SIMULATION MODELING OF OUTCOMES AND COST EFFECTIVENESS

Scott D. Ramsey, MD, PhD, Martin McIntosh, PhD, Ruth Etzioni, PhD, and Nicole Urban, ScD

Mathematical modeling techniques are now an integral component of researchers' efforts to find better ways to reduce the burden of cancer. Models are useful when investigators attempt to choose among a number of interventions or treatment options. When the direct effect of the intervention on the outcome of interest is unknown but information about intermediate components of the process is available, models are a tool for integrating the available knowledge to estimate the outcome.

Mathematical models commonly are used now to assess the relative effectiveness and cost effectiveness of different cancer screening and treatment strategies. The results may be used to guide policy decisions, to inform clinical practice, or even to assist in the design of clinical studies. Models may also be used to understand the reasons underlying population trends in disease prevalence, to highlight areas where more research is needed, and to shed light on disease characteristics.

Funding for this research was provided by Career Development Award-Grants K01-CA7-6189 (SDR); the National Cancer Institute, N01-PC-8506320 (RE); from NCI, and DAMD 17-94-J-4237 (MM, NU); and from the Department of Defense/US Army Medical Research and Development Command.

From the Cancer Prevention Research Program, Division of Public Health Sciences (SDR, MM, RE, NU), Fred Hutchinson Cancer Research Center; and Division of General Medicine, Department of Medicine (SDR), Department of Biostatistics (MM), and Department of Health Services, School of Public Health and Community Medicine (NU), University of Washington, Seattle, Washington

This article reviews the use and basic structure of models and enumerates ways to assess their quality as decision aids. The authors hope to provide an awareness of the scope of the modeling task, an understanding of the potential usefulness of models, and an ability to assess the quality of predictions derived from models.

WHAT IS A MODEL?

A model is a structured representation linking potential actions with their consequences. It is, effectively, a formalization of the steps or health states between an intervention and an outcome. Models consist of both structure and content. The structure is the formalization of the process of interest, whereas the content is the data used to inform the model. Once the structure of the model has been defined, the likelihood of progression from one health state to the next is used to determine the likelihood that the outcome of interest will occur under different scenarios.

USES OF MODELS

Models are frequently used to estimate the efficacy or cost effectiveness of interventions in the absence of data from controlled clinical trials. In studying prostate cancer, for example, models have been developed to assess the expected costs and benefits of prostate-specific antigen (PSA) screening in Medicare recipients,[11] to compare annual and biannual screening strategies,[9] and to project the outcomes of surgery, radiation therapy, or watchful waiting for localized disease.[5] Models have also been used to compare alternative screening strategies for ovarian, breast, colorectal, and cervical cancer.

As the field of mathematical modeling has evolved, models have been applied in areas other than cost effectiveness analysis. A number of studies have used models to investigate the reasons behind changing population trends in the prevalence of breast and prostate cancer. Etzioni et al modeled the role of PSA testing in recent declines in prostate mortality in the United States and concluded that testing could plausibly explain the observed declines only when extreme assumptions were made about the natural history of the disease.[10] A study by the Microsimulation Screen Analysis (MISCAN) group used a model to estimate that 70% of the observed reduction in breast cancer mortality in England and Wales is attributable to mammographic screening.[26] This group also predicted that in the Netherlands, mammography will reduce breast cancer mortality by 18% in 1999 and by 29% in the long term for women aged 55 to 74 years.[26]

Models are also being used to aid in the design of clinical trials. Indeed, analytic models may ultimately be more valuable than typical power calculations, the traditional aid in clinical trial design. Unlike typical power calculations, which use only simple concepts of effect,

size, and other factors, models allow researchers to examine a large variety of factors that may affect the number of subjects needed and the outcome of a clinical trial. For example, Urban et al have evaluated the cost effectiveness of many ovarian cancer screening options that have not been evaluated through cancer screening trials and have predicted their performance when evaluated in a trial.[25] The models are particularly useful because the cost effectiveness of these designs, if applied as public policy, can be estimated at the same time. Urban et al have learned that the most efficient trial designs are not necessarily the most cost effective.

