treatment scenarios.
Value to Decision Making	The simulation techniques would be useful to decision-makers choosing between treatment scenarios A and B. The authors attempt to make a case for the use of Bayesian MCMC techniques, however the results do not appear to be greatly different from the results obtained from the more straightforward and conventional Monte Carlo simulation.
Ease of Implementation	Actuaries who are interested in Bayesian MCMC techniques could easily educate themselves on the use of WinBUGS and apply these modeling approaches to economic decision problems.
Further Reading	The authors suggest a reference to Spiegelhalter et al. as further reading on Bayesian methods in Health Technology Assessment. They recommend the WinBUGS tutorial in the BUGS manual and the BUGS website (http://www.mrcbsu.cam.ac.uk/bugs/welcom.shtml) for an introduction to using WinBUGS.

Article

Author	Davies, R., Crabbe, D., Roderick, P., Goddard, J.A., Raftery, J., and Patel, P.
Title	**A simulation to evaluate screening for Helicobacter Pylori infection in the prevention of peptic ulcers and gastric cancers**
Source	*Health Care Management Science*, 2002. 5: 249-258.

Context

Description

This model uses discrete event simulation to estimate the population level impact of a one-time screening for Helicobacter pylori infection (H. pylori) of individuals under age 50 at the time of screening. The model simulates a population of people alive in the year 2000, simulating their progress over an 80 year period. New individuals enter model with no gastric problems. Individuals can be determined to have an H. pylori infection by screening. Individuals may also be diagnosed with gastric cancer, which puts them at risk for the outcome of death due to gastric cancer. Individuals may also present to a doctor with a suspected ulcer, and may then be subjected to opportunistic testing for H. pylori infection. Those determined to have H. pylori infection are simulated to have the infection eradicated or not eradicated. Individuals with suspected ulcer may also be stimulated to develop a peptic ulcer complication, which may subject them to the risk of mortality from that complication. The model allows individuals to also die from other causes, while in any state in the model.

Outcome of Interest

Primary outcomes in the model include the immediate risk of peptic ulcer, and the delayed risk of gastric cancer associated with H. pylori infections. Mortality, "costs," and years of life saved are also outcomes generated by the model.

More specifically, the authors developed this model in order to simulate the number of gastric cancer deaths that could prevented, the years of life that could be saved, and the costs of implementing a hypothetical, population based screening and treatment program to detect and eradicate H. pylori infections.

The model results suggest that for the population of England and Wales under age 50 in 2000, compared to the baseline results of the model, an H. pylori screening and eradication program as specified in this paper would result in roughly 16,000 total deaths prevented to age 75; 34,000 total deaths prevented; and 368,000 life-years saved. The costs of

implementing the program are estimated to be roughly £19 million (upper and lower limits of £24 million and £12 million, respectively) in the first year.

Model

Type

This paper is based on a discrete event simulation model. The simulation model begins with a prevalent population, with new individuals added to the population annually for 20 years, starting at age 20. Individuals are assumed to be free of symptoms at entry to the model, but are assigned age-specific levels of H. pylori prevalence based on published estimates.

Software

The model uses what the authors refer to as the patient oriented simulation technique (POST) and was written using the Delphi software development package from Borland Software Corporation. (http://www.borland.com/us/products/delphi/index.html)

Model Quality

Data Sources

The authors indicate that model parameters and estimates were derived as much as possible using Data from the United Kingdom. The authors state that data used to construct the simulation were obtained from published databases and peer-reviewed published papers, references for which are given in the paper.

Parameters

In this simulation, some parameters of the model were set based on published estimates, while other parameters were derived from randomly sampling distributions (presumably using Monte Carlo methods, though this is not made explicit by the authors).

Parameters based on sampling distributions include: Time to death from non-gastric cancer or ulcer; time to duodenal ulcer; time to gastric ulcer; time to gastric cancer; time to be screened; time to H.pylori infection if not currently infected. With respect to duodenal ulcer, parameters established by random sampling include the probability of complication; probability of death; and time to duodenal ulcer. With respect to gastric ulcer, parameters established by random sampling include probability of complications; probability of death; and time to gastric ulcer. With respect to gastric cancer, time to gastric cancer death is established by random sampling. Time to re-infection with H. pylori among those in whom it has previously been eradicated is also established by random sampling. Events of H. pylori infection or re-infection as well as death are also established based on random sampling.

Model parameters based on published estimates include: the relative risk of ulcers; relative risk of gastric cancers; relative risk of non-ulcer dyspepsia; compliance with screening; compliance with treatment; age-

specific H. pylori prevalence and a “lag” time between H pylori eradication and when an individual’s risk of gastric cancer are reduced.

Unit costs and ranges of costs based on published estimates (all in terms of British £, adjusted to year 2000 values) include: H. pylori screening program cost; opportunistic testing for H. pylori using the urea breath test; cost a visit to a doctor; cost of the presumed H. pylori treatment (combination of a proton pump inhibitor – omeprazole, with antibiotics clarithromycin and metranidazole); cost for gastric cancer treatment; and cost for complications of peptic ulcers.

Simplifying Assumptions

Differences in local practices and guidelines for H. pylori, testing and treatment lead to great variability in the level of opportunistic testing and treatment. To reduce additional variability related to these parameters in the simulation, the authors assume opportunistic testing has both sensitivity and specificity near 100%, as well as 100% compliance with both screening and treatment.

In the systematic screening program, individuals who are "invited" for testing, and who comply with this request are simulated to receive a serology test with sensitivity of 0.9 and specificity of 0.95. The H. pylori eradication treatment (proton pump inhibitor, combined with two antibiotics) is assumed to have an efficacy rate of 89% with variability built into the model to allow for imperfect compliance on the part of individuals with taking the medications as prescribed.

Individuals “starting” in the simulation at ages over 50 were eliminated from the model (age 50 being the oldest age at which screening was assumed to take place).

Screening is modeled to take place only once per individual in the model, with screening taking place between the ages of 40-50 for those in the “prevalent” population, and at age 40 for those new individuals who enter the simulation at age 20.

Duration / Time Perspective

Results were simulated for all individuals ages 20-50, with simulated yearly updating of parameters until all subjects “died,” or the simulation had run for 80 years.

Iterations per Scenario

The authors report that their results are based on 2,500 replications of runs totaling 2.9 million individuals ages 20 to 50 in the prevalent population and 100,000 individuals in the incident population of 20-year-olds entering the simulated population.

Validation

The only evidence of validation of the model is that the authors have assessed the gastric cancer mortality results from the simulation model to gastric cancer mortality data for England and Wales from the Office of National Statistics. The authors report that the model results fit the real world data well in general, but with a less good fit for those over age 70.

Sensitivity Analysis

The authors engaged in a systematic assessment of sensitivity of the simulation results by running a complete factorial design. Models were run for every combination of high and low range values of five sets of key model parameters, resulting in 64 different simulation runs to test sensitivity.

Evaluation

Strengths/ Weaknesses

This model is designed to assess the population level health and cost outcomes related to a population level screening in eradication program for H. pylori infection. It would not be particularly useful for simulating individual level outcomes of, say, hypothetical alternative treatments for H. pylori infection.

The authors acknowledge that the model results are particularly sensitive to several key parameters. In particular, these include the population incidence of gastric cancer, and the prevalence of H. pylori infection. Model results are also sensitive to assumptions about the extent to which H. pylori infection increases the risk of peptic ulcers and gastric cancers. In their concluding comments, the authors note that better research data are needed to establish key parameters in the simulation.

Presentation of Results

Model parameters are presented primarily in tabular form with two figures; one that presents the rates per 100,000 population of peptic ulcers, complications from peptic ulcers, and death rates, and a second that presents gastric cancer mortality by birth cohort in terms of deaths per million.

Simulation model run results are presented in a single table (Table 6)in terms of means and confidence limits for each of the specific output parameters in the model.

Results of the sensitivity analysis are presented in tabular form as well as in the text. Sensitivity analyses indicate that all parameters and pairs of parameters have a significant impact on the model results, the magnitude of which can be assessed based on the table presented in the paper (Table 7).

Interpretation of Results	The authors interpret their results as indicating that a population level screening program for H. pylori in the general population (of England and Wales) would be of benefit in terms of reduced mortality and morbidity, and would have program related costs of roughly £19 million in the first year of operations.
Value to Decision Making	This article offers an example of how to develop and use simulation modeling to estimate the health and economic impact of treatment changes for a specific condition – H. pylori infection. The value of the specific results to decision makers in the U.S. is an open question because the cost parameters are likely to be quite different in our health care system than they are in the U.K. In addition, a more recent publication has suggested that there are serious flaws in some of the key assumptions used in past research on the cost effectiveness of treatment programs. See: Fairman & Motheral. 2003. Do decision analytic models identify cost-effective treatments? A retrospective look at Helicobacter pylori eradication. J Manag Care Pharm, 9(5): 430-440.
Ease of Implementation	If available publicly, this type of model could be used by actuaries. However, the user would need to be familiar with H. pylori, as well as screening and eradication regimens, and associated costs to replicate these results for a different population, or to modify the model.
Further Reading	The paper lists two references that appear to give further information about the POST simulation tool developed and used by the authors in this paper: (1) H.T.O. Davies and R. Davies, Simulating health systems: Modeling problems and software solutions, European Journal of Operational Research, 1995, 87: 35-44. (2) R. Davies, S.C. Brailsford, P.J. Roderick, C. Canning and D. Crabbe, Using simulation modeling for evaluating screening services for diabetic retinopathy, Journal of the Operational Research in Society, 2000, 51: 476-484.

Article

Author	Eastman RC, Javitt JC, Herman WH, Dasbach EJ, Zbrozek AS, Dong F, Manninen D, Garfield SA, Copley-Merriman C, Maier W, Eastman JF, Kotsanos J, Cowie CC, Harris M
Title	**Model of Complications of NIDDM** **I. Model construction and assumptions** **II. Analysis of the health benefits and cost-effectiveness of treating NIDDM with the goal of normoglycemia**
Source	*Diabetes Care*, 1997. 20(5): 725-734, 735-744.

Context

Description

The model simulates a population of individuals who are newly diagnosed with non-insulin dependent diabetes. Individuals are projected from age at onset of NIDDM through death or to age 95. Fourteen health states are modeled, each state a complication of NIDDM: retinopathy (eye disease – 5 states), nephropathy (kidney disease – 4 states), neuropathy (nerve disease – 3 states), and cardiovascular disease (2 states). In Part II, the authors estimate the cost and quality of life impacts of additional treatment which reduces HbA_{1c} from the assumed diabetic level of 10.0% to a desired level of 7.2%.

In Part II, cost-effectiveness is analyzed from the perspective of a single-payer responsible for all direct medical costs.

Outcome of Interest

In Part I, the outcomes of interest are cumulative incidence rates of the modeled population. Outcomes for rates of retinopathy health states, nephropathy health states, and neuropathy health states are displayed graphically.
In Part II, the cost-effectiveness outcomes are expressed as the difference between standard and comprehensive treatment costs divided by the difference in effectiveness measured in quality adjusted life years (QALYs).

Model

Type

This is a stochastic model using Monte Carlo techniques. To begin a simulation run, a cohort of 10,000 hypothetical patients are assigned age, sex, and ethnicity characteristics. Each patient's annual health state is projected forward until death or age 95. The simulated patients represent a population with newly clinically diagnosed NIDDM; the patients are aged 25-74 at the start of the simulation. Monte Carlo techniques are used to progress patients through the simulation;

transitions to progressively worse health states are irreversible.

Software

The model was developed using simulation software (@Risk version 3.5b for Windows, Palisades, Inc., Newfield, NJ) and Excel Version 7 (Microsoft, Seattle, WA).

Model Quality

Data Sources

Hazard rates for transition from one health state to a more severe health state are estimated using data from the Wisconsin Epidemiologic Study of Diabetic Retinopathy (retinopathy and nephropathy rates), the Rochester Epidemiology Project (nephropathy and neuropathy rates), and the National Health and Nutrition Examination Survey (NHANES) II (neuropathy rates). The authors clearly describe demographics and transition probabilities in Table 2 of the article. NHANES II and Hispanic Health and Nutrition Examination Survey are sources for CVD risk factors, summarized and displayed in Appendix B.

For Part II's economic analysis, costs based on Medicare reimbursement races. The cost of comprehensive care was estimated from the VA Cooperative Study and the Metformin Cooperative Trial. The authors give detail on estimates of treatment intensity and cost for both standard and comprehensive care.

Life years are adjusted by utilities reported in the literature.

Parameters

The age, sex, and ethnicity characteristics of the simulated cohort are determined by sampling from pre-assigned frequency distributions. Users of the model could modify the characteristics of the cohort by modifying the frequency distributions used for sampling.

As each hypothetical patient is projected forward, the patient's health state is randomly determined using Monte Carlo techniques and the hazard rates for each health state and mortality. NIDDM complication hazard rates vary by duration of diabetes and ethnicity.

In Part II, the effect of glycemic control is modeled by adjusting the incidence rates for complications. Standard care is represented by a model run with HbA_{1c} of 10.0%; it is compared to a comprehensive care model where HbA_{1c} is 7.2%. The authors model decreasing treatment compliance by specifying HbA_{1c} between 7.2% and 10.0%, however their description for doing so is incomplete.

Simplifying Assumptions	The model is limited to ages 25-74 at incidence of clinically diagnosed diabetes. Risks for cardiovascular disease (smoking, cholesterol levels, and blood pressure levels) are randomly assigned to a patient at the start of the simulation, but it is not clear whether the assigned values can be changed during the projection. The authors did not model peripheral vascular disease and their model does not accumulate statistics on compound health states. The authors assume multiples of the average hazard rate for various ethnic populations, and those assumptions are clearly described in the text and in Table 2. Assumptions about progression of disease are also described in the text and in Table 2.
Duration / Time Perspective	Hypothetical patients in the cohort are projected from starting age through death or age 95.
Iterations per Scenario	The authors are silent about number of iterations per scenario; we can assume that each scenario is one iteration of a 10,000 member cohort.
Validation	Hazard rates from the model simulation are compared to their source rates. The authors produced graphs comparing the modeled cumulative rates and the source rates.
Sensitivity Analysis	In Part II, costs and QALYs are discounted at a 3% rate. Rates of 5% and 7.5% are also evaluated.

