Development of a Simulation Model of Colorectal Cancer

STEPHEN ROBERTS
North Carolina State University
LIJUN WANG
Vanderbilt University Medical Center
ROBERT KLEIN
Medical Decision Modeling, Inc.
and
REID NESS and ROBERT DITTUS
Vanderbilt University Medical Center

Colorectal cancer (CRC) is deadly if not found early. Any protocols developed for screening and surveillance and any policy decisions regarding the availability of CRC resources should consider the nature of the disease and its impact over time on costs and quality-adjusted life years in a population. Simulation models can provide a flexible representation needed for such analysis; however, the development of a credible simulation model of the natural history of CRC is hindered by limited data and incomplete knowledge. To accommodate the extensive modeling and remodeling required to produce a credible model, we created an object-oriented simulation platform driven by a model-independent database within the .NET environment. The object-oriented structure not only encapsulated the needs of a simulation replication but created an extensible framework for specialization of the CRC components. This robust framework allowed development to focus modeling on the CRC events and their event relationships, conveniently facilitating extensive revision during model construction. As a second-generation CRC modeling activity, this model development benefited from prior experience with data sources and modeling difficulties. A graphical user interface makes the model accessible by displaying existing scenarios, showing input variables and their values, and permitting the creation of new scenarios and changes to its input. Output from the simulation is captured in familiar tabbed worksheets and stored in the database. The eventual CRC model was conceptualized through a series of assumptions that conformed to beliefs and data

This research was supported in part by grant RO1CA92653 from the National Cancer Institute.
Authors' addresses: S. Roberts, Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC 27695-7906; email: roberts@ncsu.edu; R. Klein, Medical Decision Modeling, Inc., 3600 Woodview Trace, Suite 317, Indianapolis, IN 46268-3124; email: rwk@mdm-inc.com; L. Wang, R. Ness, and R. Dittus, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN 37232-8300; email: (lijun.wang, reid.ness, robert.dittus)@vanderbilt.edu.

regarding the natural history of CRC. Throughout the development cycle, extensive verification and validation calibrated the model. The result is a simulation model that characterizes the natural history of CRC with sufficient accuracy to provide an effective means of evaluating numerous issues regarding the burden of this disease on individuals and society. Generalizations from this study are offered regarding the use of discrete-event simulation in disease modeling and medical decision making.

Categories and Subject Descriptors: I.6.3 [**Simulation and Modeling**]: Applications; I.6.5 [**Simulation and Modeling**]: Model Development; I.6.4 [**Simulation and Modeling**]: Model Validation and Analysis

General Terms: Design, Measurement, Verification

Additional Key Words and Phrases: Medical applications, medical decision making, colorectal cancer

ACM Reference Format:

Roberts, S., Wang, L., Klein R., Ness, R., and Dittus, R. 2007. Development of a simulation model of colorectal cancer. ACM Trans. Model. Comput. Simul. 18, 1, Article 4 (December 2007), 30 pages. DOI = 10.1145/1315575.1315579 http://doi.acm.org/10.1145/ 1315575.1315579

1. INTRODUCTION

Colorectal cancer (CRC) is an important health problem, especially in Western societies. In the United States about 150,000 [Cancer Stat Fact Sheets 2006] people are diagnosed with CRC each year. CRC is usually not symptomatic until its later stages. If not found early, it is quite deadly with less than 10% of persons diagnosed with late-stage CRC having a 5-year survival [Cancer Stat Fact Sheets 2006]. On the other hand, 5-year survival exceeds 90% when local disease is diagnosed.

Colorectal cancer is believed to evolve from nonvisible mutated cell collections in the lining of the colon. Ordinarily cells within the colon die every 4 to 5 days, but these mutated cells seem to have no definite lifetime and become invulnerable to the body's normal defense mechanisms. Over time these cells multiply and usually appear as polyps in the colon. If these are precancerous polyps, they are called *adenomas* (not all polyps are precancerous). If not removed, the precancerous cells continue to divide and may spread from their locale of origin throughout the region to eventually be transmitted by the blood and lymph systems to other parts of the body. There the cancer cells will invade vital organs, compromise organ function, and eventually kill the individual.

Beginning with the National Polyp Study [Winawer et al. 1993], there is extensive evidence and widespread acceptance that the removal of adenomatous polyps and early cancers will reduce morbidity and mortality associated with CRC [Pignone et al. 2002]. Screening and surveillance guidelines or protocols have been designed to intervene in the natural course of the disease [Smith et al. 2004]. However, low adherence to any recommendation to screen for CRC [Vernon 1997] complicates the expected results.

Although screening appears to be cost-effective relative to no screening, the single best screening strategy remains unknown [Pignone et al. 2002]. Clinical trials of medical interventions for CRC such as screening provide the most compelling evidence, yet comprehensive trials are difficult for many reasons including (a) the development cycle of CRC may exceed 20 years, (b) there is

constant evolution of medical technology in screening and treatment, (c) any results observed are limited by the age, gender, race, family history, personal health, and other covariates in the population being studied, (d) the costs of screening, surveillance, and treatment vary by locality, (e) the compliance with recommended screening and surveillance must be incorporated, and (f) clinical trials are limited by the ethics of experimenting with accepted clinical practice.

Thus the question of the natural history of CRC and the cost-effectiveness for CRC screening has been the subject of numerous modeling efforts dating from 1980 [Eddy 1980]. Since that time, a number of models of natural history and/or screening cost-effectiveness have been proposed [Pignone et al. 2005]. Almost all the published natural history models use a Markovian framework. Discrete-state Markov models are relatively easy to create and can provide an approximation to general CRC statistics such as incidence and mortality. Yet these models have some weaknesses that are important in modeling the natural history of CRC. First, Markov models create a geometric or exponential state occupancy, neither of which has ever been validated in a CRC context. Second, the memoryless property of the Markov model creates a state space explosion as patient history attributes, important in CRC, are added. Third, these analytical models cannot track concurrent multiple adenoma precursors to CRC within individuals because the state trajectories interact and there is competition for time. Fourth, complex stochastic inputs, such as multivariate and time-dependent representation details needed to model incidence and progression of adenomas, cannot be incorporated. Finally, most of the Markovian models are executed in Monte Carlo fashion, often using a fixed time update (for example, TreeAge (available online at `http://www.treeage.com`)) that potentially creates event collisions and changes the analytic character of the approach, which is the most valuable part of a strictly mathematical approach. Because the detailed modeling of CRC requires model agility and since Markov-based models restrict model flexibility, we prefer a discrete-event simulation (DES) approach that removes these barriers (albeit at the price of more complex input and output analysis). More generally, DES has been recommended for the broad evaluation of disease prevention and treatment [Davies et al. 2003] and for pharmacoeconomic analysis [Caro 2005]. A limited comparison of Markov models versus DES in health economics is offered by Karnon [2003].

The first simulation model of CRC is apparently the MISCAN-Colon (MC) model [Loeve et al. 1999], based on the MISCAN [Habbema et al. 1984] epidemiology framework. This model was promoted by the National Cancer Institute (NCI) and revised in two NCI model review meetings in 1996 and 1997, which "validated" the structure and input of the model [Loeve et al. 2000]. The model was later used to estimate the capacity of endoscopic CRC screening [Brown et al. 2003] and more recently to argue for regression of adenomas [Loeve et al. 2004]. Currently the MC model is being used as the primary analysis tool for the CISNET project (go online to `http://cisnet.cancer.gov/`). The second simulation model, developed independently from the first, which we will call the *Vanderbilt* (V) model, was developed by several coauthors of this article to study CRC screening in a cohort population at risk [Ness et al. 2000]. Both the MC model and the V model were reviewed by Pignone et al. [2002] and both were

employed in the more recent Institute of Medicine workshop [Pignone et al. 2005].

In this article we will describe the development of a new generation CRC simulation model called the *Vanderbilt/NC State* (V/NCS) model. Our objective is to describe the model construction process of the V/NCS model of the natural history of CRC, presenting the basic methodological contribution rather than details on the burden of CRC in society or determining the cost-effectiveness of screening alternatives or issues in cancer policy. Our aim is to make clear how V/NCS was developed, verified, calibrated, and validated. Furthermore, we offer some generalizations for simulating medical decisions based on our experience.

The overall simulation design philosophy and its motivation are described. Acquiring and calibrating valid input for the model was an arduous task consisting of analysis of actual patient records, detailed review of CRC related literature, and employing clinical estimations. In fact, the struggle with data acquisition and its representation in the simulation was a constant endeavor throughout the modeling process, often serving as a means of verifying the behavior of the model and validating its outcomes.