Finally, models may be used to shed light on disease characteristics. One of the applications of the MISCAN breast cancer model was to estimate the mean duration of the preclinical stage of the disease and the sensitivity of mammography.[28] The model developers varied the values of these parameters until the model-predicted mortality reduction caused by mammography matched the results of a randomized trial conducted in Sweden.

MODEL STRUCTURE

There are several different types of models. A recent publication differentiates between macro- and micro-level models.[19] Macro-level models move groups or populations through different health states using standard, population-based statistics such as disease incidence and mortality rates. In micro-level models, individuals are the basic modeling units, so micro-level models are able to capture disease histories at the individual level.

Decision trees and Markov models are two commonly used macro-level models. A decision tree is a sequential graph that presents the problem as a sequence of decisions and consequences. For example, Ortega and colleagues constructed a decision tree to estimate the cost effectiveness of the combination of paclitaxel and cisplatin compared with cyclophosphamide and cisplatin for advanced ovarian cancer[22] (Fig. 1). The alternative decisions (represented in Fig. 1 by a square decision node) are the two chemotherapy regimens. After the decision, patients follow a decision path where there is a probability of either death or survival (shown with a round chance node). Survival is further divided into survival with response or survival with no response. The ends of the tree at the far right are end health states, which are assigned values (in this case, utility values determined by a survey of patients and healthy controls). To estimate the expected value of the alternative initial strategies, the trees are folded back or averaged out by multiplying outcome values by the probability of the outcome and summing across all possible outcomes for each decision choice. Costs and outcomes are compared for each alternative strategy.[30]

Markov models are frequently used to summarize more complex decision problems. In decision analysis, a Markov model is used to try

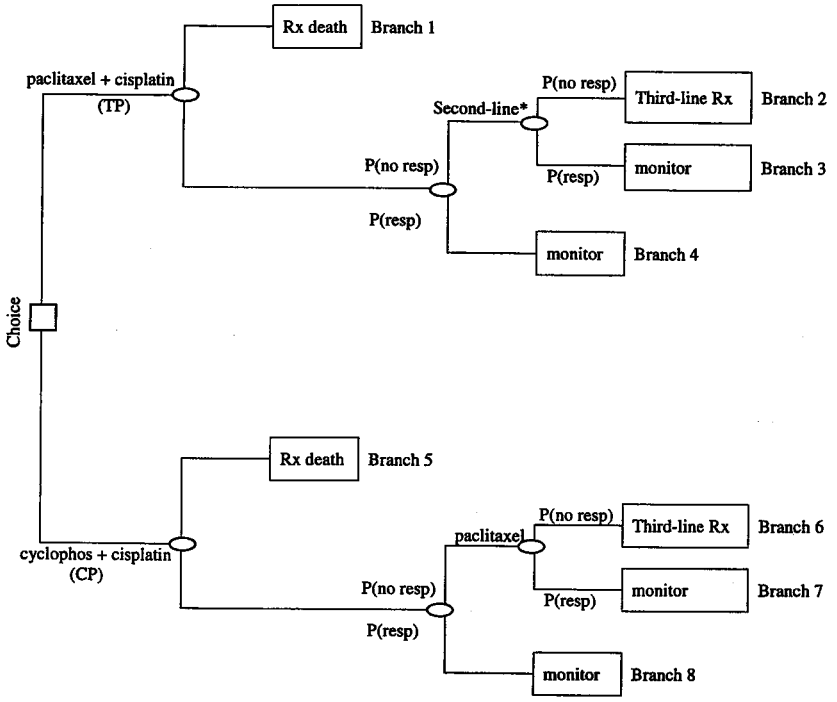

Figure 1. Decision tree. Paclitaxel and cisplatin versus cisplatin and cyclophosphamide as first-line therapy for ovarian cancer. *Second-line therapy = ifosfamide, tamoxifen, altretamine.

to represent more accurately complex processes that involve transitions in and out of various states of health.[2] Cancer patients, for example, may have relapses and remissions after the diagnosis is made. Markov models allow decision analysts to track patients as they move in and out of distinct health states over time. For example, Ng and colleagues constructed a Markov model to compare the life expectancy and the quality-adjusted life expectancy of patients with early-stage, favorable-prognosis Hodgkin's disease managed with and without staging laparotomy.[20] Several health states were considered, including refractory disease, remission, relapse, secondary leukemia, secondary solid tumor, secondary non-Hodgkin's lymphoma, and death (Fig. 2).