Evaluation

Strengths/ Weaknesses	The articles provide clear and straightforward description of how the hazard rates were developed and how the model projects hypothetical patients from year to year. Part I gives an easy to read table with hazard rates, and appendices describing hazard rate development, a glossary of terms and phrases, and tables of risk factors for cardiovascular disease.
Presentation of Results	In Part II, the authors display the cumulative incidence of blindness, end stage renal disease (ESRD), first lower-extremity amputation (LEA), and survival and compare the rates between standard care and comprehensive care. They also summarize the present values of cost of treatment, effectiveness (life years and QALYs), and incremental cost-effectiveness for the standard care and comprehensive care scenarios. There is a net increase in cost of care under the comprehensive scenario: cost of eye disease, renal disease, neuropathy/LEA decrease while cost of general and diabetes-related medical care and new coronary heart disease increase. The average incremental cost /QALY gained is $16,002.

The authors also investigate various subpopulations to find where the investment in glycemic control has the greatest returns. They discover that the greatest efficiency is for younger patients and for those with the highest absolute risk o complications under standard care: the cost/QALY is lowest for minorities and those with higher HbA_{1c}.

Interpretation of Results

Value to Decision Making

The authors' intent in Part I was to develop a model that could be used to analyze strategies to prevent complications in NIDDM. In Part II, the model is used to estimate the health benefits and economics of treating NIDDM with the goal of normoglycemia (HbA_{1c} of 7.2%).

This pair of articles is valuable in three ways. First, the articles provide a summary of descriptive statistics on demographics and disease progression rates for a patient population newly diagnosed with NIDDM. Second, the articles offer a detailed example of how to develop and use simulation modeling to estimate the health and economic impact of treatment changes. Third, the results from Part II provide information that could prove useful when making decisions about investment in a particular change: glycemic control.

Ease of Implementation

If available publicly, this type of model could be used by actuaries. However, the user should be familiar with NIDDM and expected treatment impacts in order to modify parameters to fit a specific population and question.

Further Reading

This pair of articles was published in 1997. Although the authors mentioned in Part I that "this article is the first in a series of studies …" we did not uncover additional published articles on this modeling effort.

Article

Author	Keeler EB, Malkin JD, Goldman DP, Buchanan JL
Title	**Can Medical Savings Accounts for the Nonelderly Reduce Health Care Costs?**
Source	JAMA, 1996. 275(21):1666-1671.

Context

Description	This is a simulation of one year of health care expenditures for a nationally representative sample of the under-65 insured population.
Outcome of Interest	The authors used stochastic simulation to estimate the changes in national health expenditures that would result from introduction of an insurance benefit designed as a catastrophic health insurance coupled with a medical savings account.

Model

Type	The authors do not state the model type, but it appears to be a microsimulation of insured families. In earlier work, the authors used data from the Rand Health Insurance Experiment (HIE) to develop an expected rate of treated episodes of medical care, assuming full insurance coverage, and the associated cost of those episodes. In this work, the authors randomly generate a number of episodes per individual for one year conditional on the expected rate, and the cost of those episodes. Some of the episodes that would be treated under full coverage would not be treated under an insurance plan with cost-sharing. The model simulates the number of treated episodes depending on the cost-sharing, which varies by insurance benefit. In addition, the authors develop a selection model to simulate the insurance benefit selected when multiple offerings are available.
Software	The authors are silent about the type of software used to develop the models.

Model Quality

Data Sources	Results from the Rand HIE were used to estimate rates of treated episodes under free care (100% coverage). Health Care Financing Administration and National Medical Expenditure Survey data were used to update the episode costs to reflect changes in health spending from the time of HIE to 1996.

	23,157 observations from the 1993 Current Population Survey was used to represent the national population.
Parameters	In the simulation, the number of treated episodes each year under free care is generated by random draws and is conditional on the expected rate. The choice model (of insurance benefit plan) uses a number of variables, one of which is expected out-of-pocket spending. However, it is unclear if random variables are used to compute expected out-of-pocket spending.
Simplifying Assumptions	Marginal tax rates of 15% and 40% were assumed for low and high income families, respectively. Marginal tax rates of 32% and 45% for low and high income families were used for employer contributions that are exempt from Social Security and federal and state income taxes. Insurance plan choice simulations assume that $1 of HMO care is worth $0.90 of FFS care; the assumption represents the value individuals put on choice of physician and freedom from managed care oversight. Funds in an MSA are assumed to have value interchangeable with other savings dollars.
Duration / Time Perspective	The model projects health care expenditures for one year. It uses experience from the previous year to develop premium rates and expected family expenditures.
Iterations per Scenario	There are two scenarios: the first with all households insured by a catastrophic plan plus a medical savings account that is employee funded, the second with multiple offerings where households select (through the choice model) into the insurance plan that has the greatest utility for them. In the second scenario, there are three insurance plans: employee funded MSA with catastrophic insurance, employer funded MSA with catastrophic insurance, and HMO. The model simulates 23,157 households for each scenario.
Validation	The authors state that the model of response to cost-sharing (episodes not treated because of out-of-pocket cost) has been tested extensively and they site references to that work.
Sensitivity Analysis	The authors tested sensitivity of results to various assumptions ($1 of HMO care valued as $0.90 of unmanaged care, both employer and employee funded MSAs offered, and funds in MSAs valued the same as money). They find that the assumptions have an impact on which benefit plans are selected, but little impact on aggregate spending.

Evaluation

Strengths/ Weaknesses

The model recognizes the interdependence among: tax-advantaged medical spending, cost-sharing (deductible, coinsurance), treatment-seeking for episodes of medical care, and out-of-pocket maximum. The results of these interdependencies on health care expenditures are not easy to anticipate without using a model. This microsimulation model addresses the tax implications and financial implications of the various insurance benefit designs.

The health expenditure model relies on behaviors observed in the Rand HIE. The data from the HIE are old, and one might argue that they are too old to be valuable. However, the authors recognize the age of the HIE data and assert that the behaviors about money and health care are constant. In addition, the authors have updated the episodes to reflect changing health care service use, less frequent hospitalization, for example.

The authors assume that households consider tax-implications when making selections of an insurance benefit plan and seeking care for medical episodes. Those assumptions may not be realistic. Tax implications are complex and perhaps beyond the grasp of the average family making health insurance and care decisions.

Presentation of Results

Authors present a very simple, illustrative graphic of the after-tax out-of-pocket for various insurance benefit designs and various medical expenses. They use that graphic to explain the differences in out-of-pocket expenditure due to tax-implications for the various benefit designs. The results of the simulations are shown in two tables. The first table compares the average spending, changes (relative to the FFS plan) in spending, tax cost, value to the consumer, value to society of the alternative designs (HMO and MSAs). The authors also show estimates of waste, after-tax out-of-pocket, risk, and the % of households with spending over the out-of-pocket maximum. The authors also use a table to show the results of insurance benefit choice; total average spending and income are shown for each insurance benefit offered.

Interpretation of Results

Results are interpreted as the societal impact of insurance benefit choice.

Value to Decision Making

This article carefully explains the cost-sharing and tax-incentives of various insurance benefit designs and how those incentives can modify care-seeking behavior. However, one might question whether employees are sophisticated enough to consider tax implications when

choosing an insurance benefit or when deciding whether to seek care for a medical episode.

Ease of Implementation

The authors rely on previous work from the Rand HIE to develop this simulation model. Actuaries would need to invest considerable time and resources to develop similar models. That said, the concepts used in the Keeler et al. model may be implemented in actuarial practice. Actuaries could easily use experience data to develop an insurance benefit choice model which could be used to project select of multiple benefit plans. Perhaps more difficult, but still possible, actuaries could also use their own administrative claims data, which has been grouped into episodes of care, to model care-seeking behavior as it varies by insurance benefit design. Both of those applications would require some econometric modeling skills.

Further Reading

Goldman DP, Buchanan JL, Keeler KB. Simulating the Impact of Medical Savings Accounts on Small Business. *Health Services Research*, 2000, 35(1 Part I):53-75.

Article

Author	Kong DF, Eisenstein EL, Sketch MH Jr, Zidar JP, Ryan RJ, Harrington RA, Newman MF, Smith PK, Mark DB, Califf RM
Title	**Economic impact of drug-eluting stents on hospital systems: A disease-state model**
Source	American Heart Journal, 2004. 147:449-56.

Context

Description	The model simulated the financial consequences of treatment pattern changes when drug-eluting stents (coated stents) were added as a treatment option for patients with coronary heart disease. The model population consisted of patients with coronary heart disease at a single hospital (Duke Medical Center) and the outcome was the financial impact on that individual hospital.
Outcome of Interest	The authors were interested in financial results: the contribution margin (revenue less variable costs) and net profit (revenue less total costs).

Model

Type	This was a state transition model. Five treatment options for coronary heart disease were modeled at a single hospital. Individual patients were followed for one to two years during treatment for coronary heart disease. Cohorts of patients entered the model each simulated year.
Software	The model was developed using SAS software.

Model Quality

Data Sources	Experience data came from Duke University Medical Center: the Duke Databank for Cardiovascular Disease from January 2000 through December 2000, with 3,112 patients in database during this time period; the Duke cost-accounting data during the same sampling period; and a survey of 13 (11 respondents) faculty interventional cardiologists responsible for selecting treatment strategies at Duke Hospital.
Parameters	Clinical parameters included revascularization rates, restenosis rates, and number of stents per patient. Clinical parameters were estimated from two sources: 1) the Duke Databank for Cardiovascular Disease and 2) published literature from two trials (RAVEL and SIRIUS). Cost and reimbursement parameters were modeled from the Duke cost-accounting data.

	Patient outcomes (symptoms relieved and exit model or recurrent symptoms and additional treatment) were determined at random from a uniform distribution.
Simplifying Assumptions	Fixed costs were assumed the same for bare and coated stents. Bare stents were assumed to cost $1000 each; coated stents were assumed to cost $3200 each. Variable costs (excluding cost of stent) were assumed to be the same for bare and coated stents, and based on the variable costs of bare stents. The number of new patients was assumed to remain constant in each year of the simulation at the level of the base year. Patient volume varied by necessity of repeat procedures, which varied by type of procedure and was substantially lower for coated stents than bare stents. The diversion of patients from existing therapies to the new therapy (coated stents) was a key assumption. Survey responses were used to estimate physician behavior for selecting treatment strategies; mean survey responses were used to estimate the change in treatment to coated stents.
Duration / Time Perspective	The model projected 5 years of financial results, but the authors did not mention discounting over the time period.
Iterations per Scenario	The authors used 1000 model runs for each set of parameters. Simulation probabilities represented patient outcomes (relieved of symptoms, restenosis, repeat revascularization).
Validation	Authors did not address whether the base-line scenario (before transition to coated stents) validated with actual net profit amounts.
Sensitivity Analysis	The transition from current therapy mix to therapy mix including coated stents was tested using parameters from cardiologist survey responses. The Authors also tested the model for sensitivity to two assumptions: 1) risk of repeat revascularization and 2) revascularization rate for patients diverted from CABG or medical therapy.

Evaluation

Strengths/ Weaknesses	This was a fairly straightforward model. The key assumptions were the net profits from each treatment option and the mix of treatments used. From the baseline data provided, the hospital had greatest profit per case for CABG ($2,369) and angioplasty ($387). Even before considering the addition of coated stents, the hospital experienced a $29 loss per bare stent case and $841 per medically treated case. The gains and losses were not highlighted in the paper. Because the coated stents were

	assumed to cost more than bare stents ($3200 vs. $1000) without a comparable revenue increase, any shift from bare to coated stents increased losses. Because the authors assumed shifts from profitable procedures (CABG and angioplasty) to non-profitable procedures (coated stents), the results were not surprising. The application of the simulation was intended to support and quantify the expected losses. The stochastic portion of the model was the outcome for each patient with the possibility of repeat treatment. Scenarios were used to model various treatment distributions. Scenarios were also used to estimate effects of various reimbursement. However, financial impact appeared to result from use of coated stents (a loss procedure) in place on CABG (a profitable procedure).
Presentation of Results	The authors presented mean losses and 95% confidence intervals for year one and after. Losses were compared to the base line (before treatment converted to coated stents).
Interpretation of Results	The authors considered financial changes for the cardiac unit, the hospital as a whole, and similar hospitals.
Value to Decision Making	The simulation quantified a potential financial problem to inform policy makers of how changing treatments and reimbursement policies can impact hospitals.
Ease of Implementation	This type of model can be easily implemented in an actuarial setting. The data for baseline parameter values would be available from the employer or client. Assumptions about treatment were generated from a survey polling expert opinion, a form of data gathering that is accessible to practicing actuaries.
Further Reading	None.

Article

Author	Lauer, J.A., Röhrich, K., Wirth, H., Charette, C., Gribble, S., and Murray, C.J.L.
Title	**PopMod: a longitudinal population nmodel with two interacting disease states**
Source	Cost Effectiveness and Resource Allocation, 2003. 1(6)

Context

Description

PopMod is the first published "multi-state dynamic" life table model, designed to simulate the mortality experience of a simulated population under the force of two disease conditions that are allowed to interact, as well as a "background" force of mortality. It differs from other multi-state life table models in that it does not assume statistical independence between competing causes of death (a typical assumption of competing risks models) and because it models age and time independently.

The primary purposes of PopMod are; (1) describing the evolution of a population's health over time for typical demographic purposes, and (2) providing a model to yield comparable measures of "effectiveness" for intervention and cost-effectiveness studies.

Outcome of Interest

Projection of an arbitrary population forward in time at yearly intervals, with output including the size of age and sex specific population subgroups at each interval - corresponding to the l_x column of a standard life table. The model also outputs time at risk for these subgroups during yearly intervals – corresponding to the life table column L_x.

If estimates of disease severity are available for the two diseases being modeled (diseases that can be modeled exclude all "transmissible" diseases) the model will output all standard life table measures along with other summary measures of population health.

Model

Type

PopMod is perhaps best described using the term chosen by the authors – a dynamic multi-state life table model. It is a "compartment model" of differential equations modeling the inflow and outflow between model states, and in which the assumption is made that the risks of mortality are conditional on disease states. As noted by the authors, it has similarities to "incidence, prevalence, mortality" (IPM) models but with the additional detail of specifying a population with age and sex structure typically absent from IPM models. The authors also note the

model's similarities with the Prevent model (discussed in the review of Naidoo, 1997) – another population model for analyzing age and sex structured population outcomes, but which is limited to modeling a single disease condition.

The model contains six population states, with birth and death being special states that can only "send" or "absorb" from the four primary model states which represent; (1) those with neither of the diseases being modeled, (2) those with the first of the two diseases modeled, (3) those with the second of the two diseases modeled, and (4) those with both diseases modeled. The model includes twelve hazard rate transitions between these model states.