This article is organized into several sections. The next section presents the general environment for modeling CRC and why we chose a database-driven, object-oriented simulation. We also present the model construction strategy and programming tools. In the third section we present the general simulation framework that supports the modeling. This framework provides a base simulation platform on which the CRC model was constructed by inheriting objects and events from it. The fourth section presents more details about the CRC simulation platform. The fifth section describes the verification, calibration, and validation effort, which has been a major activity leading the actual formation of the model. The eventual goal is that the model will provide a behavior that can replicate the natural history of CRC. In the sixth section, we present some results that illustrate the use of the model, recognizing that there are many potential applications. Finally, we discuss our experience in developing a simulation model of a medical process and offer some general perspectives.

2. SIMULATION MODELING OF CRC

Modeling any clinical process can best be described as volatile, requiring frequent and substantial changes to the way the model behaves and often the reformulation of the information database that drives the model. There are several factors that contribute to this volatility. Biomedical "theory" that forms the information base on which a model is constructed is often incomplete. For example, the exact pathways from abnormal cells to colorectal adenoma to CRC, sometimes called the *adenoma-carcinoma sequence*, are not fully understood. Simulation models of such pathways should be regarded as "grand hypotheses" that incorporate a consistent and defensible set of assumptions, but should not substitute for fact. Also, random processes and random variables best describe information that is known. For some of these variable values, such as the overall incidence of CRC, there is some information. For other values, there is little or

no information, such as the incidence of nonvisible adenomas. While we know, for example, that 15% of CRCs appear rather spontaneously, we do not know what percentage of nonvisible adenomas exhibits this behavior. Finally, the relationship between what knowledge there is about pathways and the actual supporting data is tenuous and dynamic. Sometimes the lack of complete information prevents the direct modeling of some behavior. For example, data on the occurrence of symptomatic CRC may not be directly related to the cancer stage change in CRC, so stage changes and the appearance of symptoms may need to be modeled indirectly.

The volatility in modeling is sometimes viewed as manifest "uncertainty" in model parameters and model structure [Manning et al. 1996]. But in simulation, since modeling consists of an evolution of models, the uncertainty also pertains to the model evolution trajectory. A general simulation modeling effort must accommodate any number of evolutionary patterns and not be prejudiced by the orientation perspective of a simulation modeling language such as the queuing network framework of Arena [Kelton et al. 2004], the location character of ProModel [Harrell et al. 2004], the movement in three-dimensional space of AutoMod [Banks 2004], or other such simulation modeling language concepts, even though all of these are both valuable and have programmable extensions. Commercial simulation programming languages such as SLX (go online to `http://www.wolverinesoftware.com`) and CSIM (go online to `http://www.mesquite.com`) would have been applicable, but were not considered further to avoid the commercial licensing and distribution complications for any model that might be used or extended by others.

We favored a purely object-oriented approach [Joines and Roberts 1998] with which we were more familiar (both conceptually and at a programming level). An object-oriented simulation (OOS) was attractive because of the power to abstract not just the functionality of CRC behavior but the data structures within a simulation model describing CRC objects. There are some noncommercial epidemiological simulation resources such as SIMEX (go online to `http://www.la.utexas.edu/lab/software/lib/simex/NMSR.html`) and MISCAN ([Habbema et al. 1984]) but their detailed implementations (such as simulation fundamentals) are unknown and the software was unavailable. Furthermore, many similar tools are oriented to simulating communicable and infectious disease, neither of which is applicable to CRC. Thus, by constructing the simulation ourselves, we can insure that the software is fully available and the details accessible.

2.1 Implementing the Simulation

The CRC model has been constructed using a general object-oriented, discrete-event simulation platform implemented in the Microsoft .NET environment to be efficient and adaptable in the (present and future) Windows operating system. We therefore decided to work almost exclusively in the Microsoft realm including using Microsoft Excel for output and Microsoft Access for storing/retrieval of data. The Microsoft .NET Visual Studio interactive development environment provided a flexible development platform and Microsoft Visual

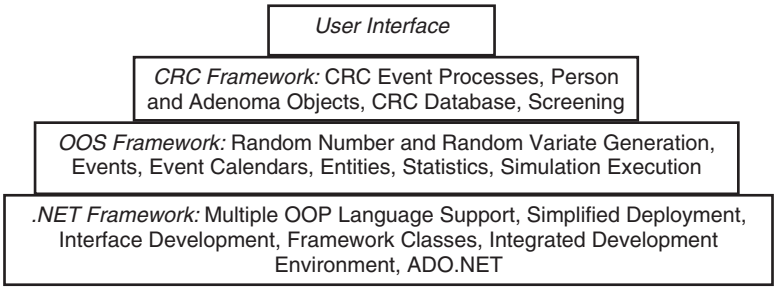

Fig. 1. The object hierarchy.

SourceSafe was used for version control—important in a multi-institutional project. VB.NET was chosen as the primary programming language owing to its historic interface development environment, although the fact that all .NET languages employ the Common Language Runtime (CLR) made this choice one of convenience rather than function. In fact, our model uses a managed C++.NET random number generator.

The implementation employs a four-tier object hierarchy, illustrated in Figure 1.

The base of the hierarchy is the .NET framework, which provides the multiple language support, interface development, framework classes, Integrated Development Environment (IDE), and ADO.NET that is used to interface to the simulation database. Using that base, we constructed an OOS platform framework which consists of the random number and random variate generation, the event class, entity class, event calendars, statistics classes, and simulation execution. The CRC objects framework provides the model components for the CRC project, namely, the person and adenoma objects, the CRC event processes, the CRC screening methods, and the CRC database. Last, the user interface promotes use of the model for analysis, but without additional programming.

2.2 Primary Simulation Objects

Overall, our design employs an OOS, driven by a model-independent database. The primary object in the CRC simulation is the person. The simulation of the events within the lifetime of a single person determines the length of each simulation replication. The replication will be terminated when the person dies or when statistics collection ends. CRC is not communicable and each person is medically treated without regard to any other person. Multiple replications are therefore appropriate for output analysis and, the statistical problems associated with dependent observations within a simulation are not present [Law and Kelton 2000]. Together the replications are used to create a population at risk. Individuals in the population will have their own properties such as birth date, gender, race, and family history of CRC. The evolution of adenomas in any given individual is dependent only on that individual's characteristics.

The population at risk may consist of a cohort population whose properties are prespecified (such as 50-year-old white females with no family history of CRC). Or the population at risk may correspond to a population profile (such

as the population in the U.S. in the year 2000). Statistics can be collected according to the cohort whose endpoint is death or a certain age or collected for the entire population each year for some given number of years providing a dynamic population analysis.

Within each person object, secondary objects are adenomas, which are collections belonging to each person; each adenoma (several may be present) has its own development cycle and impact on survival. However, the trajectories of adenomas may interact. For example if an adenoma is discovered, any surgery or colonoscopic polypectomy procedure will remove all detected adenomas. Moreover, adenoma incidences are correlated through common risk factors.

A simulated person has a potential lifetime devoid of CRC. Nonetheless this lifetime can be altered by the incidence of one or more adenomas, whose transition to CRC may cause a CRC-related death to the person.

The specification of the simulation model for the natural history of CRC includes the various events and may include external actions such as screening or treatment experienced in the "life" of the person and their adenomas. Clearly, interventions such as screening and removal of adenomas and early cancers can cause important downstream changes in the future CRC events. Finally, an independent database removes the data specification from the model and updates to the data are restricted to database changes, rather than changes in the model (and vice versa).

2.3 Simulation Database

Typically, the input for a simulation model and the model, which manipulates the inputs, are interwoven. Changes in inputs can be a source of inadvertent changes to the actual model structure. There was a frequent need to change, modify, and refine the input for the CRC simulation. Many of the inputs to the CRC model required the "fitting" of unknown variables to accomplish target output and the simulation had to be executed many times to find input values that obtained proper output behavior. Furthermore, since a simulation of the natural history of CRC is itself a complex and controversial topic, there are frequent changes to the scale of the model (such as the addition of gender and race) as well as the details of the input (such as the incorporation of histology).

Inputs for the CRC simulation model are generally random variables that provide the stochastic/statistical processes within the simulation. Such inputs will typically be objects known by their name and parameters (which may describe a random variable or random process). But since the simulation is an OOS, the simulation can accept any "random variable object," interpret it when the simulation initializes, and use it throughout the simulation. Therefore a change from a lognormal input to a beta input is isolated to changing the database. Features and further advantages of the database input are discussed in Appendix A: CRC Simulation Database.