Although decision trees and Markov models are appropriate for many decision problems, their usefulness is limited in certain applications. First, Markov models tend to be event-driven simulations with relatively inflexible time structures. The modeler must often ignore the disease process to make the model fit observed events. This inflexibility limits the realism of the model in reflecting the natural course of the disease. Moreover, the structure of Markov models does not allow the analysis of interaction among cofactors (e.g., patient, environment, treat-

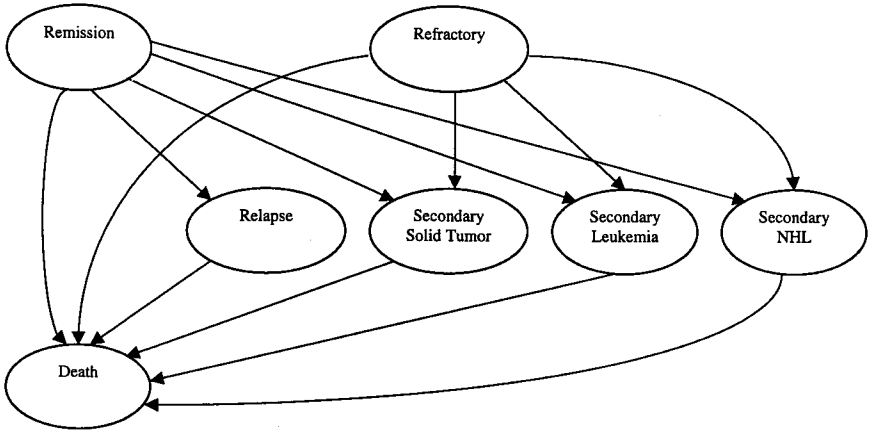

Figure 2. Markov model.

ment) that may be important. Markov models are often built so that transition out of one state is independent of the prior transitions, an assumption that may be questionable for many medical situations. Because of these limitations, microsimulation models have become popular for addressing many issues in cancer prevention and treatment.[8, 27]

Microsimulation models consider populations as collections of individuals with varying characteristics. An example is the Ovarian Cancer Screening Simulation program, a comprehensive representation of ovarian cancer biology, detection, screening behavior, interventions, and costs in a simulation of a defined population of women.[25] The model represents individual women, whose ages and risk factors for cancer vary as they enter the model. The likelihood of an ovarian tumor occurring and its detection through screening vary, depending on the characteristics of the individual and the intervention that is being considered. Similarly, the PSA Screening Simulation model represents individual men, some of whom develop prostate cancer, who progress through the pathologic stages of disease, and whose disease is either detected without PSA screening or who die of other causes. The PSA screening is then superimposed according to a defined schedule, which may depend on the characteristics of the individual.[11] The likelihood of detection by PSA screening depends on the PSA level, which grows as the disease develops. The rate of PSA growth is correlated with the rate of disease progression and changes more slowly in men with slowly progressing disease than in men with aggressive disease.

To predict the outcome of a particular intervention in a population, microsimulation models are run many times, so a large number of life histories for individuals who have slightly different characteristics as they enter the model can be aggregated and analyzed. An important advantage of microsimulation modeling is the level of detail modeled,

which often makes computing the probability of a number of different outcomes straightforward, even if they were not originally deemed to be of interest. For example, although the PSA Screening Simulation model was designed to estimate the years of life saved by screening, it was also possible to estimate the number of deaths prevented and to project the rate of overdiagnosis of prostate cancer resulting from screening.[10]