Software The modeling software was not discussed by the authors.

Model Quality

Data Sources This paper only describes the model structure, but does not present any model results, hypothetical or real. Minimum model input that the analyst must have available include (presumably age and sex specific) rates of the prevalence, incidence, remission and cause-specific mortality rates for the two diseases to be modeled.

Parameters Input parameters would appear to be a starting (presimulation) age and sex structure of an arbitrary population, along with specification of incidence, prevalence, remission and cause-specific mortality rates to be applied to the population. The six model states along with the transition rates, or hazard rates for flows between these states (as described above in Model Type) represent the other essential model parameters.

Simplifying Assumptions Simulated individuals are not allowed to simultaneously acquire both diseases of interest, nor are individuals allowed to have two simultaneous causes of death. These transitions are constrained to not occur in the model.

The most limiting assumption of the model is the assumption of homogeneity of individuals within each of the six model states. The authors describe this as "compression," and note that the assumption of homogeneity within model states may be, on the one hand, a reasonable simplifying assumption, but on the other, may not adequately address potential confounding if there are important heterogeneities across individuals within the states described by the model.

Duration / Time Perspective PopMod is designed to be run over a simulated period of 100 years. Presumably, individuals are projected for either a lifetime, or for the

	portion of their lives during which a particular disease may be operative (i.e. heart disease and stroke are typically not a substantial concern in populations under age 35). The perspective is of the population.
Iterations per Scenario	This article describes the model and its applications; it does not discuss iterations or sample size per scenario.
Validation	The authors do not describe validation of the PopMod model, but provide a fairly in-depth technical description of the model, it's assumptions, and the numerical algorithms it employs.
Sensitivity Analysis	The paper does not present any model results, nor are issues of model quality or sensitivity analyses discussed.

Evaluation

Strengths/ Weaknesses	This model resembles standard multi-state life-table techniques more than it does microsimulation models. The strengths it has over more traditional multi-state life-tables include the ability to simultaneously model the mortality effects of two competing causes of death that are NOT assumed to be statistically independent, along with the ability to model the effects of age separately from the effects of time. The model can, apparently, be used as one component of estimating cost effectiveness of population level health interventions (see Evans at al.) though this is not described in the article. Weaknesses include the requirement for fairly detailed input data on the disease states to be modeled. State-transition models such as PopMod do not allow for specification of individual level heterogeneity such as microsimulation models do. The authors point this out, and note that if the analyst suspects that such heterogeneity in the population exists that may confound the projections provided by multi-state life-table methods, that a microsimulation method may be preferable.
Presentation of Results	The purpose of the article is to describe PopMod and its potential; no model results are presented in the article.
Interpretation of Results	The purpose of the article is to describe PopMod and its potential; no model results are presented in the article.
Value to Decision Making	From the article, it is not perfectly clear what the value of the model might be to decision making. However, a 2005 paper published in the British Medical Journal (David B Evans, Tessa Tan-Torres Edejer, Taghreed Adam, Stephen S Lim, "Methods to assess the costs and

health effects of interventions for improving health in developing countries," *BMJ* 2005;331;1137-1140) demonstrates the potential usefulness of PopMod as part of a larger modeling process to assess the costs and health impacts of health promoting interventions deployed in developing nations.

Ease of Implementation

The PopMod model was designed to be generic, making it potentially applicable to multiple populations and multiple disease conditions. However, the level of input data required by the model in terms of disease specific information are still quite demanding, and may limit the utility of the model.

The model is apparently publicly available from the WHO website, upon emailed request and a licensing agreement.

Further Reading

This article is available online at:
http://www.resource-allocation.com/content/1/1/6

The PopMod model is available by request from WHO here:
http://www.who.int/choice/toolkit/pop_mod/en/index.html

Article

Author	Macdonald AS, Waters HR, Wekwete CT
Title	**A Model for Coronary Heart Disease and Stroke with Applications to Critical Illness Insurance Underwriting** **I: The Model** **II: Applications**
Source	North American Actuarial Journal, 2005. Volume 9 (1):13-40 and 41-56

Context

Description	This pair of articles describes modeling rates of incidence (called intensities in the papers) for selected insurable events and applications of those models in critical illness insurance pricing. The stated motivation for developing the models is to estimate the impact of including genetic information in insurance underwriting.
Outcome of Interest	In Part I, the outcomes of interest are the rates of incidence of a coronary heart disease event or stroke in a population without a prior CHD event or stroke. In Part II, outcomes of interest are rates of incidence of cancer, kidney failure, or total and permanent disability and critical insurance premiums developed using all of the models.

Model

Type	All models are of the same form: a continuous time Markov model with a finite state space, where time is equivalent to a person's age. The essence of the papers is the development of the parameters for rates (λ) of incidences of insurable events in a population, assuming a Poisson distribution for the number of occurrences, and rates of transition between states. The rate of incidence is formulated as the exponential of a linear combination of factors: $\lambda = \exp(\alpha + \beta x + \gamma_s)$ where, for example, x is age and s is an indicator for sex. In Part I, models for transient states are developed (hypertension, hypercholesterolemia, diabetes) and absorbing state models for CHD events and stroke. For the CHD and stroke model, the factors are age, sex, smoking status, diabetes, BMI, hypercholesterolemia, and hypertension. In Part II, models for cancer, kidney failure, and total or permanent disability are developed. Subpopulations of insurance purchasers are defined by age, sex, and BMI category. Premiums for selected subpopulations are displayed, but the methodology used to develop the premiums from the rates of

incidence is not well described; however, a reference for the methodology is provided. Simulations for selected subpopulations (male; ages 35, 45, 55 at issue; normal BMI) are also displayed, but the description isn't thorough enough for the reader to understand what parameters were allowed to be random (sampled from a distribution).

The authors assumed a Poisson distribution for the number of occurrences of insurable event and estimated the Poisson parameter (the rate of incidence) using a generalized linear model with a log link to smooth the estimates. It wasn't clear how they actually did this – at first it appeared that they ran Poisson regressions, because they had the individual level data available, but after reading Part II, it seems as though they just started with occurrence/exposure rates for each of the cells and smoothed from there.

Software The authors did not describe the software used.

Model Quality

Data Sources The authors used multiple years of individual level data from the Farmingham Heart Study to estimate parameters for the hypertension, hypercholesterolemia, diabetes, stroke, and CHD event models. As the authors note, the data are from 1949 through 1987 and although causes of CHD and stroke have not changed over time, treatment for the causes (hypercholesterolemia and hyptertension) have changed and the rates of outcome (CHD event or stoke) have been modified through that treatment. Actuaries using the authors' models may want to consider changes in treatment before using the parameters as given.

Aggregate data sources were from the Office of National Statistics and Department of Health (England) were used to model rates of cancer incidence. The U.S. Renal Data System Annual Data Reports for 1997-2000 and previously published prevalence data were used for modeling incidence rates of ESRD. Mortality rates were estimated from the English Life Table No. 15.

Parameters In the simulation of premium rates for selected subpopulations (male; ages 35, 45, 55 at issue; normal BMI), the authors state that 49 parameters were random variables and assumed to have a multivariate normal joint distribution. Those parameters were sampled for each simulation run. Unfortunately, it isn't clear what those 49 parameters were and the authors do not provide justification to assume they have a multivariate normal joint distribution.

Simplifying Assumptions	The authors assumed that only a handful of variables were important predictors of CHD event or stroke: age, sex, smoking status, BMI, hypertension, hypercholesterolemia, and diabetes. They mentioned the value of including other variables, but recognized the usefulness of their work will be limited if actuaries do not have access to variables that might enter the models.
Duration / Time Perspective	Premiums for 10, 20, and 30 year term critical insurance policies were displayed.
Iterations per Scenario	10,000 iterations for each subpopulation (sex, age, BMI) and insurance term simulation.
Validation	Incidence rate models were validated against the Framingham data, Morbidity Statistics from General Practice, The Health Survey for England, and Hospital Episodes Statistics. The authors suggested adjustment to the rates based on their validation exercise.
Sensitivity Analysis	To estimate bounds on the effect of genetic information on premiums, the authors tested several scenarios. First, they tested the impact of 5 times the incidence rates of hypertension, hypercholesterolemia, and diabetes (each separately). Next, they tested the impact of 50 times the incidence rates of those same conditions. Finally, they tested the impact of 5 times the incidence rate for the end states: CHD events and stroke.

Evaluation

Strengths/ Weaknesses	Part I of this pair of articles is strengthened by the complete description of data used for the development of incidence rates. The description of the models developed in Part II are very sketchy in comparison. The stated motivation was to develop models that could be used to quantify the impact of including genetic information in insurance underwriting. For the events considered, genetic information is not available or its influence is confounded by the influence of the environment and health behaviors. Lacking such information, the authors ran several scenario tests to estimate what various increases might mean in terms of premium change. Those scenarios are useful to the extent they provide lower and upper bounds to the impact of using genetic information in pricing critical illness insurance. Although the authors provided detailed description in some sections of this pair of papers, in other sections their methods were described by one or two brief sentences. For example, a more complete description of the premium simulation would have been helpful.

Presentation of Results

The authors present the significant factors for each of the rate of incidence models. Validation results for those models was presented graphically. Tables were used to present premiums for selected subpopulations and insurance terms; similar tables were used to present the impact of scenario tests (5x or 50x rates of incidence) and the impact was shown as a percentage increase over the base case.

Interpretation of Results

The authors conclude: 1) the most significant risk for critical illness premium (the factor that has the greatest impact on premium) is diabetes, followed by hypertension (Stage II or III), 2) the combination of risks has an additive effect on premiums vs. a multiplicative effect, and 3) genetic links that operate only through risk factors (hypertension, hypercholesterolemia, diabetes) are unlikely to have a significant impact on premium levels, 4) their models can be used as underwriting tools.

Value to Decision Making

The authors develop these rates of incidence with the intention that they will be used to estimate the impact of genetic information on critical illness insurance premiums, however, the rates could be used for any purpose where the goal was to analyze the incidence of hypertension, hypercholesterolemia, diabetes, or the frequency of CHD events or strokes in a population. Macdonald et al. refer to similar models developed within the medical/health management sciences, in particular, they cite Babad et al. (2002), which is also included in this literature review. Macdonald et al. note three differences:

	Macdonald et al.	Babad et al.
Motivation	Quantify insurance risk	Intervention strategies to reduce risk of CHD
Model form	Continuous time Markov	Discrete-time Markov
Range of CHD outcomes	Only outcomes that trigger an insurance claim	Broader range of outcomes included

This comparison suggests that models, similar to these developed particularly for an actuarial audience, are being developed more generally in other disciplines and may be adapted for actuarial purposes.

Ease of Implementation

The authors purposely selected model variables that would be available to actuaries who were pricing critical illness insurance, therefore, the models should be easily implemented. However, an actuary should be cautious about directly applying the model parameters as developed by Macdonald et al. without validating that the parameterization is appropriate for the population likely to purchase the insurance product.

Further Reading

Angus Macdonald's home page with several publications available for downloading: http://www.ma.hw.ac.uk/~angus/
Framingham Heart Study web site:
http://www.nhlbi.nih.gov/about/framingham/

Article

Author	Marquis, MS
Title	**Adverse selection with multiple choice among health insurance plans: A simulation analysis**
Source	Journal of Health Economics, 1992. 11:129-151

Context

Description

The goal of this paper was to use simulation techniques to describe adverse selection behavior in two situations: a market with multiple health benefit coverage choices and a market for multiple supplementary insurance packages purchased to cover the cost sharing of a basic health benefit coverage. Adverse selection depends on whether the insured can accurately forecast his own risk and whether that forecast influences his coverage choice.

Outcome of Interest

The outcome of interest was the difference in benefit plan market share between community rated and experience rated insurer pricing. The author measured the extent of adverse selection by simulating the impact of coverage selection on market equilibrium – comparing the types of insurance coverage selected and the equilibrium premiums when the premium was community or experience rated. The premiums were assumed to be equivalent to the expenditure of the insured pool, plus administrative fee. Premiums which were larger under experience rating than under community rating were taken as evidence of adverse selection; people with greater health expenditures choose more generous coverage when it was community rated (when the premium was subsidized by people with lower health expenditures). Premiums were assumed to be either uniform or age/sex adjusted for the community and experience rating scenarios.

Model

Type

The author's description is that of a microsimulation model, although not described as such in the article. The model has 3 components: a plan choice model, a health expenditure model, and an insurer pricing model. Stochastic variables were used in the insurance plan choice and health expenditure. Families of various sizes, including individuals, were simulated to choose a health insurance plan (or supplemental policy); the choice was influenced by the premium, expected expenditures, and family demographics. A stochastic term for plan choice was drawn from a multivariate normal distribution. Actual health expenditures for

families were a function of plan and family attributes; they were simulated using a two equation model of demand: the first equation was for the probability of any expenditure and the second equation for the total annual expenditure for families with positive spending. Both equations had stochastic terms.

1,326 families were simulated.

Software The author did not describe software used for the simulation model.

Model Quality

Data Sources The author used data from the Rand Health Insurance Experiment to develop equations for health plan choice and health expenditure. In addition to expenditure data, the author used results from two HIE surveys. One survey asked HIE subjects to estimate their health care expenditures in the next 12 months the second asked subjects about their preference for stop-loss coverage. Preference for stop-loss was used to estimate the plan choice equation.

Parameters Parameters in the health plan choice and health expenditures equations were developed from the Rand HIE data.

Simplifying Assumptions The author assumed that insurance premiums cover all expenditures plus an administrative fee. Premiums were not adjusted for family risk. Four choices were simulated: 100% coverage, and three plans with 95% coinsurance and stop loss at $500, $1000, and $1500.

Health expenditures were simulated for the full (100%) coverage plan and then adjusted for the 95% coinsurance plans. Using results from previous work by Keeler et al. (cited in the article), the author assumed that the rate of spending with 95% cost-sharing was 55% that of the full coverage plan. Simulated full coverage expenditures were adjusted for each of the coinsurance plans to be 55% of the full coverage amount up to the stop-loss, and 100% of full-coverage thereafter.

The insurer pricing model assumed that the insurer set premiums to cover claims cost plus an administrative fee of 15% of claims cost. The pricing model assumed that the insurer did not adjust premiums for expected risks of covered lives.

Duration / Time Perspective The simulation continues until the premiums stabilize; authors report four iterations for stabilization. The simulation describes the adverse selection exhibited for a single year of health plan choice.