2.4 User Interface

The population being examined (mix of ages, birth years, races, genders, etc), the screening strategy (what, if any, screening and the screening

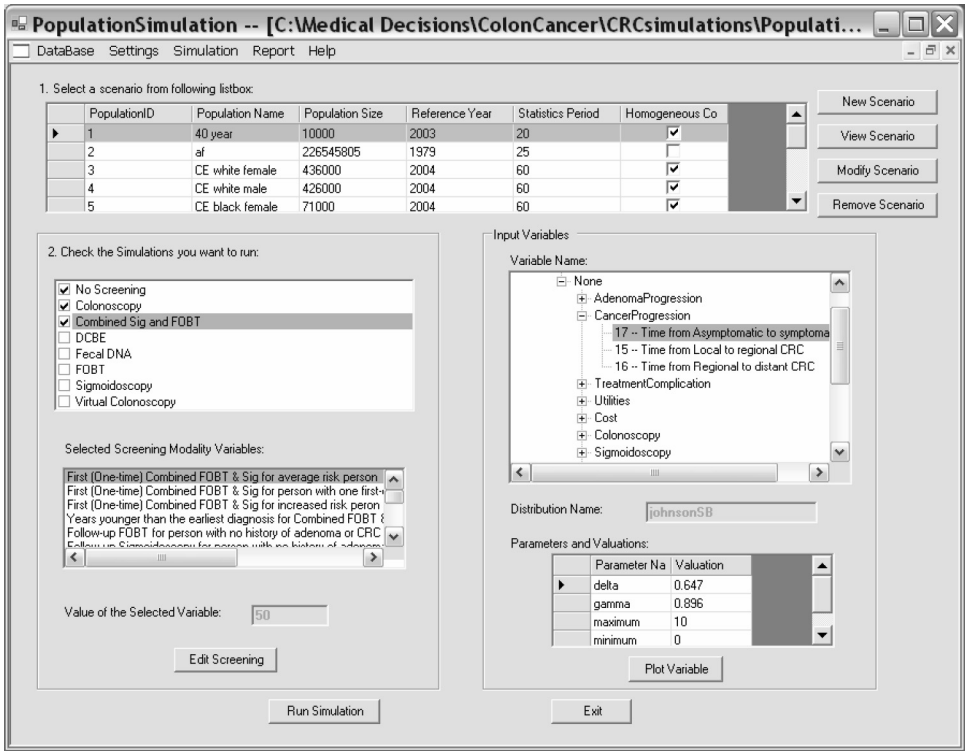

Fig. 2. Main scenario display.

characteristics), and the types of statistics to be collected and displayed are specified through the user interface. The main screen for the simulation is given in Figure 2.

Existing scenarios are displayed and may be viewed, modified, or removed. New scenarios may be created. Any number of screening options may be selected for the simulation and screening modality variables may be viewed and edited. CRC input variables are listed and their values or distribution parameters displayed. If a distribution is being used for input, then the visualization component ("Plot Variable" button) can be used to display and possibly edit the distribution shape (density function, distribution function, hazard function) and compute statistical properties such as mean, mode, etc. Input data needed for the structure of the simulation that has been validated by the research team may be viewed but cannot be changed in the user interface (such data is considered "fixed" via the validation and so can only be changed through the database manager).

An Excel workbook subdivided into detailed worksheets is used to display output, thus making Excel a needed component for the simulation. Tailoring the output is limited to a few options, so the standard worksheets are the primary results of the simulated scenarios. Standard worksheets include the following: (1) information on the population being simulated including the initial

population according to age (5-year intervals), gender, race, and family history. Included is the population size and the reference (or initial) year, the statistics collection period, the new births/in-migration over the years, and the screening strategy, if any. (2) Adenoma incidence by year and location in the colon. (3) Adenomas found is by year and by location in the colon, and according to whether the adenoma was advanced. (4) Cancers found reported by year according to location and cancer stage (local, regional, or distant). (5) Compliance by year and individual compliance characteristics. (6) Five-year cancer survival by cancer stage and also the number of CRC deaths, non-CRC deaths, number of survivors, and the survival rate. (7) The natural deaths, the cancer deaths, and the survivors reported by year. (8) Costs and effectiveness are reported for the undiscounted rate and the discounted rates at 3% and 5%, and 95% confidence intervals for the undiscounted values. The values reported include the cancer costs, the screening costs, the surveillance costs, the total costs, the quality-adjusted life years, and the total life years. For more detailed analyses, the cancer discovery times, cancer care durations, and cancer terminal charge times are available in another worksheet by year for the various cancer stages. Means and confidence interval bounds are included. (9) A worksheet on colonoscopy tests is available that gives the time of screening, diagnostic and background colonoscopy tests by year and includes confidence intervals and numbers of observations. (10) Finally, a summary is available that provides population statistics including the number of different types of screening tests performed, the number of polypectomies, biopsies, bleeds, perforations, symptomatic cancers, people with CRC, surgical and CRC deaths, costs at different discount rates, quality-adjusted life years at different discount rates, and life years at different discount rates.

To facilitate input modeling for CRC, especially since the data environment was limited and because much of the input depended on the opinions of "experts" (clinical personnel), a special piece of software called *VIM* was developed that permits the subjective specification of an input distribution. Appendix B: Input Modeling discusses the VIM software and its use. This software was also the source for the "Plot Variable" option in the user interface (refer again to Figure 2).

In contrast to other CRC models and medical decision models in general, we offer a Web site to download the simulation software, its database, and VIM (go online to `http://www.ie.ncsu.edu/robertsresearch/PopulationSimulation/`). Further, this Web site contains a User's Manual for those using the software in addition to the "Help" function within the software.

3. THE SIMULATION FRAMEWORK

The simulation framework provides the mechanics that formulate and execute the simulation employing the object-oriented implementation strategy. This framework was designed to contain the simulation components found in many discrete simulations, but focuses on the simulation of independent entities constrained only by time and competing events (not queuing for limited resources). This framework uses some classes from the .NET framework.

3.1 Random Number and Random Variate Generation

The random number generator used in the simulation framework is the combined multiple recursive generator proposed by L'Ecuyer [1999] and given as an object-oriented implementation (C++) in L'Ecuyer et al. [2002]. Conversion to the .NET managed environment made the random number generator native to the other simulation code so that derived classes could use all the properties from this base class for random variate generation.

The simulation framework employs inverse transform generators for all the random variables to attach a unique source of randomness to each variable in the model. Consequently, identical populations of individuals who are at risk for CRC have "common random numbers" in different simulation experiments to reduce the sampling error throughout the simulation and statistically enhance the potential comparison of interesting scenarios [Law and Kelton 2000]. Additional information on the use and generation of random numbers can be found in Appendix C: Random Numbers and Random Variates.

3.2 Simulation Events and Entities

The simulation framework defines two base classes: SimulationEvent and SimulationEntity. SimulationEvent is an abstract class (must be inherited) that stores an event time for each object, a read-only method to get the event time for the object, and a virtual method called EventProcess. Overriding the Event Process in derived events causes the simulation to execute the appropriate event process. The actual next-event processing is handled within the SimulationEntity object since it is the lifetime of the person that controls the length of the replication. The EventsCollection is a third class in the framework that defines the overall event calendar which collects SimulationEvents through scheduling new events, removing events, counting events, and clearing the calendar of events. Actually, the EventsCollection calendar is potentially a collection of event calendars, so the management of events must be managed within the collection of event calendars. In the CRC simulation model, the overall event calendar manages death events, screening events, and other events. Adenoma events are managed in separate calendars, which dramatically improved the simulation execution time because the computer time spent searching the event calendar for entries when adenomas are removed during or after screening is greatly reduced.

The SimulationEntity class defines the base entity so that instantiated objects will have their own unique ID numbers. Most importantly, this class contains the Activate function which executes the discrete-event simulation by getting the next event, updating the global clock, and then calling the Event-Process of the event. Thus all the mechanics of the simulation are separated from the CRC specializations, keeping the development of CRC independent from the underlying simulation layer.

3.3 Scenario Control

Scenario control is accomplished in the main routine called *Simulate*. It starts the dialog with the user to set the scenario and begins the simulation. Here a

SimulationEntity (actually a derived entity) is created and its Activate function is invoked, which initiates the event processing cascade. In the CRC simulation, it is a person at risk in the population who has his events processed. Events for that person are processed until death or the end of the statistics collection period, which causes the event processor to exit and clear any remaining events. Statistics for the entity are collected both during and at the end of each replication. Finally, the next substream is indexed so the next replication samples all needed random numbers at a known point in each random number stream. Furthermore this indexing insures that identical persons in a defined population are repeated regardless of the screening or treatment method.

3.4 Statistics

The statistics frame includes an observation-based statistics class that maintains a running mean and variance to obtain numerically stable statistics (see Chan and Lewis [1979] and West [1979]) as well as the count, minimum, and maximum properties. A special class is used as a member of a class inherited from the observation based statistics to create cells of "breakdown" statistics for subgroup analysis. Breakdown statistics are collected on various important subsets, such as lifetimes broken down by age, race, gender, etc.

The statistics frame also consists of a count breakdown class that is inherited from the breakdown statistics, so counts that are recorded as observation statistics are collected. Such counts provide "incident" information in the CRC model.

Finally the statistics frame includes a time-persistent statistics class. Time-persistent statistics are needed in CRC for discounting quality-adjusted life years when the discounting is performed continuously to reflect the continuous nature of the statistics such as life years.