EVALUATING UNCERTAINTY IN MODELING

Like clinical trials, the results generated by models contain uncertainty. In clinical trials, the uncertainty is usually quantified in a confidence interval, or a *P* value. Uncertainties in simulation models are a bit more complex. In modeling, both the structure and the content of the model contain uncertainty. Uncertainty relating to the content of the model is known as *parameter uncertainty.* Parameter uncertainty arises because the data used to inform the model are not known with certainty. Often, many different values of the same parameter may all be plausible. Uncertainty about the structure of the model is known as *model uncertainty.* Model uncertainty arises because the structure of the model is only the modeler's best estimate of how a chain of events progresses in the real world. For example, the modeler may be uncertain whether tumors respond to a particular chemotherapy regimen in a linear relationship to increasing dose or whether the effects of the therapy decrease as the dose is raised. Finally, microsimulation models contain a third type of uncertainty known as *seed uncertainty.* Microsimulation models are run multiple times, each with a different starting point, or seed. For example, the seed may be a group of individuals with a certain number of characteristics, such as age, gender, and risk factors for developing a particular cancer. Because the individual characteristics are created by random draw from a range of possibilities for each characteristic, the model will generate slightly different results each time it is run. Thus, the predictions of a single model run represent only one of several possible predictions.

Parameter, model, and seed uncertainty each affect the results generated by the model in important ways. How each can be assessed is discussed briefly here. An overview addressing many of these aspects for food stamp microsimulation models is described in detail by Thurston and Zaslavsky[24] and Zaslavsky and Thurston.[31]

Parameter Uncertainty

Models contain a number of input parameters, most or all of which cannot be known with complete certainty. Some parameters may be informed by published literature; other parameters may have completely subjective sources. Modelers address this uncertainly through a process

known as *sensitivity analysis*. Weinstein and Feinberg describe sensitivity analysis as "systematically varying the different structural assumptions that are built into the decision tree (i.e., factors included or excluded) and the numerical assessments (i.e., probabilities and utilities) to see if the conclusions change."[30] If varying the values for the parameters that are used to inform the model does not appreciably affect the results, this stability provides reassurance about the robustness of the model. For example, the cost effectiveness of screening for ovarian cancer is not very sensitive to assumptions about the magnitude of the savings in treatment costs attributable to earlier detection.[25]

Sensitivity analysis is an important first step towards analyzing the robustness of a model, but it has limitations. First, the upper and lower limit for each parameter are often selected subjectively, reflecting the modeler's assumptions about the best and worst case rather than the true possible range. Second, varying an individual input parameter singly while holding all others constant (known as one-way sensitivity analysis) ignores important interactions among parameters. Third, modelers may choose some parameters as known or fixed (and thus not include them in the sensitivity analysis), when in fact they are not.

In general, one-way sensitivity analysis will produce a falsely low estimate of the uncertainty concerning the estimates generated by the model. To avoid this problem of overconfidence in the results, some decision analysts now advocate reflecting parameter uncertainty with multivariate, probabilistic sensitivity analysis. The analysts must first describe a distributional shape reflecting their belief of the true values for each of the input parameters, rather than giving only upper and lower limits. For example, instead of setting the upper and lower range of ages for individuals in a cancer screening program, the analyst would specify a distribution that would reflect the expected numbers of individuals in each age range for the entire population. Next, the analyst runs the model many times (e.g., 1000), each time making a random draw from each parameter's distribution. The parameters can be drawn one at a time for a probabilistic one-way sensitivity analysis or simultaneously for a multiway sensitivity analysis.[5] This process will produce a large number of individual predictions, each somewhat different from the others, and will produce a histogram of predictions that is similar to the range of outcomes observed for patients in a large clinical trial. The modeler can then compute means and standard deviations from the histogram, just as is done for a clinical trial.

Model Uncertainty

Model uncertainty is perhaps the most difficult uncertainty to evaluate. Because most models are built de novo to answer a specific clinical or policy question, no formal method for evaluating model uncertainty has yet been developed. Researchers should take care to present their representations and to justify their adequacy using sound theories. If

there is substantial doubt about the form of the model, restructuring the model and re-running the analysis is appropriate. If changing the structure does not substantially alter the results, this consistency increases confidence that the outcomes will not be greatly altered by changes in this aspect of the model.