Iterations per Scenario	The author computed 50 iterations for each time period simulated. Each iteration represented a model solution; variation among the iterations was due to the stochastic terms in the plan choice and health expenditure equations.
Validation	The author suggested validation with the Rand Health Insurance Experiment data through a footnote; the validation discussion was minimal.
Sensitivity Analysis	The article did not address how sensitive the outcome could be to the model parameters.

Evaluation

Strengths/ Weaknesses	This article from 1992 addresses adverse selection, question that will always be relevant for pricing health actuaries. The article includes a nice summary of the Rand Health Insurance Experiment and describes how the data from the HIE was used to simulate adverse selection. Major components of the HIE data used for this analysis – survey data on expected expenditures and preference for stop-loss coverage – would not be available to an actuary attempting a similar analysis. However, an actuary could assume various levels of insured forecasting ability and preference for stop-loss, then estimate the degree of each required for adverse selection to have a significant impact in a market.
Presentation of Results	The author presents selection effects by showing the equilibrium premium for each of the four benefit plans simulated and the distribution of individuals or families in each plan. The change in premium and distribution from community rated to experience rated premium simulations is evidence of adverse selection.
Interpretation of Results	The simulation showed that individuals and families predicted to choose the most generous plans at community rated premiums were also predicted to have higher health care expenditures. For the family plans, adverse selection lead to experience-rated premiums for full coverage that were almost 40% higher than the community rated premium – which lead to zero demand for the full coverage plans. Individual and age/sex results were also described. When the pricing model was modified so that premiums were adjusted for age and sex, the adverse selection results were dampened, but not eliminated. Results for the supplementary insurance product were analyzed similarly. The simulation showed little adverse selection in the supplementary market, because supplementary insurance is generally underpriced relative to its value and therefore attractive to both high and low risks.

Value to Decision Making	The value of the simulation was the measure of adverse selection. Various health benefit designs could be tested using this type of model to further understand the impact of adverse selection, health plan demand, and utilization of health services.
Ease of Implementation	The microsimulation model is fairly straightforward, however the amount of effort required to develop the health plan choice equations and the health expenditure equations was significant.
Further Reading	None.

Article

Author	Muldoon JM, Stoddart GL
Title	**Publicly Financed Competition in Health Care Delivery**
Source	Journal of Health Economics, 1989. 8:313-338

Context

Description

The paper describes a microsimulation (although not stochastic) to evaluate the impact of a system of formal competition in representative market area in Ontario, Canada. Two types of financing were examined: fee-for-service and capitation. Public health insurance provided access without additional charges to the lower cost financing arrangement – the differences in average per capita cost between financing arrangements was translated into an "enrollment charge" and paid by the insured.

The difference in average per capita cost arises from different hospitalization rates and ambulatory charges in the two market sectors and the distribution of the total insured population between the two sectors.

Outcome of Interest

The outcome of interest was the ambulatory and hospital services used; the purpose of the model was to estimate the impact that competition would have on ambulatory and hospital expenditures.

Model

Type

The authors used a deterministic microsimulation model.

Software

The DYNAMO simulation package was used [Puch-Roberts Associates Inc. 1984]

Model Quality

Data Sources

Initial data values (except enrollment elasticity) were based on experience of Sault Ste. Marie, Ontario, where FFE and capitation modalities have existed for 25 years (at the time of this publication).

Parameters

This was a deterministic model; no parameters were random.

Initial rates were drawn from previous work by the author and based upon rates in Sault St. Marie, Ontario.

	Hospitalization utilization rates for the capitation market were set at 0.85 patient days per person and the rates for the FFS market were set at 1.21 patient days per person. Ambulatory care expenditure in the FFS market was initially assumed to be $102.68 per person; in the capitation market it was initially assumed to be $87.23. Initial market split was assumed to be 50/50. Enrollment elasticity was assumed to be -0.25, based on a synthesis of studies in the U.S.
Simplifying Assumptions	The model assumed that competition takes place within a single community of fixed size and health services could be obtained from either a fee-for-service or capitation arrangement. Hospital and ambulatory utilization were separated. Insureds enroll once per year and react to the enrollment charge; insureds can move between providers financed by either FFS or capitation. For the simulation, optional or additional types of services which might result in benefit competition between the two types of providers are ignored. Average health status of insureds and quality of care within each financing arrangement was assumed to be similar. Each new insured in a modality generates the average cost of the existing enrollees and capacity adjustments occur instantaneously. Insureds are price sensitive and therefore respond to changes in the enrollment charge.
Duration / Time Perspective	Model runs were 10 years; present values were computed using two discount rates (5% and 10%) and the results discounted at 10% were presented in the paper.
Iterations per Scenario	Because this was a deterministic microsimulation, only one iteration per scenario was required.
Validation	The authors did not discuss validation of their model results.
Sensitivity Analysis	A value of 1.03 was tested for the FFS hospitalization patient days per person. Alternative initial market splits were tested, ranging from 50% FFS to 95% FFS.

Enrollment price elasticity in the ranges -0.004 to -0.64 were tested.

Evaluation

Strengths/ Weaknesses

This was a simplistic modeling of the impact of FFS and capitation financing of health care on utilization and expenditure. The authors made several simplifying assumptions (see above) which make their results hard to defend, although their conclusions seem reasonable, especially with the knowledge we have gained in the 17 years since their work was published.

Presentation of Results

Authors present tables of results from their baseline run and five different modeling exercises. The authors show present value of health care expenditures, savings from baseline, and savings as a percentage of baseline. Results were presented using both 5% and 10% discount rates.

For one of the models, the authors show yearly detail (4 of 10 years) of savings attributable to various components (capitation, response to enrollment charge, etc.).

Interpretation of Results

The authors conclude that a capitation financing arrangement has the potential to decrease overall utilization and expenditures. This seems obvious to us now, along with the complicating considerations of provider response to capitation and risk taking and patient response to managed care's capitation methods.

Value to Decision Making

A model used by the authors has some value for decision making because it uses simplified modeling techniques to illustrate a long-term (10 year) impact of financing changes and consumer response to an enrollment charge. However, the simplifying assumptions leave the decision-maker with model results that were based on some very non-realistic assumptions.

Ease of Implementation

This model could be easily implemented by actuaries, and they may want to consider such a model as a first step into the microsimulation modeling world. However, actuaries may consider the benefits of this type of modeling only as educational, rather than an improvement of model sophistication over their current approaches.

The authors do provide a detailed index with variable definitions and model equations.

Article

Author	Warner, K.E., Smith, R.J., Smith, D.G., and Fries, B.E.
Title	**Health and Economic Implications of a Work-Site Smoking-Cessation Program: A Simulation Analysis**
Source	*Journal of Occupational & Environmental Medicine*, 1996. Volume 38(10):981-992

Context

Description

A simulation designed to estimate the financial impact and cost effectiveness of a simulated smoking cessation intervention in a hypothetical blue-collar (manufacturing) firm with an initial 10,000 employees, of whom 3160 are smokers at the outset of simulation. The model simulates the flow of employees through the firm (new workers enter, current workers leave, retire or die, current workers who are smokers quit or do not quit smoking).

The simulation can be run assuming the presence or absence of a work-site smoking cessation intervention program, and can be run in a "full-firm" mode (in which workers who leave, the firm are replaced), or in a "cohort" mode (in which an initial cohort of 10,000 workers is followed through the model, without replacement as they die, retire, or quit the firm). The full-firm model is run with annual intervals of change for 50 simulated years, whereas the cohort model is run over 80 simulated years to observe the full mortality experience of the cohort of workers.

Individual simulated employees are assigned age, gender, tenure with the firm and smoking characteristics at the outset of the simulation. Yearly information recorded on each simulated individual includes these characteristics (updated annually where appropriate), salary, and simulated smoking related costs.

Each smoker in the simulation is assigned a "background" annual probability of quitting smoking, based on age-specific population estimates of quit-rates. Quit rates are incremented up in the presence of the simulated cessation intervention program, according to the assumptions of participation rate, and quit rates among program participants.

The model is run twice, once with a smoking cessation intervention in place and once without, and results compared to assess the effect of the intervention. People move between smoking statuses based on

assumptions about quit rates and smoking prevalence; employees join and leave the firm by quitting, retiring or dying; and mortality rates are applied by age and smoking status.

Outcome of Interest

Primary model outcomes are the simulated costs of the intervention as well as several classes of benefits (e.g. smoking cessation, health, financial benefits to the simulated firm). From these elements, cost effectiveness and benefit-cost are derived.

Health benefits and cost-effectiveness of the smoking cessation program are based on model results from the cohort model – taking primarily the societal perspective. The full-firm model is used to estimate the "bottom line" results that would be of most interest from the firm perspective.

All model outcomes were discounted at 3.5% per year.

Model

Type

The model used here is a microsimulation, state transition model. The model is stochastic, discrete event, object oriented. "Objects" are current and former employees who are modeled as passing through the "events" of the simulation. The stochastic nature of the model pertains to the use of Monte Carlo techniques to assign the probabilities of certain events occurring to model objects (e.g. quitting smoking).

The model can be run on either a cohort basis or on a full firm model. The difference between these is that in a cohort model workers who died, retire, or leave the firm are not replaced where as in the full firm model, they are replaced according to specify parameters.

Software

Simulation was written in Objective-C language

Model Quality

Data Sources

- Background smoking quit rate based on historic rate of quitting among all US smokers (US DHHS Pub. No. (CDC) 90-8416).
- Permanent smoking cessation rate of 15% based on published estimates in AHCPR publication No. 96-0692).
- Probabilistic risk of dying based on age, sex, smoking status based on American Cancer Society's Cancer Prevention Study II – adjusted to reflect national rates of mortality.
- Age and sex specific chances of retiring based on data from Peracchi and Welch ("Labor force transitions of older men and women." J Labor Econ. 1994; 12:210-242).
- Probabilistic risk of leaving the firm based on age and sex specific data for 1991 from Bureau of Labor Statistics.

- Smoking status assigned to initial and new hire workers based on age and sex specific data on blue collar respondents from the 1990 National Health Interview Survey. NHIS also used to specify years quit among former smokers.
- Age and sex specific health care cost parameters based on National Medical Expenditures Survey II, adjusted to data on medical expenditures among employees of manufacturing firms based on US Chamber of Commerce's 1992 Employee Benefits Report.
- Differentiation of health care costs between smokers and non-smokers based on Bartlett et al. (Bartlett JC, Miller LS, Rice DP, Max WB – MMWR. 1994; 43:469-472).
- Absenteeism rates by age, sex, and smoking status based on NHIS.
- Age specific salary was based on data about manufacturing workers' earnings from US Dept. of Commerce. ("Money Income of Households, Families, and Persons in the United States: 1991. Bureau of the Census, Current Population Reports, Series P-60, No. 180).
-

Parameters

Age, sex, tenure with firm, smoking status, salary, health care costs, turnover rates, mortality rates, new hire rates, presence or absence of worksite smoking cessation program.

Simplifying Assumptions

- Male smokers assumed to have started smoking at age 18, female smokers at age 21.
- Current smoker participation rate in cessation program = 30%
- Background quit rate = 2.5%
- Quit rate = 15% among participants
- Cessation intervention cost per participant = $150
- Had to make assumptions about health care expenditures among the elderly because there wasn't enough empirical data
- Firm assumed to be self-insured for health care.
- Firm assumed to have a defined contribution pension plan.
- Smoking cessation rates assumed to not vary by age or sex.
- No effects of the cessation program on employees' family members.
- Some potential "intangible" benefits of smoking cessation program are not modeled (e.g. increased job satisfaction or firm loyalty).
- No control made for the possibility that those participating in health promotion programs may increase their health care utilization of primary and preventive care services.

Duration / Time Perspective

Multiple perspectives. The simulation model loops in annual intervals, and simulation results are presented for 1, 3, 5, 10, 25, 50 and 85 year durations. The cohort model is used to generate results from a primarily societal perspective and the "full-firm" model is used to generate results primarily from the firm's perspective.

Iterations per Scenario

Both models (cohort and full-firm) were run with 1000 iterations – each using a different random number seed for all probabilistic events. Sensitivity analyses were conducted using 500 iterations each.

Validation

The authors note that their model results are consistent with and "confirm" conventional wisdom that work-site smoking cessation programs likely generate economic benefits in excess of their costs. Beyond this, there appears to be no attempt to validate the model results through comparison with other real-world or simulated results.

Sensitivity Analysis

Twelve sensitivity analyses were performed, testing the impact of changing the value of variables representing relevant assumptions of the model. Typically, these values were altered one at a time, but the authors also present results of a "worst case" analysis in which multiple assumptions were varied so that the model would be purposefully biased against support of the findings in the "base case" analysis.

Sensitivity analyses suggested modest alterations of quantitative results of the models, but the authors indicate that in no situation did they alter the "essential qualitative findings." Cost per life-year saved ranged from roughly $500 to about $2300, and benefit-cost ratios all indicated that the work-site smoking cessation program would eventually show economic returns – with financial savings being observed within 5 years in all but the "worst case" scenario.

Evaluation

Strengths/ Weaknesses

Strengths:

- Incorporation of realistic aspects of work-sites such as modeling of worker turnover and following employees who quit as well as those who stay.
- Ability to model both a "societal" perspective as well as a firm based perspective and compare the two.
- Both short and long-term implications of a cessation program are modeled.
- The substantial impact of "background" rates of smoking quitting are explicitly included and demonstrated.
- Specifies a level of "reality" that is quite high by allowing age and sex specific transition rates in a number of places in the simulation.

Limitations:

- The most serious limitation of this model is identified by the authors as the omission of what implications the smoking cessation program may have on a defined benefit pension plan and the Social Security system. The benefits modeled (whether

to the firm or to society) may be offset partially, or completely, if deaths averted or life years saved ultimately end up substantially increasing pension or Social Security costs.

- Smoking quit rates are not age and sex specific.

Presentation of Results

Financial implications of covering smoking cessation services are presented in tabular form, first for the results of the cohort model run, and next for the results of the full-firm run.

For the cohort model, model output presented includes: number of smoking quits, deaths postponed, life-years saved, program costs, cost per quit, cost per death postponed, and cost per life-year saved for each of the 1, 3, 5, 10, 25, and 85 year time intervals.

For the full-firm model, model output presented includes: reduced medical expenditures for the various employee classes (workers, retirees, etc.), reductions in absenteeism, on-the-job productivity gains, reductions in life-insurance, total firm benefits, program cost, net benefit for the firm, and firms benefit-cost ratio, all for intervals 1, 3, 5, 10, 25, and 50 years.