4. THE CRC SIMULATION

The CRC specialization of the simulation platform contains the largest segment of program code in the simulation. This framework presumes the simulation of (independent) individuals and consists of constant declarations that are used throughout the model. These also provide descriptors for the health states, cancer stages, adenoma locations, compliance, histology, adenoma type, symptoms, risk effect members, screening types, family history identifiers, and database input variable identifiers. Most of these were rather stable throughout the development with the most frequent changes being made to the input variable list (see Appendix A: CRC Simulation Database).

4.1 Variable Input from the Database

Several modules in the CRC framework are devoted to obtaining the input variables for the simulation. At the start of the simulation, a user-specified CRC database is opened. The connection to the database and queries to it are done through ADO.NET. A routine is devoted to interpreting the input variables from the database at the time the simulation is initiated. Another routine associates a distribution (of which a constant is the degenerate case) with input

parameters and seeds, thus setting up the sampling from the input variables. Finally, the internal simulation variables are assigned the distributions from which they are sampled.

A routine in these modules recalls from the database the natural lifetime distribution (an empirical continuous distribution) without CRC for this person based on the person's gender, race, age, and reference year for age. To insure lifetimes from the year 1900, the lifetime distribution functions for all possible age, gender, race, and family history groups were constructed and entered into the database. Existing life tables for each race and gender with birth years 1968 to present [Wilmoth 2003] and tables for gender with birth years from 1900 to present within the Berkeley Mortality Database [Wilmoth 2003] had to be modified to provide gender-based data before 1968 and adjust for race for birth years from 1990. Details of this adjustment are found in Cubbage [2004]. No other CRC model accounts for individuals in the population according to age, race, gender, and family history categories. For example, the MC model uses only an externally generated age-distributed population.

Finally, it should be observed that many of the random variables within the V/NCS natural history model are often represented by Johnson S_B distributions and the time-dependent variables are usually modeled by nonhomogeneous Poisson processes (NHPP) (see Law and Kelton [2000] for reference to these and other random distributions and processes). The choice of the Johnson S_B distribution was made because of the lack of data and the need to represent a wide variety of bounded, nonsymmetrical shapes, both characteristics of biomedical variables. A Johnson S_B distribution has excellent flexibility because of its four parameters. The presence of bounds in biomedical variables corresponds easily with the "bounded" Johnson. Parameters for the Johnson S_B were obtained from either the use of minimum, mode, and maximum estimates (with the standard deviation being one-sixth of the range) or direct estimation using VIM. The NHPP model was chosen because of the belief that the Poisson rate applies to adenoma incidence and that the rate function varies with time, although the NHPP model may understate variability. Nevertheless the NHPP proved adequate and, in fact, all incidence data is created in terms of rates per 100,000 people subdivided by year or age groupings which are directly modeled with the NHPP. The NHPPs in this work use piecewise linear rate functions, which also appropriately model our intuitive beliefs about the rate changes.

Surprisingly many medical simulation models, such as MC, employ only constants and exponential distributions for random variables, probably because of the ease of estimation. The uses of exponentials are especially problematic given their impractical distribution characteristics. As a consequence, only mean values are valid and variability analysis of output is very limited.

4.2 Persons and Adenomas

The central object in the CRC simulation is the person object, which is instantiated from the Person class that is derived from the entity class. For each replication of the simulation, a person object is created and that object follows in its own pathway among events until that person dies and ends the replication.

Person objects have a number of properties, which generally fall into six groups as described in Appendix D: Person and Adenoma Object Details.

A related class of objects associated with each person in this model is the collection of colorectal adenomas for that person. Whenever an adenoma is created (from the Adenoma class), it is assigned characteristics that determine if and when it will become CRC. Once formed, a CRC can progress through various stages and lead to patient death. One or more adenoma or CRCs may also be found through screening or diagnostic testing and then be removed.

In the instantiation of the person object, all the characteristics of the person are established (usually from random variables) or initialized to beginning values. Several of the CRC characteristics such as incidence and progression are affected by personal risk. These effects modify the base rate functions for the NHPP processes that describe incidence of adenomas and their progression for each person, since risk is a personal characteristic.

The scenario object is instantiated from the Scenario class only once at the beginning of the simulation run to define the characteristics of the population to be studied and the intervention that is to be performed. Parameters for each random variable are assigned based on scenario definitions. Parameters for each scenario are stored and retrieved from the database, which also stores the calibrated variables for the natural history simulation.

4.3 CRC Events

The events in the CRC model can be viewed as a generic event graph (see Schruben [1983]). The event graph displays all events in the simulation and, when relevant, what conditions are required for an event to be scheduled. If no conditions are stated, the event is always scheduled. Events in the CRC simulation are documented by their description, predecessor events (events that must take place prior to this event occurring), condition for next events scheduled, and statistics collected. The event graph for the V/NCS model is given in Figure 3. The rounded, rectangular, bolded events are the death events.

Although there were many issues related to implementing the mechanics of the simulation and in the definition of the CRC objects, most of the effort in evolving the model was the identification of events and the event behavior, including scheduling new events. However, this localization reduced the complexity of the simulation model as the trajectory of model evolution was primarily restricted to data and events.

4.4 The “Grand Hypothesis”

During the development of the CRC simulation a number of questions arose regarding the natural history of CRC and the potential of various screening technologies. Many of the questions related to natural history had no definitive answer. Since the full course of CRC is not completely understood, the model was formulated on the basis of educated assumptions regarding the process of CRC accepted by our medical consultants. We refer to the assumptions together as a *Grand Hypothesis* to remind all that we work with a model of CRC. The

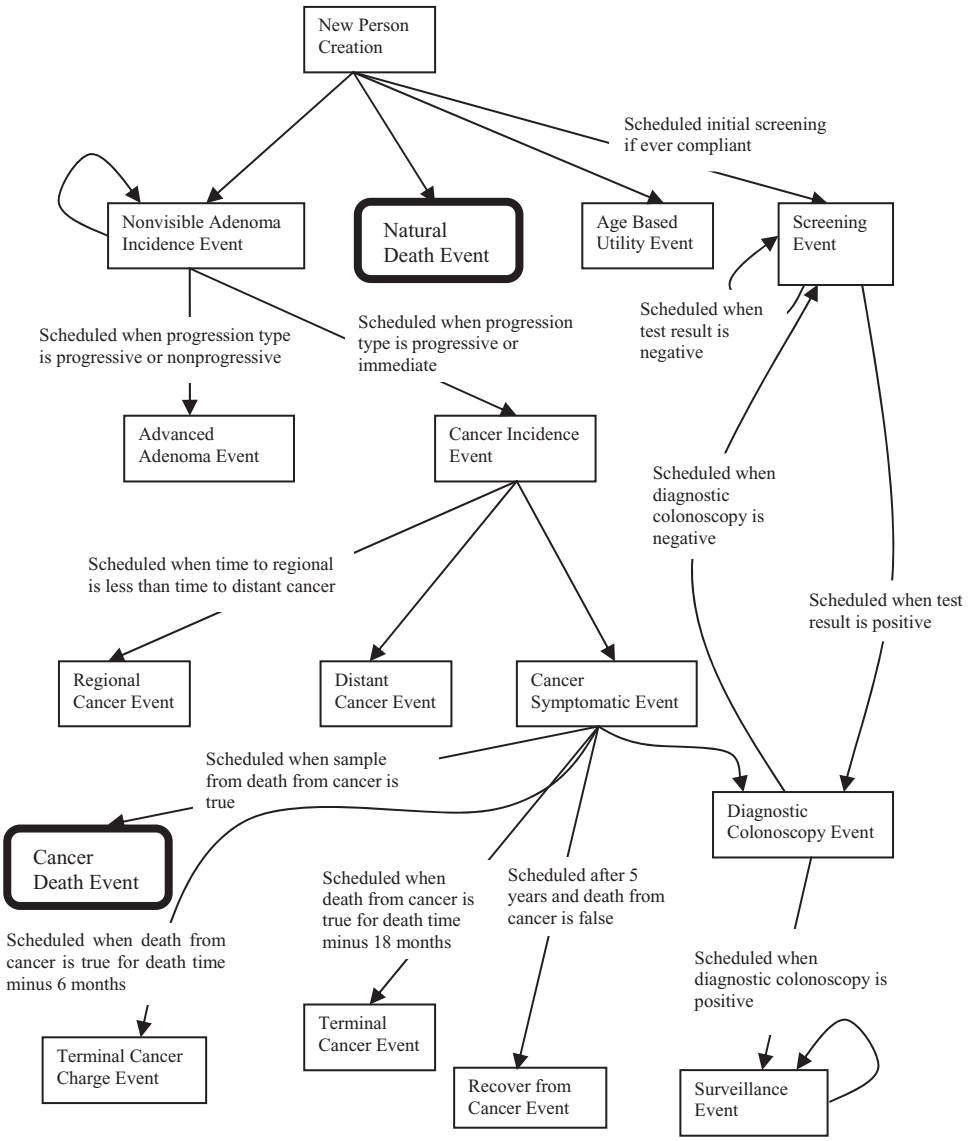

Fig. 3. Event graph for V/NCS model.

set of assumptions regarding the natural history are found in Appendix E of the online supplement to this article, available in the ACM Digital Library.