Seed Uncertainty

Seed uncertainty can be addressed in a relatively simple way. To determine whether the results generated for a single cohort of individuals would change with a new cohort, the modeler runs the analysis several times, each time using the same number of patients in the cohort. This repetition in essence simulates the procedure of conducting multiple clinical trials, each with the same number of subjects but with the characteristics of the individual subjects varying slightly from trial to trial. The mean outcome and standard deviation around the outcome can then be derived from this fixed effects "meta-analysis" of simulation trials. Confidence is increased if, for the population being predicted, there is little variability between trials.

Model Validation

All three types of uncertainty about models can be evaluated together through a process that is commonly known as *model validation.* Model validation is a process of configuring the model to represent a population studied in an external observational or randomized study, running the model, and determining whether the results generated by the model match the real data. The ability of the model to make predictions that are corroborated by external data is called *external validity.*

The process used to validate the Ovarian Cancer Screening Model[25a] can illustrate model validation. Although large-scale clinical trials of ovarian cancer screening have not been conducted, a pilot study has been performed that includes a prevalence screening followed by a randomized, controlled trial in 22,000 post-menopausal women.[16] The ovarian cancer model was configured to match the sample size, age, and cancer incidence of the clinical trial and the clinical trial protocol (screening frequency, prevalence screen, follow-up, and so forth) in the actual study. The model was then run 1000 times, producing 1000 subject-specific estimates of important outcomes, such as the number of screen-detected cancers, their stage, and their lead time. The values of the 1000 replications were compared with the results from the actual trial. If a large proportion of the outcomes generated by each of the 1000 runs are close to those observed in the trial, the model is said to have validity. If some important feature of the trial is not reproduced, the model is not validated, and content and structure should be re-evaluated to resolve the conflict. In this way, model validation is much like hypothesis

testing: one tests to see if the model is so different from the trial that one must reject the hypothesis that the model is correct.

One limitation of this type of validation is that it depends largely on the number and size of external studies that are available. A single, small clinical trial is itself subject to uncertainty regarding the degree to which the results reflect random variation versus a true effect of the intervention. For the Ovarian Cancer Screening Model, further validation will be useful when larger screening trials are completed.

MODELING CONSIDERATIONS WHEN ESTIMATING THE COST EFFECTIVENESS OF INTERVENTIONS FOR CANCER

Cost-effectiveness analysis (CEA) can be defined as a set of research methods to assess and quantify the costs and clinical consequences of medical care treatments to estimate the economic value of the treatment in relation to alternative treatments.[7] Cost-effectiveness analysis is based on well-tested microeconomic theories of human behavior.[12] A CEA of medical treatments should incorporate evidence on the clinical consequences (efficacy and safety) and the costs and relative cost effectiveness of treatment alternatives.[1]

Because CEA integrates methods and data from clinical medicine, economics, epidemiology, statistics, and quality-of-life research, some degree of modeling is almost always required.[4] Clinical trials in cancer are by necessity carefully controlled situations with a limited time horizon, and they typically contain little economic data or quality-of-life endpoints. In contrast, CEA seeks to determine costs and benefits in typical clinical settings. Interestingly, the steps that are used to transform data with high internal validity (clinical trials) to an evaluation with high external validity are a central component of the controversy over cost-effectiveness studies.[21] O'Brien has referred to this issue as the "Frankenstein's monster" form of economic evaluation, referring to the fact that cost-effectiveness models typically bring disparate information together to form the "monster" (model) that is expected to behave predictably.[21]

Although economic evaluations incorporate data generated from clinical studies, the distinct objectives of CEAs determine the structure and content of the models needed to conduct these analyses. The major issues that modelers must address for CEA include extrapolation, translating intermediate endpoints to final outcomes, incorporating the concepts of quality of life and cost into the analysis, and adjusting outcomes data obtained from carefully controlled clinical trials to outcomes that are more likely to occur in standard practice (Table 1). Each of these issues is discussed briefly here.