Sensitivity analyses are also presented in tabular form, at least for "selected" summary outcomes (cost per smoking quit at intervals 5 and 85; cost per life-year saved at interval 85; and benefit-cost ratio at interval 5 and 50) resulting from changes in model assumptions.

Interpretation of Results

The authors state that their model findings are largely consistent with the conventional wisdom about worksite smoking cessation programs (that they generate financial benefits exceeding program costs). Specifically, the note that "the cost per life-year saved of $894 is consistent with the findings from other cost-effectiveness analyses of smoking cessation." (p. 16).

They also note that, because their models are run over a substantially longer period than previous studies have examined, they may have uncovered much larger eventual financial benefits of smoking cessation programs. They argue that prior studies have likely systematically underestimated the short-term costs of such programs while overestimating the short-term benefits, and have completely missed the longer term benefits.

Value to Decision Making

This model may be particularly useful as a decision aid regarding whether work-site smoking cessation programs may be a good investment for a firm. This is both because of the care and thoroughness of work and thought put into development of the simulation, as well as the presentation of output across a broad range of timeframes from short

to medium-term to long-term. The "full-firm" model may be particularly useful to decision makers.

Importantly, however, while the model results clearly suggest economic benefits resulting from investments in such a program, the results in some ways support NOT investing in such programs because the most sizable economic benefits took many years (5, 10, 20, etc.) before they were realized. This is clearly a timeframe far longer than that typically taken by most business organizations.

Ease of Implementation

It doesn't appear that this simulation program is publicly available for use by analysts other than the developers of the program.

Further Reading

Further information about the model and data sources used can be found here: http://www.umich.edu/~rwj/techappend/

Article

Author	Warner KE, Mendez D, Smith DG
Title	The Financial Implications of Coverage of Smoking Cessation Treatment by Managed Care Organizations
Source	*Inquiry*, Spring 2004. 41: 57-69.

Context

Description

A simulation designed to estimate the financial impact and cost effectiveness of smoking cessation treatment in a hypothetical managed care organization (MCO). The model follows MCO members annually for 30 years, beginning in 2000, recording their age, smoking status, and associated health expenditures. New members join the MCO and existing members exit either voluntarily or through death.
A hypothetical managed care population starting at 450,000, with age-sex distributions based on an actual managed care population of over ½ million that is not named.

Each smoker in the simulation is assigned an annual probability of quitting smoking, based on age-specific population estimates of quit-rates. Quit rates are incremented up in the case of the MCO providing cessation coverage, according to the assumptions of participation rate, and quit rates among program participants.

The model is run twice, once with coverage for a smoking cessation benefit and once without, and results compared to assess the effect of the benefit. People move between smoking statuses based on assumptions about quit rates and smoking prevalence; members join and leave the plan; and mortality rates are applied by age and smoking status.

Outcome of Interest

The primary outcome of interest is the net financial impact of the MCO's coverage of smoking cessation services. The three components of this are: cost of the cessation program itself, the net effect of program induced changes in health care expenditures, and incremental revenues to the MCO resulting from some members living longer, thereby continuing to pay premiums. Cost expressed as PMPM. The cost of cessation programs to society is also modeled as are life-years saved.

Model

Type

The model used in this paper is a state transition model (microsimulation). {So – looks like a Markov Chain model – need to check the original article to see whether stochastic elements or not. If so, then would also be a Monte Carlo model – so MCMC microsim.}

Software

The original model was estimated using nonlinear weighted least squares estimation, through use of a minimization routine (generalized reduced gradient – GRG – algorithm) implemented in the "Solver" component of the Microsoft EXCEL spreadsheet software program. This detail is not provided in the current paper, but in a prior publication more fully describing the model used here (See further reading section below).

Model Quality

Data Sources

- Dynamic model of smoking prevalence originally estimated from National Health Interview Survey smoking status data (Mendez et al, 1998).
- Enrollment and claims data for a single year (1996-97) from an actual managed care organization (that is not named)
- Relative risks of death by smoking status are from the American Cancer Society's Cancer Prevention Study II (Mendez and Warner 2002)
- Age-specific per capita 1997 health care expenditures are from a second large health care organization (not named)
- The proportion of total costs attributable to smoking, the cost per participant of smoking cessation treatment, and the program participation rates come from published studies

Parameters

Age, smoking status, health care expenditures, turnover rates, mortality rates, coverage of cessation treatment.

Simplifying Assumptions

- Current smoker participation rate in cessation program = 10%
- Quit rate = 15% among participants
- Cessation services cost $350
- Probability of coverage-induced smoking cessation increases linearly with age
- Had to make assumptions about health care expeditures among the elderly because there wasn't enough empirical data
- The rate of decline in smoking-related health care costs after cessation is proportional to the rate of decline in mortality (may not be true for some groups, like pregnant women)
- Illness patterns and health care costs are the same for smokers who quit on their own as those who quit using a covered benefit

- Seniors expenditures are set the same way as non-seniors (average expenditure plus 15% markup)
- Once members leave the plan they don't return
- Smokers who are trying to quit wouldn't pay for cessation services on their own

Smoker status change (quit, not quit) and MCO membership change (die, leave MCO, re-enroll) were stochastic parameters.

Duration/Time Perspective

Two perspectives are applied - health plan and societal.
Results are presented for multiple time perspectives from the medium to long-term - 5, 10, and 30 years.

Iterations per scenario

Not specified, but it can be assumed that, because the model loops over annual intervals, that there were 5, 10, and 30 iterations for the three time perspectives, respectively.

Validation

Results are compared to previous, real-world AND simulation studies and are of a similar magnitude, but show a higher net cost to cessation programs than other studies.

Sensitivity Analysis

In all scenarios modeled coverage of cessation services has a small net cost to an MCO; at 5 years it is $.61 PMPM. In the cohort version of the model run from a societal viewpoint, cessation services cost $4730 per life-year saved. The results are stable across many scenarios. Authors report that substantial changes in all model variables yielded changes in PMPM costs no greater than a factor of two.

Evaluation

Strengths/ Weaknesses

This is a relatively straightforward model that appears flexible, and provides useful estimates of anticipated costs (in meaningful metric - PMPM) of implementing smoking cessation programs in an MCO. Well-defined, and reasoned assumptions are applied.

Presentation of Results

Financial implications of covering smoking cessation services are presented in tabular form, both in terms of total costs, and PMPM costs, for each of the 5, 10, and 30 year periods. The outcomes so presented include cessation program costs, change in medical costs, change in MCO revenue, and total net costs.

Sensitivity analyses are also presented in tabular form, with "low" and "high" assumptions for each model parameter, and the resultant impact on PMPM costs for each of the 5, 10, and 30 year time intervals.

Interpretation of Results	The authors state that their conclusions are stable across many scenarios, and that their simulation results lead to the conclusions that, (1) smoking cessation programs will not be profit generators for MCOs, but neither will they be large cost drivers, and (2) such smoking cessation services appear to be highly cost-effective investments in the health of MCO members, and may thereby be viewed as a societal benefit, perhaps worthy of some level of public investment.
Value to Decision Making	It's useful for managed care organizations to know that they will probably not see a positive return on investment for covering cessation services, but that it is a good value from the perspective of improving population health.
Ease of Implementation	This model doesn't appear to require special software. The technical details are available from the authors on request as well as being available in their 1998 publication (See below).
Further Reading	The current paper applies an existing simulation model to a particular application. Further details about the model were previously published in 1998: Mendez, D., Warner, K.E., and Courant, P.N. Has Smoking Cessation Ceased? Expected Trends in the Prevalence of Smoking in the United States. *American Journal of Epidemiology*, 148(3): 249-258.

Article

Author	Naidoo, B., Thorogood, M., McPherson, K., and Gunning-Schepers, L.J.
Title	**Modelling the effects of increased physical activity on coronary heart disease in England and Wales**
Source	*Journal of Epidemiology and Community Health*, 1997. 51(April,#2): 144-150.

Context

Description

This paper presents the extension of an existing simulation model (PREVENT) to include physical activity as a risk factor as well as interventions to increase physical activity. The PREVENT model is a cell based, "population attributable risk based" model for simulating population level changes in coronary heart disease mortality as a function of hypothetical interventions aimed at changing disease risk factors.

The paper simulates two strategies for increasing population levels of physical activity in England and Wales and their impact over 25 years on CHD mortality. Strategy 1 was simulated to increase by 25% the proportion of people who are moderately active among those ages 15 to 64, while Strategy 2 simulated a similar increase in the proportion of people who were vigorously active in the same age range. Simulation results suggested small reductions in CHD death rates for both men and women as a consequence of both strategies. The simulated effect sizes suggested the strategies would result in postponement of just over 12,000 deaths over a 25 year period.

Differences by age and sex suggested that concentrating strategies to increase physical activity would have the largest impact if they were focused on: (1) men rather than women, (2) those over age 45, as opposed to those who were younger, and (3) the most sedentary individuals as opposed to those who were already getting at least some physical activity.

Outcome of Interest

Percentage reduction in CHD mortality rate (relative to 1994), and years of life gained as a result of two hypothetical interventions.

Model

Type

The authors describe the PREVENT model as a cell based simulation. Elsewhere, the PREVENT model has been described as a population attributable risk model. The population attributable risk (PAR) of a particular factor can be defined in this case as the number of CHD cases

due to the presence of the risk factor (in this case, lack of physical activity) divided by the total number of CHD cases in the population. The PREVENT model is focused on the impact of potential preventive interventions in reducing PAR, and so computes a related measure the authors refer to as the potential impact fraction (PIF). This is basically an estimate of how much a particular intervention can reduce the PAR of a particular risk factor for a given disease.

The model uses life table methods to project a hypothetical population forward in time, and can be set to model risk factor exposure prevalences as characteristics of either a cohort or age group. Input data requirements for the model are fairly simple, and include disease prevalences by five year age group, sex, and risk factor exposure category. For many populations (nations) these type of data are readily available for a range of disease states and risk factors.

Software

The software used to develop the PREVENT model is not discussed in the paper. The model was developed originally in 1988 by Louise J. Gunning-Schepers (last author on the paper).

Model Quality

Data Sources

Based on information obtained from a master's thesis that made use of the PREVENT model (see URL in Further Reading below), it is apparent that there are a number of types of input data needed for the model that were not discussed in the paper and sources for which were therefore not specified. These include: (1) sex-specific population structure by age, in one year increments from less than 1 through age 95+; (2), age and sex specific all cause mortality data; (3) absolute number of births projected to occur by sex over the number of years covered by the simulation; and (4) Disease-specific mortality – in terms of rates per 100,000 for 20 five-year age groups, by sex;

Parameters for which data sources were specified:
A meta-analysis of the impact of physical inactivity on CHD was used to specify the relative risks of CHD mortality attributable to physical inactivity (Berlin & Colditz, 1990. A meta-analysis of physical activity in the prevention of coronary heart disease, *Am J Epidemiol*, 132: 612-27.)

The age- and sex-specific relationships between physical inactivity in CHD risk was specified based on a 1991 paper (Shaper & Wannamethee, 1991, *Br Heart J*, 66: 384-94.)

Parameters

Prevalence of physical activity at different intensity levels in England and Wales by sex and five year age group. An inverse, graded

	relationship between CHD risk and physical activity was specified, as were assumed latency between onset of decline in relative risk from increased activity, and the lag time until relative risk reached it's lowest level. Relative risks of CHD mortality in relation to levels of physical activity. CHD mortality by sex and age group in absence of any hypothetical intervention. Age and sex-specific information on population structure, all cause mortality data, disease-specific mortality data for the specific disease of interest. Number of births projected (by sex) for the years covered by the simulation.
Simplifying Assumptions	Disease risk factors are assumed to be independent of one another. The risk of CHD due to physical activity is assumed to begin declining immediately upon someone's taking up an increased amount of physical activity, with the risk leveling off at one year, equal to the level of those obtaining that amount of physical activity. At that point, the person's past history of physical inactivity is assumed to have no further detrimental effect on outcomes (e.g. the model "forgets" history). Increases in physical activity were assumed in the model to affect only CHD death risk and not the risks of death attributable to other diseases. It is also assumed that physical activity does not affect the prevalence of other CHD risk factors such as hypertension, hyperlipidemia and obesity – a patently false assumption. The intervention was assumed to began in 1994 and continue for 11 years – with the population being simulated for an additional 14 years after the end of the intervention. It was assumed that the hypothetical interventions would change a given age groups behavior for the entire simulation period of 25 years, and that any changes in the prevalence of physical activity in the population would result exclusively from the intervention. Other risk factors in the population were assumed to be constant (i.e. have no trend).
Duration / Time Perspective	The maximum time frame over which the PREVENT simulation can run is 50 years. However, in this paper, the simulation was run over a 25 year period. Age groups up to age 95 are modeled. The PREVENT model can be run for an entire (synthetic) cohort or for specific age-groups. In this paper, the model was run for an entire cohort

Iterations per Scenario	This simulation does not appear to be an iteration based model. It appears to resemble more of a modified life table model.
Validation	There is no discussion in this paper of validation of either the PREVENT model or of the results of the simulation presented.
Sensitivity Analysis	The only sensitivity analyses conducted, were to run the simulation with an assumption that the relative risk of CHD mortality attributable to physical inactivity is lower for older individuals than younger individuals (see Presentation of Results below).

Evaluation

Strengths/ Weaknesses	The PREVENT model can incorporate latency between exposure to a risk factor and development of a particular disease. It can also incorporate lags in the model. The simulation allows the modeling of multiple diseases and multiple risk factors and was specifically designed for modeling hypothetical interventions at the population level. Outcomes that are obtainable from the model include disease specific morbidity and mortality and person-years of life gained or lost. The PREVENT model does not provide for cost-benefit analysis and requires a fair number of strong assumptions about the independence of risk factors and diseases. The PREVENT model does not allow for modeling of potential changes to CHD morbidity that may result from hypothetical interventions.
Presentation of Results	Results of a series of simulation runs were presented in graphical (bar charts) format. The figures plotted in these charts are also included in the paper appendix. The first chart shows life years gained for men and women under age 95, who achieved an increase in either moderate (strategy 1) or vigorous activity (strategy 2). The second chart shows, years of life gained for men and women under age 95, simulated to have achieved the respective physical activity goals of each strategy, with the strategies simulated to have been focused on one of two ages groups – either those 15-44 or those 44-64. Two additional figures were also presented (each representing results from one hypothetical strategy) showing the results of a sensitivity analysis that took into account the possibility that the relative risk of inactivity on CHD may be lower for older individuals than it is for younger individuals. The results presented in these graphs suggest that modifying the relative risks downward did not have much of an impact on the results.