To elicit structural information about CRC progression given limited data, we employed a modified Delphi approach and a panel of 15 clinical and biomedical experts to provide estimates through several rounds of Web-based surveys [Liebsch 2003]. The study sought to elicit model inputs for the following four topics: (1) the proportion of CRCs that cannot be prevented through conventional screening (i.e., CRC that does not pass through a visible polyp intermediary);

(2) whether the genetic predisposition posed by an affected first-degree family member increases the adenoma incidence rate, the progression rate from adenoma to CRC, or both and to estimate the relative proportions of these two factors in affecting a person's risk of developing CRC; (3) the distribution for conversion time to CRC from an adenoma; and (4) the distribution for conversion time of asymptomatic to symptomatic CRC. The final simulation model inputs were developed using these estimates and the VIM distribution-fitting software. Because this information represents the best estimates available to date, and considering there is limited data to support formal analysis, the inputs developed as a result of this process are extremely valuable and have provided the critical characterization of the issues targeted.

5. VERIFICATION, CALIBRATION, AND VALIDATION

Microsoft Visual Studio provided an integrated development environment for constructing, executing, and testing the code and for creating robust and efficient programs for Windows. Verification was integrated into the development process through (1) stepping through the simulation source code line-by-line and routine-by-routine to verify the general program flow, (2) creating and examining an output trace file, and (3) watching the key variables within the execution of the code. The verification process focused on the execution of events to be sure that the simulation state variables were being updated and that variate generation and output statistics were computed and collected properly. To illustrate, when the adenoma incidence event is triggered, the type of progression of the adenoma must be determined. Here the simulation must sample from the proper distribution and correctly assign the adenoma to one of the three types. If the adenoma could progress to an advanced adenoma, then the simulation must correctly check the "risk affects" variable to see if risk affects progression in this individual. If it does, then the calculations are done by hand to verify that the time to next stage is correctly sampling once the adjustment due to family history is added. Last, the scheduling of the next adenoma is viewed to see if it correctly schedules the next adenoma, and to further verify that the simulation was correctly sampled from the NHPP for incidence. Similar step-by-step processing is done for the other larger events to ensure their proper functioning.

5.1 Calibration of Model Input Based on Targets

More than the assumptions, the model connects variables influencing this history through causal numerical relationships, such as events. Input values for many variables in the model are incomplete, especially those with random components having age-dependent parameters. Direct estimates of many of these input values are not present in the medical literature. Also some of these estimates, such as how long it takes cancerous adenomas to progress to more significant states, are unobservable (it would be unethical to allow such adenomas to progress without immediate action simply to observe this value). Therefore some input values within the model must be inferred from other known values.

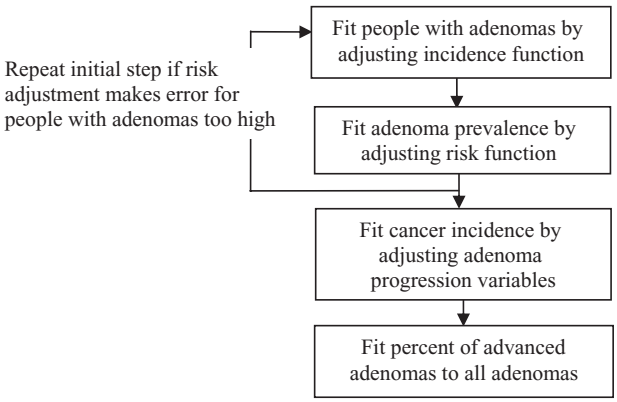

Fig. 4. Flowchart fo fitting procedure.

Although the problem of missing values in input data sets is a well-researched problem in statistics and in sampling, the complex nature of CRC input models means simple substitution of missing values for input is not possible, regardless of the missing value criteria. So to match known CRC results, a form of input fitting was developed to create model behavior that mimics what is known about the real system. This input fitting process resulted in adjustment of input parameters and their dependencies. The standard percent error computation based on absolute deviations was employed:

$$Error\% = \frac{|TargetValue - ModelOutput|}{TargetValue} * 100.$$

In general, the average error and maximum error were used to determine the goodness of a fit or calibration of data to targets (see Figure 4). The average insured that overall values were reasonable, whereas the maximum error was used to insure that the fit was adequate over the entire range of values. In addition, a qualitative visual "smoothness of fit" was used to be sure that values did not oscillate around targets.

As CRC incidence is the primary output of the natural history model, it was the primary output target that was used to "calibrate" the model. SEER (Surveillance Epidemiology and End Results statistics from the National Cancer Institute) data before screening became widespread, during 1973–1982, was employed to determine target age-dependent incidence values [U.S.Department of Health and Human Services 2004]. Adenoma prevalence in terms of both the percentage of people with adenomas and adenomas per 100 people was used as the secondary target. This data was obtained from autopsies and similar studies of age-related adenoma prevalence with and without family history (see, for instance, Ahsan et al. [1998], Boutron et al. [1995], Fuchs et al. [1994], Pariente et al. [1998], St. John et al [1993], and Winawer et al. [1996]). Most of the variables within the simulation were predetermined based upon clinical estimates or were established by analysis of comprehensive medical studies. The input variables where no distribution was known in advance included the population

adenoma incidence rate, the distribution of familial risk for CRC, and the fraction of prevalent adenomas of various types (e.g., nonprogressive). Fitting was performed for each combination of race and gender employing a 19% rate of persons with a significant family history of colorectal neoplasia using a cohort population of one million people starting at age 40 in 2003 and simulating this cohort from birth to death.

While several approaches to calibrating/fitting failed to produce satisfactory results [Cubbage 2004], the ultimate procedure used the steps described in Appendix F of the online supplement, which also contains examples of intermediate results of the procedure. The overall fitting procedure is shown in Figure 4.

After an initial estimate of the incidence function to produce persons with adenomas, the second step was to adjust the function describing the distribution of CRC risk in the population to meet the adenomas per 100 persons target. Some cycling within these first two steps was needed to produce both an incidence function and a risk function.

The third, and most complex, step was to fit cancer incidence by adjusting adenoma progression variables until (1) cancer incidence from progressing adenomas constituted 85% of the total cancer incidence, and (2) cancer incidence from immediate adenomas constituted the other 15% of the total cancers. Cancer incidence targets were obtained from SEER by eliminating the people that died in that year and multiplying that year's cancer rate to get the cancer incidence in a given year, matching on number of cancers, not just the rates.

Adenoma progression variables for each age denote the type and time of progression for each gender, race, and family history. The calibration was done by changing the rates of progressing and immediate adenoma types to create an overall fit. The process started with younger ages and moves to older ages. Ages 50–80 were considered the most important since these ages contain most of the CRCs and their precursor lesions. Thus younger-age adjustments, which impact older-age statistics, required delicate changes. In addition, a proper ratio between cancers arising from advanced adenomas and cancers arising from nonadvanced adenomas was needed; the target ratio was 2.4. Cancers arising from advanced progressive adenomas should account for 60% of total cancers and cancers arising from nonadvanced progressive adenomas should account for 25% of total cancers.

To have an appropriate overall distribution for "time to transition to CRC for progressive adenomas," the probability density function curve should reflect the combination of the "With family history" progression curve and the "Without family history" progression curve. Figure 5 shows the expected characteristics of the curve based on the survey of the 15 experts cited earlier [Liebsch 2003].

Finally, personal risk may affect adenoma progression (see Appendix E in the online supplement). If risk affects only progression, then the sampled progression value for each adenoma is obtained as the complement of a sample from relative risk (meaning a high risk produces faster progression). Successive progression times for each adenoma for an individual are serially correlated. When both incidence and progression are influenced by risk, then the sampling

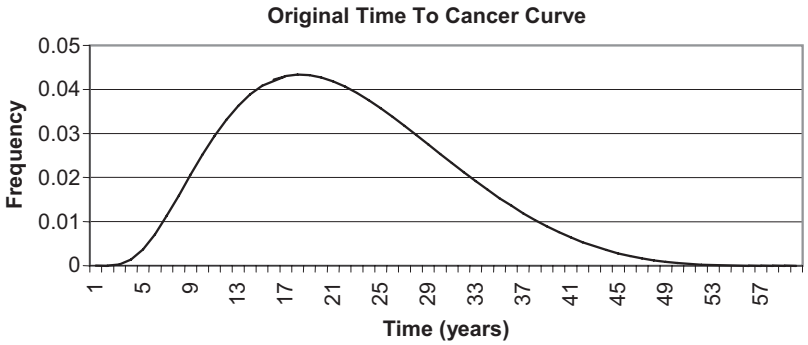

Fig. 5. Transition time to cancer.

for the progression time for each adenoma uses serially correlated samples for incidence and progression.