Clinical trials are of limited duration, often weeks or months. Cost-effectiveness analyses, in contrast, evaluate the effects of different interventions during the time horizon in which the disease and treatment

Table 1. MODELING ISSUES BETWEEN CLINICAL STUDIES AND ECONOMIC EVALUATIONS

Issue	Clinical Study	Economic Evaluation
Relevant time horizon	Duration of the study	Period that the intervention affects outcomes and costs (e.g., duration of screening, lifetime from diagnosis)
Outcomes of analysis	Clinical results (e.g., tumor response, relapse rate, survival)	Years of life saved Quality of life Cost
Aspect	Efficacy (outcomes in carefully controlled settings)	Effectiveness (outcomes in standard practice)

affect the individual's quality of life and use of health care resources in a meaningful way. For cancer screening, the time horizon is usually the age range in which screening is performed. In the case of cancer treatment, it may be the remaining years of life from the time of initial treatment, particularly if lifelong surveillance is indicated. Extrapolation models are necessary to project outcomes data from clinical trials to the time horizon that is needed for most cost-effectiveness studies. For example, the MISCAN model of fecal occult blood screening for colorectal cancer draws on data from the three large randomized trials of fecal occult blood screening (the longest of which lasted 13 years) to project the expected years-of-survival benefit for individuals over a lifetime of screening.[17] It is important to note that extrapolation models depend greatly on the procedure that is used to project observed trends to longer terms. An incorrect extrapolation can severely bias the final determination of cost effectiveness.

Many clinical trials use intermediate endpoints as the focus of the study. For example, a particular study of a chemotherapeutic regimen for lung cancer may list the percentage of patients with a tumor response as an endpoint. The most accepted CEAs, in contrast, use years of life saved or quality-adjusted life years as the major endpoint of the analysis.[13] Linking intermediate markers to years of life saved and quality-adjusted life years is strongest if high-quality studies support the *links in the chain,* namely, the effect of the intervention on the intermediate endpoint and the relationship between the intermediate endpoint and the final outcome. Often, however, one of the links in missing. Certainly, few studies have examined the relationship between intermediate cancer-related endpoints (e.g., relapse, remission) and quality of life. This shortcoming is important, because researchers are increasingly emphasizing quality-adjusted life expectancy as the endpoint for CEAs.[7, 13, 15]

When quality of life is considered in CEAs, the most widely accepted measure of effectiveness is quality-adjusted life years.[13] Quality-adjusted life years combine the life expectancy in years adjusted for an individual's perceived quality of life, measured from 0 (death) to 1 (ideal

health). The quality adjustment is derived from preference weights or health utilities.[7] The advantages of cost utility studies are that they (1) simultaneously capture changes in mortality and morbidity in the measure of effectiveness; (2) are applicable to all disease states and treatments; (3) consider patients' preferences for health outcomes; and (4) conform to normative theory of decision making under uncertainty.[29] Historically, quality-of-life data have not been collected in oncology clinical trials.[3, 18, 23] In most cases, modelers must turn to external sources for quality-of-life information, or must collect primary data from patients who are similar to those who are the focus of the model.

Cost data used in CEA models should be comprehensive and should not be limited to the assessment of the cost of therapy alone. For example, if only the costs of medications are assessed in an evaluation of chemotherapy for testicular cancer, a number of important economic parameters will be disregarded. These variables might include the direct costs associated with the use of medical resources to treat significant adverse reactions to the chemotherapy and hospitalizations and emergency department visits resulting from the disease and its treatment. Direct nonmedical care costs can also be important in cancer care. Examples include transportation costs to and from the physician and the value of the time the family spends caring for patients with cancer. Indirect costs, more correctly known as productivity costs, should also be considered. These costs address the value of work lost (or gained) as a result of the intervention. For screening, work lost during the time spent obtaining screening should be included. Time lost seeking treatment is less clear, because persons with cancer may already be off work because of their illness. There is some controversy among economists about how these costs should be integrated into cost-effectiveness studies, but most agree that they should be addressed and reported if they are likely to be affected by the intervention. Finally, the time horizon for costs in modeling should be explicit. As noted, in all but a very few instances, cancer is a disease that will affect an individual's medical care costs for several years beyond the initial date of diagnosis. Thus, oncology models must account for future expected costs resulting from relapses and long-term surveillance.