Interpretation of Results

The authors interpret their results as suggesting that interventions to increase population levels of physical activity (in order to reduce CHD mortality) would best be focused on older men, and sedentary individuals, to have the largest impact.

The reasonableness of this interpretation needs to be evaluated in light of the assumptions and unknowns of the model as it was implemented. In particular, the model assumes that physical activity increase will have a nearly instantaneous impact on CHD mortality, but that the history of physical activity for an individual does not. This seems a fairly tenuous assumption. Moreover, as the authors themselves admit, the model assumes that physical activity does not have any impact on CHD risk factors such as hypertension, hyperlipidemia, or blood sugar. We know that this is not the case, and it is particularly through such risk factor mechanisms that longer-term physical activity may have a longer-term impact on CHD mortality-effects that are not captured in this model.

Value to Decision Making

The value of the results of this paper to decision makers is limited by some of the factors and assumptions listed above. However, to the extent that the model captures something about the immediate and direct impact of physical activity on CHD mortality, it may be useful in focusing attention on the potential benefits of increasing physical activity among sedentary, older men – a group that is an increasing proportion of the populations of many nations.

Ease of Implementation

Presumably, the PREVENT software may be obtainable from the last author of the paper. If so, the types of data needed to implement the model are generally widely available, and the model appears implementable by someone with an understanding of the specific disease state(s) and risk factor(s) being targeted by hypothetical interventions.

Further Reading

This article was published in 1997. The authors reference more recent development of the PREVENT model into something called NIMPH. The only evidence we have found online of this work was a project funded in the Netherlands between 1996 and 1998: http://cordis.europa.eu/data/PROJ_BIOMED/ACTIONeqDndSESSIONeq10611200595ndDOCeq2ndTBLeqEN_PROJ.htm

There is available online from Collections Canada, a master's thesis from 1997, evaluating the usefulness and applicability of the PREVENT model for assessing the impact of changes in alcohol policy on coronary heart disease: http://www.collectionscanada.ca/obj/s4/f2/dsk2/ftp04/mq29208.pdf

Article

Author	Palmer, A.J., Roze, S., Valentine, W.J., Minshall, M.E., Foos, V., Lurati, F.M, Lammert, M., and Spinas, G.A.
Title	**I: The CORE diabetes model: projecting long-term clinical outcomes, costs and cost-effectiveness of interventions in diabetes mellitus (types 1 and 2) to support clinical and reimbursement decision-making** **II: Validation of the CORE diabetes model against epidemiological and clinical studies**
Source	Current Medical Research and Opinions, 2004. Volume 20 (Suppl 1): S5-S26 and S27-S40

Context

Description

This pair of articles describes the construction and validation of a microdata simulation model of Types 1 and 2 diabetes mellitus. The model projects both clinical and costs outcomes for populations using a series of (Markov) sub models simulating multiple complications of diabetes. The model comes preloaded with data specifying cohort characteristics, which can be modified by the user as can the economic and clinical data on which the model is based, allowing users to update the model based on new information available in the literature or to experiment with hypothetical policy interventions and their impact on model parameters. The goal of the model is to provide long-term projections of clinical and economic outcomes based on the best currently available data and to allow the comparison of different diabetes management strategies in different patient populations. The model developers intended to be a support tool for clinical and financial decision-makers as well as those who develop interventions for diabetes treatment. The model appears to be a commercial product, available for use through licensing agreement with the developing organization based in Basel, Switzerland.

Outcome of Interest

Clinical outcomes of the model include sub models for the following complications: myocardial infarction, angina, congestive heart failure, stroke, peripheral vascular disease, neuropathy, foot ulcer, retinopathy, macular edema, cataract to, nephropathy, hypoglycemia, ketoacidosis, lactic acidosis, and nonspecific mortality. Economic outcomes of the models include direct costs, as well as indirect costs, discount rates, and health state and event utilities. Total costs and quality adjusted life expectancy are available from the model in different "perspectives" can be generated, e.g. societal cost perspective, versus third-party healthcare payer perspective.

Model

Type

{This can be described as an elaboration of the more typical Monte Carlo, Markov Chain, Microsimulations that we are reviewing. The most significant departure from more basic MCMC models is the inclusion of "tracker" variables to allow state transition probabilities to depend on past states as well as current state, thereby loosening a key assumption of typical Markov Chain models.}

A set of 15 sub models (each for a different diabetes related complication) comprises the overall simulation system, with each model representing a (presumably continuous time) Markov model with a finite state space. A 16th sub model simulates changes in type 2 diabetes treatment over time based on treatment failure or side effects of treatment. For each sub model, time-, state-, time-in-state and diabetes type-specific transition probabilities simulate the progress of individuals through different disease states.

The authors do not provide specific equations for all sub models in the CORE model, but an appendix to the first article presents the equations used in the hypoglycemia sub model. Specifically, the probability of first hypoglycemic event is estimated using a Poisson regression with the following form:

$$p = \exp(\alpha + \beta \times \ln(HbA_{1c}))$$

where the probability of an initial hypoglycemic event (p) is a function of the natural log of HbA_{1c} (a measure of blood sugar control), the absolute average risk per unit time of follow-up (α) and the risk gradient associated with a given HbA_{1c} value (β). Risk of recurrent hypoglycemic events is calculated based on a similar equation, but with the addition of a quadratic trend term for HbA_{1c}.

Two general limitations of Markov models are that they are "memory-less" (transition probabilities are a function of current state, but not prior states) and they require the definition of distinct and mutually exclusive disease states. In order to overcome these limitations, the authors have paired each Markov model with Monte Carlo analysis and "tracker variables" (indicator variables that toggle from zero to one when a particular complication is simulated to have occurred) to provide the models with "memory" and the ability for the simultaneous development of multiple complications. Because the model allows for dependencies and interactions between different types of complications, for each model cycle, the order in which sub models is run is randomly changed.

After all baseline parameter settings have been defined, each of the 1,000 unique, individual "patients" is simulated over time. Real-world

probabilities of dying, progressing to a higher level in severity of disease, or of developing a particular complication are compared to random numbers generated from a uniform distribution between zero and one. If the random number drawn is less than or equal to the probability of a given event, that event is assigned as having occurred. The CORE model being a highly detailed simulation of diabetes and its complications, the number of these probabilities is large.

Software

The CORE diabetes model was designed to be used as an Internet based application through license from CORE (Center for outcomes research). As such, the model contains four separate components of 1) user interface, 2) a system of related input databases, 3), the mathematical engine of the system, and 4) and an output database.

The user interface was programmed in hypertext markup language (HTML) of the input and output databases are based on a structured query language database (SQL). The mathematical engine performed the calculations for the overall simulation was programmed in C++ (Microsoft® Visual Studio 6.0, Enterprise Edition).

Model Quality

Data Sources

Data sources used in the CORE model are extensive, vary across sub models, and in some cases can be selected by the user him/herself. In the case of the myocardial infarction sub model, for instance, users can select whether to use Framingham risk function data or a risk engine based on the UKPDS study, a large diabetes prevention trial conducted in the UK. The DCCT (Diabetes Control and Complications Trial) forms a third major source of data for the CORE simulation used to specify transition probabilities across multiple sub models. Where these major epidemiologic trials have left unanswered questions or unknown, transition probabilities, the CORE authors have carefully filled in these gaps by searching the literature for other, less well known studies that provide some information for specifying these probabilities. In all cases, the authors have carefully documented the data sources used for all aspects of the model, such that a knowledgeable reader can assess the quality of the sources.

Parameters

The number and breadth of potential parameters to the specified in the CORE model is far too extensive to list in this review. The model is preloaded with plausible default values for all model parameters, such that a user can simply run the model as specified. Nearly all parameters, however, are also subject to user specification, either in an experimental fashion, or based on updated clinical information available since the model was developed. Four input databases allow for the definition of a)

	cohort characteristics (e.g. age, gender, diabetes duration, baseline levels of lipids, blood-pressure, blood glucose levels, etc.), b) clinical factors (e.g. risk factor progression transition probabilities), c) treatment factors (e.g. screening rates and treatment impact on risk factor progression) and d) economic parameters.
Simplifying Assumptions	Because this is a highly detailed, my cousin relation model, there are very few simplifying assumptions that are explicit. Real-world data from a variety of clinical trials are used in setting the model parameters and transition probabilities. To the extent that the results from these clinical trials are not completely generalizable to real-world settings beyond the experimental setting (i.e. are deficient in their external validity) the results of the simulation will also not fully reflect the real-world.
Duration / Time Perspective	Most generally speaking, the CORE model is designed to simulate the long-term future (up to a maximum of 50 years in the budget impact analysis capacity). The duration of simulation runs is under user control. The model can be run assuming either closed or open cohorts. So, for instance, in estimating a closed cohort simulation, the model begins with a cohort with defined characteristics and runs until either all members of the cohort are simulated to have died or until a user-defined time horizon is reached. The simulation progresses at fixed, discrete time intervals, typically with cycles of one year in length. Exceptions to this are the foot ulcers sub model and the hypoglycemia sub model, which cycle at intervals of one month and three months respectively.
Iterations per Scenario	First order Monte Carlo methods are used to simulate each probability in the model with 1,000 simulations run on 1,000 unique individuals, after which 1,000 bootstrap samples are drawn in order to evaluate uncertainty in the cost effectiveness outcomes in particular, as well as in the mean effectiveness gain associated with a particular simulated intervention.
Validation	Developers of the CORE model have subjected it to extensive validation. To begin with, the model was designed and programmed by an experienced team, including experts in diabetes and health economics, who reviewed and validated the overall model and all sub models as they were developed. In order to identify and weed out programming errors and inconsistencies, the simulation tool was initially programmed using two different software tools (Data Pro decision analysis software (Tree Age Software Inc., Williamstown, MA) and C++ (Microsoft® Visual Studio 6.0, Enterprise Edition), with

inconsistencies in results from each used to identify and correct programming errors.

The authors report that 66 internal and external validation analyses were conducted across a range of competitions and outcomes from the model. Internal validation analyses were those comparing model results to come up trials data that were used in the development of the model. By contrast, external validation analyses were those comparing model results with epidemiological clinical data, which had not been used in the development of the CORE model. Overall, the model has been validated against 11 published studies that encompass the most well-known diabetes trials and epidemiologic studies in existence.

Sensitivity Analysis

There's no indication in the articles that specific sensitivity analyses were conducted.

Evaluation

Strengths/ Weaknesses

Because the CORE model was developed for commercial use, and at the expense of the CORE organization itself, this is a highly detailed model, with strong validation at multiple levels. The model was developed using an extensive array of data sources, and it's large number of parameters make it extremely flexible, and versatile. The design, development, and logic of the model are carefully documented in these two articles, and this information is further supplemented by more extensive information about the model available from an online website (see URLs below).

A principal weakness of the model is that the user must obtain a license from the CORE organization in order to use it and there are fees associated with that licensing and use.

The model is well adapted to testing variations in individual-level medical treatments for diabetes, but it is not clear how one might make use the model to evaluate population-level interventions, particularly those that are not directly aimed at treating diabetes, but which would be expected to have an impact on disease progress, complications, and outcomes – e.g. impact of a national-level legislative policy aimed at reducing tobacco use. (Compare this with the Patten, 2002 simulation model, which is setup to test variations in population-level depression treatment, but which is less well adapted to testing individual-level interventions.)

Presentation of Results

The authors present summary information from the 66 validation analyses in tabular form in the second article. Results are presented primarily in terms of distributive statistics, and where possible, as plots

of the cumulative incidence of a given transition or complication comparing CORE model results with true study results. Validation runs were conducted over the same time frames as the comparator trials were run, and with cohort characteristics specified as closely as possible to those of the comparator study.

In the second article, plots comparing model results with those from true studies are presented for a myocardial infarction, nephropathy, retinopathy, neuropathy, and overall survival

Closeness of fit between the CORE model results, and study results was assessed by plotting outcomes predicted from each, fitting a linear curve through these points, with an intercept at zero, and obtaining the correlation coefficient (R^2). The overall fit between CORE results and published trial results was quite close - $R^2 = 0.9222$ (where 1.0 would indicate a perfect fit). CORE model results seem to fit slightly better for type 1 diabetes ($R^2 = .9778$) than for type 2 diabetes ($R^2 = 0.8861$). Moreover, as would be expected, correlations were higher for the internal validations than for the external validations ($R^2 = 0.9574$ and 0.9023, respectively).

Interpretation of Results

The authors conclude: the CORE Diabetes Model accurately represents real-life results from clinical trials and epidemiologic studies. The flexibility of the model makes it a useful tool for comparing diabetes management strategies in cohorts that have varying characteristics and across different types of clinical settings.

Value to Decision Making

The very high level of detail and craftsmanship that have been put into this model make it a very flexible tool to support decisions in diabetes management and care. Because the model generates both clinical outcomes and economic outcomes from a variety of perspectives, one can envision the usefulness to decision makers charged with making financial decisions about how to spend health-care dollars in providing diabetes care, as well as to clinical decision makers confronting how to best manage a population of diabetes patients. Because the model embodies the best of what is known today about diabetes in terms of its treatment, costs, and complications, decision-makers can have a high level of confidence that the model results are truly informative. Moreover, because the model can be updated as new knowledge is generated from clinical trials and epidemiologic studies, validation of the model will continue into the future with the likelihood that the model will remain valid and useful and keep pace with existing knowledge.

Ease of Implementation	Because the CORE model was designed to be used as an off-the-shelf Internet based application, it appears to be quite easy to use and relatively straightforward to modify, provided a user is willing and able to pay for the licensing to use the model.
Further Reading	A 52 page PDF file, providing an in-depth overview of the CORE model is available for download from the main CORE web site here: http://www.thecenter.ch/cdm/cdm.asp http://www.thecenter.ch/download/cdm.pdf Complete details of all validation analyses of the model are also available here: http://www.thecenter.ch/cdmappendices/

Article

Author	Patten, S.B.
Title	**A framework for describing the impact of antidepressant medications on population health status**
Source	*Pharmacoepidemiology and Drug Safety*, 2002. 11: 549-559.

Context

Description

The primary purpose of this work is to construct a simulation model to assess the efficiency of different treatment strategies in reducing the population prevalence of depression. In particular, the model is used to assess whether a population-wide increase in anti-depressant treatment utilization would reduce the population prevalence of depression more or less efficiently than targeted increases in treatment of those with recurrent episodes. The model simulates a cohort of 1000 individuals, starting at age 15, experiencing entries to, and exits from major depressive episodes.