All three steps in the calibration required iteration, since the variables are so interdependent. Finally, the fourth step was the fit of advanced adenomas to all adenomas. Altering the time from progressive nonadvanced adenoma to advanced adenomas was done in the third step, so here the time from nonprogressive, nonadvanced adenomas was altered. This calibration only affected the percentage of advanced adenomas since nonprogressive adenomas never become cancer.

More details on the actual fit results can be found in Appendix F of the online supplement, including those related to cancer stages. These numerical results verified the model by examining the logical consistency of the statistics being reported. The calibration also provided strong evidence of validity by multiple comparisons of information produced by the model to externally developed targets.

5.2 Screening Validation

To validate the ability of the V/NCS model to simulate cancer prevention strategies over a specific period of time, we simulated the Minnesota Colon Cancer Control Study [Mandel et al. 1993]. The main objective of this study was to compare annual and biennial FOBT screening to no screening. This trial was chosen because it is the only fully reported, large clinical trial performed within the United States. An enhanced V/NCS model was designed to have the capability to replicate the Minnesota study (see Tafazzoli [2004] and Tafazzoli et al. [2005]).

The most difficult part of this process was to generate a population with the same characteristics as observed in the Minnesota trial. Most of the simulation outcomes, including the CRC mortality, were somewhat higher than the actual trial. This result may reflect a higher risk for CRC in the modeled population compared with the actual trial population (which may have been “healthier”). However, the percentage decrease in CRC mortality was within the expected decrease in death of the simulation model. This result adds substantial validation to the V/NCS model.

Table I. CRC Effect on Population Subgroups

Gender	Race	Family History	Overall Life Years Lost
Female	Black	No family history	0.172 (0.002)
Female	Black	Family history	0.356 (0.003)
Female	White	No family history	0.225 (0.002)
Female	White	Family history	0.460 (0.003)
Male	Black	No family history	0.167 (0.002)
Male	Black	Family history	0.338 (0.003)
Male	White	No family history	0.191 (0.002)
Male	White	Family history	0.389 (0.003)

Table II. Life Years, QALYS, and Cost Impact on Individuals Who Die from CRC

Gender	Race	Family History	Life Years Lost	QALYs Lost	Costs of CRC
Female	Black	No family history	10.83 (0.08)	10.24 (0.07)	$123,714 (3736)
Female	Black	Family history	11.23 (0.06)	10.62 (0.05)	$114,381 (2408)
Female	White	No family history	11.68 (0.07)	10.99 (0.06)	$124,875 (3320)
Female	White	Family history	12.15 (0.05)	11.45 (0.05)	$118,188 (2283)
Male	Black	No family history	10.19 (0.07)	9.74 (0.07)	$110,460 (3188)
Male	Black	Family history	10.62 (0.05)	10.18 (0.05)	$113,317 (2326)
Male	White	No family history	9.90 (0.06)	9.52 (0.06)	$126,345 (3290)
Male	White	Family history	10.29 (0.05)	9.91 (0.04)	$123,590 (2283)

6. RESULTS

The simulation can be used to examine a number of important aspects of CRC. From a societal perspective, the burden of the disease can be measured in life years. We simulated one million people from birth to death in each of eight cohorts representing different genders, races, and family histories using the age distribution (in 5-year increments) in the year 2000 for the U.S. population. The results are given in Table I (numbers in parentheses are standard errors). The overall impact of this disease is nearly 0.25 years/person.

Clearly, the population with the most to gain in terms of life years is the population that has a family history of the disease. Due to greater life expectancy, females lose slightly more life years from CRC than males.

People who acquire CRC are those whose lives are directly affected. The "burden" of CRC may be summarized in other terms. Along with life years lost, the quality-adjusted life years (QALYs) are an important measure of the impact CRC has on the population [Ness et al. 1999]. The quality adjustment is determined by multiplying the "utility" for each health state by the time a person spends in each respective state. The natural history model includes not only changes in quality due to CRC but also age-related changes. In addition to years of life, and the quality of those years, the other key metric for gauging the impact of a disease from society's standpoint is the average attributable cost of CRC per individual. Costs can be considered by adding the historic diagnostic and treatment costs of CRC to the natural history model. These would occur with symptomatic cancer. The average life years lost, the QALYs lost, and the costs of CRC for individuals who die from CRC can be seen in Table II (numbers in parentheses are again standard errors).

Averaged over the U.S. population profile, the CRC victim loses 10.91 years. If those who incur CRC but do not die from it are included, the average life

Table III. Effects of Colonoscopic Screening

Gender	Race	Family History	Additional Cost	Life Years Gain	Cost per Life Year Gain
Female	Black	No family history	$494	0.059	$8342
Female	Black	Family history	Cost-saving	0.129	Cost-saving
Female	White	No family history	$322	0.080	$4008
Female	White	Family history	Cost-saving	0.175	Cost-saving
Male	Black	No family history	$459	0.063	$7329
Male	Black	Family history	Cost-saving	0.128	Cost-saving
Male	White	No family history	$192	0.075	$2571
Male	White	Family history	Cost-saving	0.159	Cost-saving

years lost is 4.40 years. In most cases, the presence of family history of CRC has the largest influence. Race appears less important than gender in terms of impact on individuals. White males with no family history apparently have the smallest loss, while white females with a family history tend to lose the most years

The high burden of CRC suggests that some form of screening or effective treatment option will be worthwhile based upon the costs and life years lost associated with the disease. Screening tests are administered to asymptomatic individuals who have some risk of developing CRC. Clinical guidelines for CRC screening are published by the American Gastroenterological Association [Winawer et al. 2003]. A frequent screening procedure based on age (typically 50) is the colonoscopy, which is administered every 10 years, assuming no abnormal results and no other risk factors. Since the natural history simulation was extended to accommodate screening, the same population groups were simulated using this screening protocol. Table III shows the additional cost, the gain in life years, and the cost per life year gain in the simulation population groups.

In the case of those with a "family history," the colonoscopy screening appears to be actually cost-saving in that it lowers CRC-related costs more than the added costs of screening. There are life year gains. To combine cost and life years into a single cost-effectiveness measure, the cost per life year gain ratio is computed. These results show that the cost-effectiveness for people without family history is much smaller than $50,000 year life year gained, the typically accepted threshold of medical cost-effectiveness in the United States. More rigorous analysis of screening options for CRC using the V/NCS model can be found in Tafazzoli [2004] and Tafazzoli et al. [2005].

An aspect of screening that is often ignored is compliance with the screening protocol. While Table III is produced with "typical" compliance, the model can produce comparable statistics assuming "perfect" compliance. Table IV displays the incremental additional cost, life years gain, and corresponding cost per life year gain beyond the expected typical compliance.

Moving to perfect compliance increases life years, but also increases costs due to increased screenings (the cost to obtain perfect compliance is not included). Interestingly, the incremental cost-effectiveness of the improved compliance is quite similar to screening with typical compliance. Once again, those with family histories benefit the most by perfect compliance.

Table IV. If Perfect Compliance

Gender	Race	Family History	Additional Cost	Life Years Gain	Cost per Life Year Gain
Female	Black	No family history	$428	0.042	$10150
Female	Black	Family history	Cost-saving	0.223	Cost-saving
Female	White	No family history	$266	0.063	$4182
Female	White	Family history	Cost-saving	0.303	Cost-saving
Male	Black	No family history	$287	0.045	$6350
Male	Black	Family history	Cost-saving	0.224	Cost-saving
Male	White	No family history	$157	0.053	$2946
Male	White	Family history	Cost-saving	0.269	Cost-saving

7. DISCUSSION AND CONCLUSIONS

A simulation model of a disease such as CRC provides greater fidelity than nonsimulation approaches that cannot capture all the intricacies of natural history, screening, and treatment. We created a modeling environment that presumes the need for detailed input representations and remodeling in order to accommodate a broad representation of stochastic input and complexities in the structure of the model. The Vanderbilt-NC State simulation model of Colorectal Cancer is a second-generation model of CRC that takes full advantage of current simulation technology to create a robust, extensible simulation platform for colorectal cancer modeling. The extensible object-oriented platform and the separation of the data from the model structure provides for revision and expansion of model components based on new information.

In the medical decision-making literature, the term *microsimulation* is often used, especially by economists, to describe any Monte Carlo-based model of a "micro" nature—as opposed to those of a "macro" nature. It is unfortunate terminology since it does not describe the character of the simulation. In fact, many medical "microsimulations" are pure Monte Carlo models, not discrete-event or next-event simulations. Many of these models are like our CRC model, namely, "patient or person based." In such cases, these models are somewhat like Markovian models, but with more complex input distributions. The CRC model reported here is discrete-event driven. Frequently discrete-event simulations represent a kind of queuing system where demand outstrips supply and there is often competition for resources, which in turn means a competition for time. In our CRC model, there is competition for time but not for resources. In particular, natural death and the occurrence and advancement of adenomas compete for the life of the individual. Since this competition occurs within a complex structure in which any one of these (random) events can occur, discrete-event simulation appears to be the only way to correctly execute these events in order and to advance time properly.