Finally, although there is a growing interest in including CEAs directly in randomized clinical trials,[6] most CEAs use data from completed clinical trials because they address a question that cannot be addressed within a trial. When such data are used, the reader must ask whether the sources used to establish the efficacy of the intervention provide credible evidence that the intervention will actually work in clinical practice. There are two dimensions to this credibility: (1) the robustness of the original study design to test the efficacy of the intervention, and (2) the degree to which the original study reflects the practice styles, levels of care, and patient compliance that occur in actual clinical practice. The trial design methodology that determines the robustness of a clinical efficacy study is a core element of evidence-based medicine. The second, more subtle issue for credibility is the difference between

clinical efficacy (the success of the intervention for a narrowly defined patient population treated under the tightly controlled conditions of a clinical trial) and clinical effectiveness (the success of the intervention when used in typical practice settings). Ironically, randomized clinical trials, the studies that have the highest validity in determining efficacy, have the least validity in determining effectiveness. Accounting for the dissimilarities between the clinical trial and typical clinical practice (e.g., patient compliance, comorbidity, less frequent monitoring) is important. There are, however, no widely agreed upon methods for translating efficacy data from clinical trials to units of effectiveness.

SUMMARY

Modeling will continue to be used to address important issues in clinical practice and health policy issues that have not been adequately studied with high-quality clinical trials. The apparent ad hoc nature of models belies the methodologic rigor that is applied to create the best models in cancer prevention and care. Models have progressed from simple decision trees to extremely complex microsimulation analyses, yet all are built using a logical process based on objective evaluation of the path between intervention and outcome. The best modelers take great care to justify both the structure and content of the model and then test their assumptions using a comprehensive process of sensitivity analysis and model validation.

Like clinical trials, models sometimes produce results that are later found to be invalid as other data become available. When weighing the value of models in health care decision making, it is reasonable to consider the alternatives. In the absence of data, clinical policy decisions are often based on the recommendations of expert opinion panels or on poorly defined notions of the standard of care or medical necessity. Because such decision making rarely entails the rigorous process of data collection, synthesis, and testing that is the core of well-conducted modeling, it is usually not possible for external audiences to examine the assumptions and data that were used to derive the decisions.

One of the modeler's most challenging tasks is to make the structure and content of the model transparent to the intended audience. The purpose of this article is to clarify the process of modeling, so that readers of models are more knowledgeable about their uses, strengths, and limitations.

References

1. Banta HD, Luce BR: Health Care Technology and Its Assessment. New York, Oxford University Press, 1993
2. Beck JR, Pauker SG: The Markov process in medical prognosis. Med Decis Making 3:419–458, 1983