Recognizing that some episodes of depression resolve readily without treatment, some resolve reasonably well with treatment, and some episodes are more recalcitrant, even with treatment, the author defines four subtypes of major depression in the model: a) good prognosis of recovery, b) intermediate prognosis, even if untreated, c) intermediate prognosis if treated, and d) poor prognosis of recovery.

After the model was constructed and validated against real-world data, the author ran a series of simulations, altering assumptions about the proportion of the population represented by the two intermediate prognostic groups and varying the proportions simulated to be receiving antidepressant treatment.

The principal finding from the simulation runs is that broad changes in the rate of simulated treatment utilization result in relatively small changes in point prevalence at the population level if treatment is assumed to be given for both new (incident) cases and recurrent cases of depression. By contrast, a second series of simulations shows that reducing rates of recurrence among individuals with three or more prior episodes would be expected to much more dramatically reduce the point prevalence in the population.

Outcome of Interest

The question being addressed in this article is whether the point prevalence of depression is reduced more effectively by increasing antidepressant use overall in a population or by focusing increased long-

term treatment on those with recurrent depression. Point prevalence is approximated by using lifetime sick day proportion (LSP). The author argues that the answer from their model is that it is more beneficial to focus long-term treatment on people with recurrent depression.

Model

Type

The basic modeling strategy here can be seen as an elaboration of an epidemiologic incidence-prevalence model in which the prevalence of a condition (in this case major depression) is viewed as a "stock" in the population with incident cases represent an "inflow" to that stock with "outflow" taking place through either recovery from the condition or through mortality.

The central equation in the model appears to be that used to calculate the probability density function of recovery from a depressive episode. The equation is specified as:

$$f(t) = \sum_{j=1}^{n} p_j - p_j e^{-\lambda_j t}$$

where j represents one of four groups (a – d) with different prognoses of recovery from depression, p_j represents the proportion of the cohort in group j, and λ_j represents the rate of recovery for those in group j. The proportion of cases of depression falling into group (a) (good prognosis) was specified as 0.30, with a recovery rate of 0.45/week. The proportion in the intermediate prognostic groups (b) & (c) (treated + untreated) was 0.60 with recovery rates of 0.11/week and 0.07/week in the treated and untreated subgroups, respectively. The proportion of cases falling into the poor prognosis group (d) was set at 0.10 with a recovery rate of 0.02/week.

The basic incidence-prevalence model is elaborated in this work to take into account some of the unique aspects of major depression. These include the fact that it is difficult to define a single incidence rate for depression since recurrent episodes are common in these may best be characterized as recurrences rather than new incidences. In addition, there may be multiple recovery rates from major depression and some cases appear to be chronic and recalcitrant to treatment.

In this simulation, the initial flow of (incident) cases is from the never depressed state into one of the four prognostic depression group states, with the proportion entering each group being the probability p_j. Outflows from these four states is determined by the recovery rates specified by λ_j. Those recovering from an initial depressive episode are

then "at risk" of making a transition into a first recurrence, following the same parameters as for the initial occurrence, and likewise up to three episodes total.

Differential equations are used to characterize flows between states within the model. An arbitrary cohort of 1000 hypothetical individuals starting at age 15 is specified, with 171 of these presumed to have major depression at baseline. Inflows to the depressed state and outflows from depressed to either recovered or deceased are simulated forward for this cohort over 960 subsequent months - up to age 95 years.

Software

Numerical solutions for the model were obtained using Scientist for Windows (MicroMath Scientific Software, Salt Lake City, Utah, 1995). Scientist is a numerical integration program designed for fitting model equations to experimental data. (see: http://www.micromath.com/)

Model Quality

Data Sources

The two primary sources of data used to parameterize the model were the Canadian National Population Health Survey, 1994-95 (NPHS) and the 1994 U.S. National Comorbidity Survey. The NPHS provided a measure to identify individuals with high probability of having had an episode of major depression during the prior year - the Composite International Diagnostic Interview Short Form for Major Depression (CIDI-SFMD). Data from the U.S. National Comorbidity Survey indicate that subjects scoring four or more on the CIDI-SFMD had an 80% probability of having had a major depressive episode in the prior year. To further refine this measure to a point prevalence estimate, subjects were only considered to have a currently active depressive episode if they also reported "more than usual distress" on a distress scale.

The proportion of these individuals in the NPHS who also reported being on antidepressant medication at the time of the survey was used to estimate the antidepressant utilization rate.

Data from the NPHS on the duration of major depressive episodes was used to estimate rates of recovery from depression.

Mortality was parameterized in the model based on data from a 1999 systematic review of mortality and depression by Wulsin et al.. Depression recurrence rates were parameterized based on data from reviews conducted by Thase (1999) and Keller and Hanks (1994).

Parameters

All key model parameters (which have been described throughout this review) were set based on either estimates derived from the primary

data sources (e.g. Canadian NPHS, U.S. NCS), from literature review of pertinent publications, or in the case of mortality, on a polynomial regression model fit to age-specific mortality rates from Canadian data.

Simplifying Assumptions

Because the simulation of the cohort is based on cross-sectional data of a sample of people across the age range of adulthood (a synthetic cohort approach), there is an implicit stable population assumption that the age-specific experience of depression in the sample would be constant across true age cohorts.

Several model parameters are based on fairly crude estimates, some of which based on the cross-sectional survey data from the, NPHS and others of which were derived from published review articles. Depression was conceived of in this model, as an either/or condition. Rates of recovery from depression in the face of treatment were estimated to occur uniformly over time. The mortality rate of those with major depression is modeled by simply multiplying the general population rate by a factor of 1.7. Recovery rates from depression were estimated based on the reported duration of depressive episodes in the NPHS. The likelihood of a relapse resulting in an individual falling into any one of the four prognostic groups was assumed to be independent of the duration of previous depressive episodes. Finally, recovery rates from depression were not conditioned on episode number.

Duration / Time Perspective

Models describe the hypothetical person-time experience of major depression of a hypothetical cohort of 1,000 individuals over their lifetime, starting at age 15 through death or age 95.

Iterations per Scenario

The authors are silent about number of iterations per simulated scenario; we can assume that each scenario is one iteration of a 1,000 member cohort.

Validation

Initial simulation was conducted specifying incident rates of depression based on external data (from the Stirling County Study – Murphy JM, et al., Psychol Med 2000; 30: 505-514) and assuming 1,000 non-depressed individuals with no prior episodes of depression. The author notes that the resulting simulation to a poor job of capturing age-specific prevalence rates (presumably based on the NPHS data). Based on this information, the author specified subsequent simulations using a cohort assumed to have 17.1% in a first remission from depression, and no new incident cases. Simulations based on these parameterization's result in age-specific prevalence rates that are similar to other published data and estimates of lifetime experience of depression (in terms of person months depressed) comparable to the observed prevalence of depression in the NPHS.

Sensitivity Analysis There is no indication that any sensitivity analyses were conducted.

Evaluation

Strengths/ Weaknesses

The article clearly describes the development of the model, the sources of data used and the assumptions built in. This is a relatively straightforward modeling exercise with a relatively modest amount of detail built in to the simulation. The author is forthright about the fact that many of the parameterizations are based on crude estimates. To the extent that any key model parameters are based on cross-sectional data from the Canadian NPHS, the generalizability of the overall simulation model may be questioned. The simplifying assumptions appear plausible, but the work is light on validation and appears to be absent any sensitivity analyses.

The model is specifically setup to test variations in population-level interventions pertaining to depression treatment. The population-health focus precludes the use of this model to test treatment variations at the individual level. (Compare this with the CORE diabetes simulation model, Palmer, 2004, which is setup to test variations in individual-level medical treatments for diabetes, but is less well adapted to testing population-level interventions.)

Presentation of Results

Results are presented only graphically in the paper. Four figures are included. Figure 1 depicts the proportion of a sample recovering from depression by a month, based on the NPHS data and compared to estimates derived from equation used to simulate the probability density function of recovery from depressive episode. (See above) Figure 2 depicts the simulated person months of survival in a cohort of 1000 individuals starting at age 15, along with the simulated person months lived with depression in the cohort. The ratio of these two curves represents the Lifetime Sickday Proportion (LSP) which was used in this work to approximate point prevalence of depression. Figure 3, depicts the simulated prevalence of depression, by the proportion of the intermediate prognostic group that is simulated to have received treatment. Figure 4 depicts the simulated prevalence of depression in relation to changes in recurrence rates among those with three or more episodes of depression.

Interpretation of Results

The author indicates in the discussion section that: “Simulations using the model suggest that efforts to increase treatment utilization, in isolation, will not have a major impact on the point prevalence of major depression. On the other hand, targeted efforts at increasing utilization of long-term preventive treatment in those with highly recurrent disorders, would be expected to have a much more dramatic impact."

	He goes on to suggest that the results suggest value in changing public health strategies toward a chronic disease model – emphasizing long-term management – and away from the acute treatment model.
Value to Decision Making	The primary contribution this work makes to decision making is to suggest that, under the assumptions of the simulation, a larger population health impact in terms of reduced prevalence of major depression might be obtained through intensive (selective?) treatment to reduce depression relapse rates among those with highly recurrent depression.
Ease of Implementation	Adequate information on variable parameterization, equation specification, data sources and methods are provided in the paper that would presumably allow one with access to the Scientist for Windows software to replicate and modify the simulation model constructed here.
Further Reading	None.

Article

Author	Tengs TO, Osgood ND, Lin TH – Medical Care article Tengs TO, Osgood ND, Chen LL – Preventive Medicine article
Title / Source	**Public Health Impact of Changes in Smoking Behavior Results from the Tobacco Policy Model** / Medical Care, 2001. 39(10):1131-1141
Title / Source	**The Cost-Effectiveness of Intensive National School-Based Anti-Tobacco Education: Results from the Tobacco Policy Model** / Preventive Medicine, 2001. 33:558-570

Context

Description

The Tobacco Policy Model is a system dynamics model of the US population. It simulates birth, death, aging, and changes in smoking behavior. Transitions are assumed to occur annually; transition probabilities can vary by age, gender, smoking status, exposure to nicotine *in utero*, and year.

Outcome of Interest

Medical Care article outcome is interest in Quality Adjusted Life Years (QALYs), in particular, the change in QALYs that results from various scenarios of tobacco related behaviors. The authors compare the QALYs from a base scenario (current rates of smoking initiation, cessation, and relapse) with three modified scenarios: 1) a 10% decrease in rates of initiation, 2) a 10% increase in rates of cessation, and 3) a 10% decrease in rates of relapse among former smokers.

Preventive Medicine article outcome is cost-effectiveness measured in dollars per QALY. The authors estimate the incremental change in life-years, QALYs, and medical costs for the entire U.S. population as a result of school-based intensive anti-tobacco education over 25 and 50 years.

Model

Type

The Tobacco Policy Model is a Markovian system dynamics computer simulation model.

Software

The model was constructed using Vensim 4.0 software.

Model Quality

Data Sources

Several data sources were used; variables and data sources are:

- US population age and gender – US Bureau of Census
- Number of current, former, and never smokers – Behavioral Risk Factors Surveillance Survey and the National School-Based Youth Risk Behavior Survey and the Teenager Attitude and Practice Survey II (TAPS II)
- Probability of smoking initiation – Current Population Survey, Tobacco Use Supplement
- Probability of cessation and relapse – National Health Interview Survey and TAPS II
- Probability of live birth – Census
- Mortality data for current, former, and never smokers – 1992 National Health and Nutrition Examination Survey I, Epidemiologic Follow-up Study (assumed infants born to mothers who smoke during pregnancy have a 58% higher infant mortality rate than if mother did not smoke), Census
- Quality of life implications of health problems due to smoking – estimates for adults derived from the Quality of Well Being Scale, estimates for children derived from the health status from National Health Interview Survey combined with the Health Utilities Index.
- Cost of intensive national educational program (for Preventive Medicine article) estimated from "Towards No Tobacco Use" Project. Costs include to time train educators, educator teaching time, and classroom materials.
- Cost of medical care (for the Preventive Medicine article) were obtained from Hodgson, published in Milbank Quarterly, 1992. Ultimate sources were National Health Interview Survey, National Nursing Home Survey, NHANES Epidemiologic Followup Study, the National Medical Care Utilization and Expenditure Survey, and Medicare data files.

Parameters

Rates of smoking initiation, cessation, and relapse are the main stochastic parameters of this model. Population demographics are developed with birth and death parameters.

Simplifying Assumptions

The authors state several simplifying assumptions: the main scenarios assume that 30% of the mortality differential between smokers and nonsmokers is due to smoking (the sensitivity of this assumption was tested); number of years smoked, time since a former smoker quit, and quantity smoked is not included in the model; social or environmental factors are not included; effects of second hand smoke are not modeled.

	Total population medical costs (Preventive Medicine article) were computed by multiplying the number of persons in each group (age, sex, smoking status) by the average cost for that group, then summing over all groups. Annual costs were discounted at 3% annual rate for present value calculations. Effectiveness of intensive anti-smoking education programs (Preventive Medicine article) was estimated by "the reduction of increase in the prevalence of current smoking from the point at which the program began to that year." The authors considered three estimates at the end of the 2-year program: 5%, 30%, and 56% and 2 estimates of dissipation of effectiveness. Therefore, they modeled six scenarios (3 effectiveness estimates X 2 dissipation estimates). The model, as described in the Preventive Medicine article, did not estimate the longer term impact of reducing teenage smoking initiation, such as potential to influence peers and therefore further reduce initiation rates and long term impact on future generations due to non-smoking parents (present day teenagers). In addition, changes in cost, mortality, and quality of life due to exposure to second hand smoke were not modeled.
Duration / Time Perspective	The model projects 100 years.
Iterations per Scenario	The authors do not describe the number of iterations per scenario; results are presented as point estimates, suggesting that each scenario consisted of one iteration.
Validation	The authors do not discuss validation of the model. However, they recommend that the relative difference between the base scenario and the alternative scenarios is the meaningful result (vs. absolute outcomes).
Sensitivity Analysis	Medical Care article described use of 5,000 iterations to test the sensitivity of the mortality attribution assumption (30% of mortality differential between smokers and non-smokers due to smoking). In sensitivity testing, the attribution percentage was assumed to be normal with a mean of 0.03, a standard deviation of 0.08, and a range of 0.00-0.60. Monte Carlo simulation was used for the sensitivity analysis. Preventive Medicine article described sensitivity analysis for uncertainty of medical costs, quality of life, and mortality parameters. In all cases, model outcomes display less variation in early projection years and more variation (uncertainty of outcome) in later projection years.