Random number and random variate generation is important in any simulation application. The use of a high-quality random number generator with well-spaced seeds along with the use of inverse-transform random variate generators provides not only a high-quality statistical process but also promotes correlation-induction variance reduction. By seeding each source of variation and by transforming only one random number into a random variate, different system configurations have synchronized random numbers. The use of

inverse-transform methods also facilitates antithetic sampling via complementary random numbers. The real value of synchronization was evident as alternative screening methods were considered. The additional computational burden of synchronization was of little concern since current computing speeds make the random number and random variate generation almost negligible in a discrete-event simulation.

The connection between computer models and programming is inescapable. Although our choice of working in the Microsoft realm may be viewed as a strong reliance on specific commercial products, it did facilitate the many program development requirements needed for the implementation of an object-oriented, discrete-event simulation of CRC, a database interaction, and a user-friendly interface. In particular, the creation of a database for the collection and maintenance of input for the simulations proved exceptionally useful as the model evolved. When attributes like gender and race are added to the model scale, many data items are affected and the data demand grows geometrically to the existing size of the database. By separating the data from the model structure, the model can be manipulated independently from the data and likewise the data can be changed without directly affecting the model structure.

There are many virtues of an object-oriented simulation. The key to successfully using an object-oriented simulation is in the design of the fundamental objects. Since we had extensive experience with object-oriented simulations (see Joines and Roberts [1998]), we wanted the basic simulation objects to be autonomous so that objects for the CRC model could be derived from the basic objects. We modified the general event handling by associating an event calendar with each adenoma because the evolution of the adenomas "competes" for the life of a person. This change significantly reduced execution time since events associated with adenomas that are removed during screening could be canceled without searching the event calendar. Other applications may need similar customization [Davies et al. 1993].

The actual modeling of CRC posed many challenges. Initially, we had to settle on the adenoma-carcinoma sequence, about which there was some controversy. There were numerous "assumptions" needed to describe the details that presented stochastic or distribution processes. We came to describe all of these as a "grand hypothesis" which we believe incorporates a consistent and defensible set of assumptions. There are no standards for revealing the details of a medical simulation model, so the grand hypothesis is a mechanism for making clear that a simulation model is in fact a model and not a substitute for fact.

The collection and representation of data was an arduous task that consumed much of the development effort. Although clinical trial data is helpful, it is only one set of information, usually gathered from a population with strict inclusion and exclusion criteria. It does not reveal what would happen if a different population were the experimental focus. Thus clinical trials must be augmented by critical clinical judgment. In large part, the model is a reflection of this clinical judgment, particularly that of the model coauthors. In our case, we were able to conduct a Delphi-like study to characterize some of the key questions, but the role of clinical judgment cannot be understated. Its influence extends from data

to model structure. It is another reason why model development transparency for medical decisions is critical. Fortunately, the subject of eliciting probability distributions and other stochastic inputs from "experts" is justifiably receiving more analytical attention [O'Hagan et al. 2006].

Our input modeling generally focused on specifying Johnson S_B and nonhomogenous Poisson processes when empirical distributions were insufficient or unavailable. The choice of the Johnson S_B was quite satisfactory, conforming to the intuitive belief and experience (asymmetric and bounded) about a wide variety of clinical parameters. Parameter estimates for the Johnson S_B were obtainable, although the beta distribution might have easily substituted for the Johnson S_B. The use of the nonhomogenous Poisson process was more often a matter of convenience since a piecewise linear estimate of the rate function was relatively easy to obtain (typically from national databases). However some concern remains with the adequacy of its representation, especially the variance about the rate due to the Poisson assumption. Unfortunately the data available did not allow us further investigation; nevertheless, how to best represent these stochastic processes should remain an open question.

The CRC model treats the life of one person as a replication since the trajectories of the natural history of CRC in individuals do not interact. By maintaining individual birth dates, either cohort or population simulations can be modeled (population statistics can be collected yearly from a date of reference). Simulating one person at a time may seem counterintuitive when in many discrete-event simulations the entities interact and statistics among them are correlated. However, when persons are considered individually as in the CRC model, the happy consequence is that person-based statistics are independent and the method of replications for output analysis applies. While this circumstance may not occur in all medical decision-making applications, it does appear to be a general result (for example, see Stevenson et al. [2005]). Nevertheless, a simulation structure that considers one person at a time is harder to extend to processes where patients need to compete for limited medical resources (such as doctors or hospitals) or where the process of treating one patient affects the treatment of another (as in a communicable disease).

Deciding on appropriate "measures of effectiveness" for a medical decision is particularly problematic in health care since so many interests are affected. We chose to measure cost, life years, and quality-adjusted life years. Cost is a major concern of society at large and life years clearly represent individual interests. Since the outcomes of decisions will be realized over time, discounting is appropriate for both the costs and life years [Gold et al. 1996]. Discrete discounting is applied to costs since they usually occur at specific points in time. However, in the case of life years, the discounting had to be continuously applied. Furthermore the whole notion of life years devoid of quality-of-life considerations does not reflect most values. But how to adjust quantity of life to include quality is controversial and no single quality measure is universally accepted. In our case, some work on utilities existed [Ness et al. 1999] and we used those "utility" adjustments to create quality-adjusted life years. Recognizing the need for making this adjustment, even if it is somewhat arbitrary, seems to us to be much preferred to ignoring the issue completely.

A simulation model of a medical decision is not an endpoint in the medical decision. Consequently, model development is often relegated to a footnote or brief paragraph in the description of the results from the use of a model (especially in the medical literature). There is increasing interest in producing more "transparent" medical decision-making models (see Weinstein et al. [2003]); however, almost none of the concerns address the complexity of simulation model development. No doubt part of the reason is that model development and model implementation are so closely allied. Nonetheless, clearly documenting a simulation model for public review minimizes the potential for its misuse.

A simulation model numerically provides a joint distribution of its outcomes as a function of all its inputs. Because a joint distribution of the outcome has more than just a mean value, there is the opportunity to incorporate variability into the analysis of alternatives. This topic is the subject of further work.

APPENDIX A. CRC SIMULATION DATABASE

Since the simulation model and its data are separated to the greatest extent possible, we store not only all simulation input information in a database but also all the output from any simulation experiment. Input and output are collected together with the simulation experimental conditions into a scenario template. The presence of the database allows previously constructed scenarios to be edited and run.

An Access database was chosen because (a) it is a relational database, which promotes easy storage of input through conventional tables and direct retrieval through standard SQL queries; (b) it is a distributed component of Microsoft Office, which is in widespread use and available on many desktops; and (c) it easily interfaces with .NET, our programming target. It is important to note that although Access is needed for development or revision of the database, it is not needed for simulation execution even though the simulation uses the Access database.

Within our CRC project, the CRC simulation "model" is not a single model but an instance from a model template. Each scenario within the family provides specific functionality, usually related to screening, surveillance, and treatment protocols of CRC. Each simulation scenario uses different variables for its execution and thus requires different input data. However, the base CRC natural history simulation structure is driven by a fixed set of variables in the database and all models draw upon this source for any specific scenario.

A summary of the relationships among the major tables within the database is shown in Figure 6. These include the population, distribution, input, and output variable definitions and specifications.

APPENDIX B. INPUT MODELING

VIM (Visual Input Modeler) [Roberts 2004] was inspired by earlier work with VISIFIT [DeBrota et al. 1989]. Although the software can display a wide range of distributions with multiple statistical characteristics and provide graphs of the density function, distribution function, and hazard function which interactively change according to changes in the distribution parameters, the

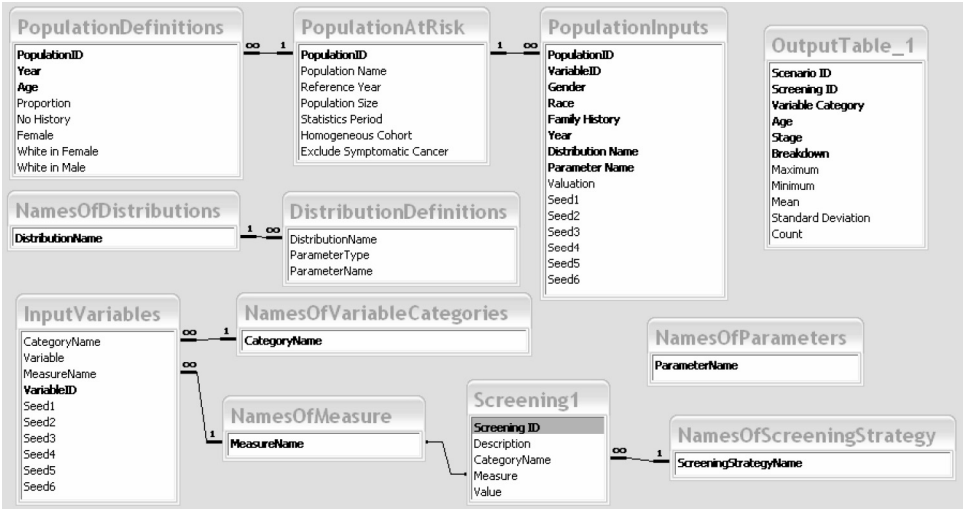

Fig. 6. Relationships among tables in database.