3. Brandberg Y: Quality of life in clinical trials: Assessment and utility with special reference to rHuEPO. Med Oncol 15(suppl 1):S8–12, 1998
4. Buxton MJ, Drummond MF, Van Hout BA, et al: Modelling in economic evaluation: An unavoidable fact of life [editorial] [see comments]. Health Econ 6:217–227, 1997
5. Cronin KA, Legler JM, Etzioni RD: Assessing uncertainty in microsimulation modelling with application to cancer screening interventions. Stat Med 17:2509–2523, 1998
6. Drummond MF, Davies L: Economic analysis alongside clinical trials. Revisiting the methodological issues. Int J Technol Assess Health Care 7:561–573, 1991
7. Drummond MF, O'Brien B, Stoddart GL, et al: Methods for the Economic Evaluation of Health Care Programmes, ed. 2. New York, Oxford University Press, 1997
8. Eddy DM, Shwartz M: Mathematical models in screening. *In* Schottenfeld D, Fraumeni JF (eds): Cancer Epidemiology and Prevention. Philadelphia, WB Saunders, 1982
9. Etzioni R, Cha R, Cowen ME: Serial prostate specific antigen screening for prostate cancer: A computer model evaluates competing strategies. J Urol 162:741–748, 1999
10. Etzioni R, Legler JM, Feuer EJ, et al: Cancer surveillance series: Interpreting trends in prostate cancer–Part III: Quantifying the link between population prostate-specific antigen testing and recent declines in prostate cancer mortality. J Natl Cancer Inst 91:1033–1039, 1999
11. Etzioni R, Thompson IM: Prostate cancer and computer models: Background, limitations, and potential. Urologic Oncology 2:57–64, 1996
12. Garber AM, Phelps CE: Economic foundations of cost-effectiveness analysis. J Health Econ 16:1–31, 1997
13. Gold MR, Siegal JE, Russell LB, et al: Cost-Effectiveness in Health and Medicine. New York, Oxford University Press, 1996
14. Muir Gray JA: Evidence Based Healthcare: How to Make Health Policy and Management Decisions. New York, Churchill Livingstone, 1997
15. Haddix AC, Teutsch SM, Shaffer PA, et al: Prevention Effectiveness: A Guide to Decision Analysis and Economic Evaluation. New York, Oxford University Press, 1996
16. Jacobs IJ, Skates SJ, MacDonald N, et al: Screening for ovarian cancer: A pilot randomised controlled trial. Lancet 353:1207–1210, 1999
17. Loeve F, Boer R, van Oortmarssen GJ, et al: The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Comput Biomed Res 32:13–33, 1999
18. Michael M, Tannock IF: Measuring health-related quality of life in clinical trials that evaluate the role of chemotherapy in cancer treatment. CMAJ 158:1727–1734, 1998
19. National Cancer Institute: Cancer Intervention and Surveillance Modeling Network (CISNET). Available at: http://grants.nih.gov/grants/guide/rfa-files/RFA-CA-99-013.html. Accessed 1999
20. Ng AK, Weeks JC, Mauch PM, et al: Laparotomy versus no laparotomy in the management of early-stage, favorable-prognosis Hodgkin's disease: A decision analysis. J Clin Oncol 17:241–252, 1999
21. O'Brien B: Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? Medical Care 34(suppl 12):DS99–108, 1996
22. Ortega A, Dranitsaris G, Sturgeon J, et al: Cost-utility analysis of paclitaxel in combination with cisplatin for patients with advanced ovarian cancer. Gynecol Oncol 66:454–463, 1997
23. Sanders C, Egger M, Donovan J, et al: Reporting on quality of life in randomised controlled trials: Bibliographic study erratum published BMJ 317:1353, 1998; BMJ 317:1191–1194, 1998
24. Thurston SW, Zaslavsky AM: Variance estimation in microsimulation models of the food stamp program. American Statistical Association Proceedings of the Social Statistics Section, 1996, pp 4–9
25. Urban N, Drescher C, Etzioni R, et al: Use of a stochastic simulation model to identify an efficient protocol for ovarian cancer screening. Control Clin Trials 18:251–270, 1997
25a. Urban N, McIntosh M, Clarke L, et al: Socioeconomics of Ovarian Cancer Screening. Oxford, Isis Medical Media, in press
26. van den Akker-van Marle E, de Koning H, Boer R, et al: Reduction in breast cancer mortality due to the introduction of mass screening in the Netherlands: Comparison with the United Kingdom. J Med Screen 6:30–34, 1999

27. van Oortmarssen GJ, Boer R, Habbema JD: Modelling issues in cancer screening. Stat Methods Med Res 4:33–54, 1995
28. van Oortmarssen GJ, Habbema JD, van der Maas PJ, et al: A model for breast cancer screening. Cancer 66:1601–1612, 1990
29. von Neumann J, Morgenstern O: Theory of Games and Economic Behaviour. Princeton, NJ, Princeton University Press, 1994
30. Weinstein MC, Fineberg HV: Clinical Decision Analysis. Philadelphia, WB Saunders, 1980
31. Zaslavsky A, Thurston S: Error analysis of food stamp microsimulation models: Further results. American Statistical Association Proceedings of the Social Statistics Section, 1995, pp 151–156

Address reprint requests to

Scott D. Ramsey, MD, PhD
Fred Hutchinson Cancer Research Center
1100 Fairview Avenue North, MP 900
PO Box 19024
Seattle, WA 98109–1024

e-mail sramsey@u.washington.edu