Evaluation

Strengths/ Weaknesses	The articles provides a fairly good description of the data sources and parameters used in the Tobacco Policy Model, but further detail on the model is lacking. More detail on parameter values and sources is presented in the Preventive Medicine article. The pair of articles provides very good examples of how a system dynamics model can be used to evaluate various policy changes.
Presentation of Results	Medical Care article scenario results are presented simply – either as a table of cumulative population QALY increases (discounted at 3% per year) or as a graph of population QALY gains per year over the 100 year projection. The author's description of results is clear and supported by the tables and graphs: young people are most benefited by policies to reduce initiation, middle-aged people are most benefited by cessation policies, and the oldest age group receives most benefit by policies which reduce relapse. Preventive Medicine article scenario results are also presented in tables of present values of main outcomes; intensive education programs were not cost saving.
Interpretation of Results	The authors provide a nice discussion focusing on the timing of results; the graphs and discussion highlight the long delay in impact from current policy change. They also stress that the potential impact of various strategies is dependent on the size of the population targeted. For example, policies that discourage initiation have little impact on the segment of the population ages 60-69 because few in that age group take up smoking – policies aimed at reducing relapse are most effective in that age group.
Value to Decision Making	The value of the Tobacco Policy Model for decision-making is its simplicity and long time horizon; QALYs are the only outcome.
Ease of Implementation	It appears to be a very simple model with only a handful of demographic parameters. The model is initialized with a population representing the U.S. population in 2001, by age, gender, smoking status (current, former, never). Three changes in smoking behavior were modeled: initiation (from never to current), cessation (from current to former), and relapse (from former to current).

Article

Author	Wolfson, MC
Title	**POHEM – a framework for understanding and modeling the health of human populations**
Source	World Health Statistics Quarterly, 1994. 47:157-176

Context

Description

Population Health Model (POHEM) is a microsimulation of a national population. It projects the life course of individuals and their family units using a variety of socio-economic and health status characteristics. The model was developed as a tool to create information beyond the usual health statistics related to resource use; it was developed to create information about the Canadian population's health status. In addition to its value in measuring population health, the model is a tool for uniting disparate health information, supporting resource allocation decision-making, and carrying out health science research.

Outcome of Interest

POHEM projects individuals and their families from birth to death. Many population level outcomes can be measured. Examples described in the paper are: disease specific average age at onset and the percentage of the population affected, life expectancy, and healthy life expectancy. Diseases modeled include lung cancer, heart disease, and arthritis (this list has expanded since the article was published). Outcomes are presented for males and females separately.

Model

Type

POHEM is a microsimulation model − it creates simulated individuals and family structures (male-female pairs and children) and projects the full life cycle of the family. POHEM includes several state variables, which represent individual characteristics, and probability functions for movement between states.

Software

The author did not describe the type of software used to develop POHEM, but he did comment that POHEM runs on standard personal computers.

Model Quality

Data Sources	The author lists several state variables and the sources used for initial values and transition probabilities. The states are: socio-economic status, health risk factors, diseases, functional status, costs (for lung cancer module), and health. Several variables are described for each state. Data sources include national statistics and published research; the sources are too many to be listed in this review.
Parameters	The author implies that most, if not all, state variables are stochastic, however the model is not described in enough detail to confirm.
Simplifying Assumptions	In this 1994 publication, the author describes POHEM as a work in progress. Modules for heart disease and lung cancer were developed first; the author states that a breast cancer module was underway (recent publications suggested additional modules have been completed). The author's goal was a broader perspective on health and health-related processes, and he writes that the disease-oriented simulation processes are a limitation. However, the processes which can be modeled are those that have been described by epidemiological studies – study availability is a limiting constraint. Where data were lacking, assumptions were developed through expert opinion, consensus plan, or re-analysis of existing data.
Duration / Time Perspective	Individuals are projected for a lifetime, up to age 95. Family units are projected through the death of the last adult member. The perspective is of the population.
Iterations per Scenario	This article describes the model and its applications; it does not discuss iterations or sample size per scenario.
Validation	The author does not describe validation of the POHEM model, however, there is discussion about the reasonableness of model outcomes and inconsistencies that may develop when using various assumptions. The author considers identification of such inconsistencies to be of value because it identifies the need for further understanding of the phenomenon at the population level.
Sensitivity Analysis	The author is silent on sensitivity analysis.

Evaluation

Strengths/ Weaknesses	This is a very complete article that describes the impetus for model development, the reasons for choosing microsimulation as the model type, a description of microsimulation in general, and examples of how the POHEM model can be used to answer questions related to resource allocation. The article is weak in its detailed description of parameter values and transition probabilities, however, given the age of the article, the interested reader may find more recent detail by accessing the

Statistics Canada Web site (given below under further reading).

The article states that cost data were only available for the lung cancer module, however, more recent work suggests additional disease modules with cost data.

Presentation of Results

The article is most valuable as a description of microsimulation modeling and its application to population health. The article does include illustrations to highlight population health status measures, chronic disease burdens, statistical coherence (inconsistent results), health interventions (cholesterol lowering interventions), and health research applications. One important message of the illustrations is the ability to see overall population impacts of an intervention beyond the impact of the intervention on a particular condition. For example, cholesterol lowering interventions will likely have a positive impact on heart disease mortality, but the overall population impact is shown when all causes of death are considered.

Interpretation of Results

The purpose of the article is to describe POHEM and its potential; interpretation of results for a particular application of the model are sketchy.

Value to Decision Making

The author describes several key roles for POHEM: 1) to produce summary indicators of population health, 2) to provide coherence to health information, 3) to support decision-making, and 4) as a tool for basic science research.

Ease of Implementation

A model the scope of POHEM has many data needs. For the model to be applied to a particular population, descriptors and transition probabilities should be modified for the population. The author comments that POHEM was designed for the Canadian population, but with the intention that other nations could modify parameters to fit their population. A similar modification could be made for a smaller insured population, with careful selection of population parameters.

Further Reading

The article is old (published in 1994), but the Statistics Canada Web site (http://www.statcan.ca/english/spsd/Pohem.htm) includes updated information about POHEM and a contact email address for the interested reader. A number of published articles related to POHEM are listed below.

Published Articles related to POHEM

Maroun J, Ng E, Berthelot JM, Le Petit C, Dahrouge S, Flanagan WM, Walker H,Evans WK. Lifetime costs of colon and rectal cancer management in Canada. Chronic Dis Can. 2003 Fall;24(4):91-101.

Flanagan WM, Le Petit C, Berthelot JM, White KJ, Coombs BA, Jones-McLean E. Potential impact of population-based colorectal cancer screening in Canada. Chronic Dis Can. 2003 Fall;24(4):81-8.

Will BP, Nobrega KM, Berthelot JM, Flanagan W, Wolfson MC, Logan DM, Evans WK. First do no harm: extending the debate on the provision of preventive tamoxifen. Br J Cancer. 2001 Nov 2;85(9):1280-8.

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 Sep;37(14):1797-804.

Will BP, Berthelot JM, Le Petit C, Tomiak EM, Verma S, Evans WK. Estimates of the lifetime costs of breast cancer treatment in Canada. Eur J Cancer. 2000 Apr;36(6):724-35.

Berkowitz N, Gupta S, Silberman G. Estimates of the lifetime direct costs of treatment for metastatic breast cancer. Value Health. 2000 Jan-Feb;3(1):23-30.

Earle CC, Evans WK. Cost-effectivenes of paclitaxel plus cisplatin in advanced non-small-cell lung cancer. Br J Cancer. 1999 May;80(5-6):815-20.

Will BP, Le Petit C, Berthelot JM, Tomiak EM, Verma S, Evans WK. Diagnostic and therapeutic approaches for nonmetastatic breast cancer in Canada, and their associated costs. Br J Cancer. 1999 Mar;79(9-10):1428-36.

Evans WK. Cost-effectiveness of vinorelbine alone or vinorelbine plus cisplatin for stage IV NSCLC. Oncology (Williston Park). 1998 Mar;12(3 Suppl 4):18-25; discussion 25-6.

Earle CC, Evans WK. A comparison of the costs of paclitaxel and best supportive care in stage IV non-small-cell lung cancer. Cancer Prev Control. 1997 Oct;1(4):282-8.

Evans WK, Le Chevalier T. The cost-effectiveness of navelbine alone or in combination with cisplatin in comparison to other chemotherapy regimens and best supportive care in stage IV non-small cell lung cancer. Eur J Cancer. 1996 Dec;32A(13):2249-55.

Evans WK. An estimate of the cost effectiveness of gemcitabine in stage IV non-small cell lung cancer. Semin Oncol. 1996 Oct;23(5 Suppl 10):82-9.

Evans WK, Will BP, Berthelot JM, Wolfson MC. Estimating the cost of lung cancer diagnosis and treatment in Canada: the POHEM model. Can J Oncol. 1995 Dec;5(4):408-19.

Evans WK, Will BP, Berthelot JM, Wolfson MC. The cost of managing lung cancer in Canada. Oncology (Williston Park). 1995 Nov;9(11 Suppl):147-53. Review.

Sapirie S. What does "health futures" mean to WHO and the world?

World Health Stat Q. 1994;47(3-4):98-100.

Wolfson MC. POHEM--a framework for understanding and modelling the health of human populations. World Health Stat Q. 1994;47(3-4):157-76.

Tugwell P, Chambers L, Torrance G, Reynolds D, Wolfson M, Bennett K, Badley E, Jamieson E, Stock S. The population health impact of arthritis. POHEM Workshop Group. J Rheumatol. 1993 Jun;20(6):1048-51.

Article

Author	Yang Z, Gilleskie DB, Norton EC
Title	**Prescription drugs, Medical care, and Health outcomes: a model of elderly health dynamics**
Source	National Bureau of Economic Research Working Paper 10964, December 2004

Context

Description

The authors present two unanswered questions about the Medicare Prescription Drug Bill: 1) will it improve the health of elderly Americans? 2) what will it cost? The authors address the complicated relationship among drug use, morbidity, mortality, and health care expenditures. Short-term spending may decrease if drug coverage increases the use of health improving drugs which leads to a decrease in hospitalization and associated expenditures. However, long-term spending may increase if drug coverage decreases mortality, which may increase lifetime medical care spending. Therefore, the authors address not only the change in drug expenditures, but also the impact of expanded drug coverage on total health care expenditures.

Outcome of Interest

The authors measure the change in drug expenditures due to an expanded prescription drug coverage by modeling annual individual utilization and health transitions over time. Their estimate is an aggregate increase in drug expenditures of 12%-17% over 5 years, which they note is smaller than estimates based on static models that do not include the consequences of increased drug use on morbidity, mortality, and total health care expenditures.

Model

Type

The authors are economists. They estimate a system of equations that represent supplemental insurance coverage, dynamic drug and other medical care demand, and health production. They describe their modeling effort as a dynamic model with correlated errors and say that it is an appropriate modeling technique "when studying complex behavior over time where changes in the composition of individual characteristics is associated with the behavior of interest." Their description of the simulation modeling leads us to conclude it is a microsimulation of the 14,439 individuals in the data sample.

Software	The authors do not describe the software they used.

Model Quality

Data Sources	The authors used publicly available data from the Medicare Current Beneficiary Survey (MCBS), longitudinal individual-level data from 1992-1998. The authors describe the MCBS and which variables they use for their modeling effort. Their final sample included 14,439 people with 42,174 person-year observations. All expenditures in the sample were adjusted to 1998 dollars.
Parameters	The model includes time-dependent variables for health status, supplemental insurance coverage, medical care consumption (prescription drugs, hospitalization, and physician services), demographic information, and health shocks. Health transitions are dynamic – they depend on lagged values of health status, medical expenditures, demographic information, and health shocks. Health status had six categories: no functional impairment, any IADLs only, 1 or 2 ADLs, 3 or 4 ADLs, 5 or 6 ADLs, and death. Health shocks were represented by five diseases or injuries: cancer, heart disease, cerebrovascular diseases, respiratory system diseases, and hip and other body part fractures. The stochastic represent the unobserved heterogeneity, both permanent and time-varying.
Simplifying Assumptions	The authors assume that all elderly persons are eligible for Medicare drug coverage. They categorize supplemental health insurance coverage as either 1) Medicaid, 2) any private insurance with a drug benefit, and 3) any private insurance without a drug benefit. They do not model changes in health insurance over time.
Duration / Time Perspective	The simulations are run for a five year duration.
Iterations per Scenario	They generate 400 simulations for each individual per scenario, stochastic variables were from the unobserved heterogeneity distributions (both permanent and time-varying).
Validation	The authors validate their model by comparing model results with the data used to develop it (MCBS sample).

Sensitivity Analysis

The authors simulate three scenarios: 1) no supplemental insurance beyond Medicare -- no drug coverage; 2) coverage by Medicaid, and 3) coverage by private insurance with a drug benefit.

Evaluation

Strengths/ Weaknesses

The authors provide a complete literature review including the state of drug coverage in the elderly population before the Medicare drug plan, the demand for prescription drugs by the elderly, and the impact of prescription drugs on the health of the elderly. The literature review sets up the need for the type of model they estimate, a dynamic model with correlated errors.

The article includes detailed discussion on unobserved heterogeneity and correlated errors across the system of equations. The authors also present the system of equations that they model. However, the discussion of heterogeneity, correlated errors, and model estimation may be beyond the reach of an actuary who without a familiarity of econometrics.

The value of the article is the discussion around the need for dynamic modeling to answer complex policy questions and the presentation of one method to do so on a topic of great interest to actuaries: Medicare drug coverage.

Presentation of Results

The authors present all results in tables. They provide descriptive statistics of the sample used to develop the models. They also provide parameter estimates from several equations, but the authors admit that translating parameter estimates into effects is difficult and so they prefer to explain effects using the simulation model results.

Simulation model results are presented in tables and discussed at a high level in the paper.

Interpretation of Results

The authors note that the increase in health care spending after drug coverage is introduced comes from individuals who experience health declines, yet survive. People who would have survived regardless of the drug benefit increase their drug expenditures a moderate amount; people who would have died if no drug benefit were available to them spend over 50% more on drugs and over 20% more on hospital and physician services.

Value to Decision Making

The stochastic simulation presented in this paper offers value over more common static modeling because it recognizes the potential for health care to affect morbidity and mortality, therefore causing long-term impact on health and health care expenditures.

Ease of Implementation	An actuary may find implementation of this econometric modeling to be beyond his or her training.
Further Reading	None.