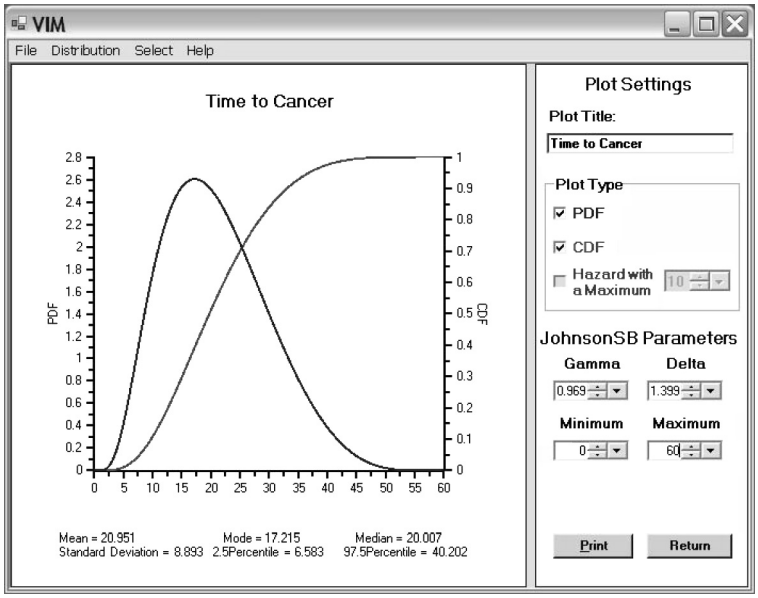

Fig. 7. VIM interactive display.

same functionality is available in commercial products like Stat:Fit (`http://www.geerms.com`), BestFit (`http://www.palisade.com`), ExpertFit (`http://www.averill-law.com`), and EasyFit (`http://www.mathwave.com`). However, VIM was integrated into the CRC simulation user interface so CRC input could be visually presented and characteristics could be potentially modified as desired. The VIM main display that also corresponds to the "Plot Variable" button in Figure 2 is shown in Figure 7.

APPENDIX C. RANDOM NUMBERS AND RANDOM VARIATES

To further reduce runtime for the analysis of a single scenario, antithetic variates can be employed (up to 40% reductions in runtime were observed using antithetic variates). The use of antithetic variates is facilitated by the ease with which complementary random numbers are generated from the random number generator and the use of inverse transform generators to obtain antithetic variates with maximum negative correlation. Total time for random variate generation within the simulation was less than 1% of computational time.

To insure that each input variable has its own random number source while the model is being changed and to keep changes in the structure of the simulation from affecting the input (for instance, with screening options), the automatic seed selection in the original generator [L'Ecuyer et al. 2002] had to be separated so the generator would use stored seeds (each stream depends on a vector of six particular integers). Seeds for each variable are now stored with the random variable definition in the database. A separate program computes appropriately spaced seeds (1.7×10^{38} numbers apart). Whenever a new random variable is added to the database, its stream seeds are also specified.

In the simulation, each new person (replication) in the model begins sampling from the next substream for the random variable. With a generator of 1.8×10^{19} streams being used for the random variates, where each person is associated with one of 2.3×10^{15} substreams, and each substream has about 7.6×10^{22} random numbers available to simulate the person, the magnitudes of these characteristics are far larger than any of our needs.

Random variates are generated through "inheritance" in an object-oriented design. Since inheritance is a fundamental design mechanism in .NET, all random variates are inherited from the base random number class of the (modified) random number generator and designed to sample from the random object class stream seed. Empirically based distributions use functional inverses that are numerically constructed. The random variates available are Johnson S_B, Johnson S_U, beta, gamma, Weibull, triangular, exponential, Erlang, normal, lognormal, uniform, Dirichlet, empirical discrete, empirical continuous, some multivariate distributions, and the nonhomogenous Poisson process with a piecewise linear rate function.

Multivariate distributions and serially correlated samples are employed both in the simulation and its subsequent analysis. The multivariate beta, as well as other multivariate distributions, are generated using the NORTA procedures described in Law and Kelton [2000]. The multivariate beta is needed to express the correlation that is expected among costs and test characteristics when performing a probabilistic sensitivity analysis [Tafazzoli 2004]. This process is also applied to the serial correlation desired for the interarrival times of adenomas.

APPENDIX D. PERSON AND ADENOMA OBJECT DETAILS

The person object properties can be subdivided into six sets of variables. The *common variables* define this individual in the population and consist of (1) Birth Year, which in conjunction with year of simulation computes age; (2) Gender, male or female; (3) Race, white or black; and (4) Family History,which

is a binary variable denoting the presence or absence of CRC in a first-degree relative. The person is also described by certain *unique variables* that include (1) ID, the unique identification number used to track the individual through the simulation; (2) Natural Life, which is the expected time of death without CRC; (3) Overall Risk, which determines the individual's risk for CRC; and (4) the compliance or extent to which this person adheres to screening regimens. The *adenoma variables* define properties of the adenoma objects associated with a person and include (1) Number of Adenomas, which keeps track of all adenomas and CRC formed in a person; (2) Health State, which describes the health state as determined by adenoma status; (3) Cancer State, which reflects CRC stage; (4) Symptomatic, which indicates whether or not the person has symptoms related to CRC; (5) Resection, which notes which sections of the colon have been removed due to CRC. The *cost variables* used in the model include (1) Accumulated Cancer Care Cost, which tracks costs associated with CRC treatment to date for a person; and (2) Accumulated Screening Cost, which tracks costs associated with the performance of a diagnostic test, pathology on biopsy and polypectomy specimen, and treatment of test complications. There are also *utility variables* [Ness et al. 1999] including (1) Utility, a quality of life measurement associated with the current CRC-related health state; (2) Accumulated Life, which accumulates the life years associated with a person; and (3) Accumulated Quality of Life, which accumulates the quality-adjusted life years associated with a person. Finally, there are *screening variables* that include (1) Adherence, being the person's maximum adherence rate with any specific screening test; (2) Screening Strategy, denoting the screening strategy being employed; (3) Diagnostic, which denotes whether or not the person is awaiting diagnostic colonoscopy (for symptoms or to follow up a positive screening test); and (4) Surveillance, which denotes whether or not the patient is following a surveillance regimen.

Each adenoma object holds the following characteristics (1) ID, a unique tracking number; (2) Progression Type: immediately progressive to CRC, progressive to CRC, nonprogressive; (3) Adenoma Location: right colon, left colon, sigmoid, upper rectum, or lower rectum; (4) Histology (microscopic structure and size) of the adenoma: nonadvanced adenoma, advanced adenoma, or CRC. It is important to note that the V/NCS model is the only model of CRC to incorporate histology into the adenoma evolution. Other models use adenoma size, which is somewhat correlated with histology, however histology is a better clinical predictor of CRC development.

ACKNOWLEDGMENTS

The authors thank Cindy Liebsch, Dan Cubbage, and Ali Tafazzoli whose master's theses research was fundamental to this work.

REFERENCES

Ahsan, H., Neugut, A. I., Garbowski, G. C., Jacobson, J. S., Forde, K. A., Treat, M. R., and Waye, J. D. 1998. Family history of colorectal adenomatous polyps and increased risk for colorectal cancer. *Ann. Intern. Med. 128*, 900–905.

Banks, J. 2004. *Getting Started with AutoMod*. Brooks Automation, Inc., Chelmsford, MA.

Boutron, M. C., Faivre, J., Quipourt, V., Senesse, P., and Michiels, C. 1995. Family history of colorectal tumours and implications for the adenoma-carcinoma sequence: A case control study. *Gut 37*, 830–834.

Brown, M. L., Klabunde, C. N., and Mysliwiec, P. A. 2003. Current capacity for endoscopic colorectal cancer screening in the United States: Data from the National Cancer Institute Survey of Colorectal Cancer Screening Practices. *Amer. J. Med. 115*, 129–133.

Cancer Stat Fact Sheets. 2006. Cancer of the Colon and Rectum, National Cancer Institute. Available online at `http://seer.cancer.gov/statfacts/html/colorect.html`.

Caro, J. J. 2005. Pharmacoeconomic analyses using discrete event simulation. *Pharmacoeconomics 23*, 323–332.

Chan, T. F. and Lewis, J. G. 1979. Computing standard deviations. *Commun. ACM 22*, 526–531.

Cubbage, D. 2004. Simulation of colorectal cancer: The natural history of disease. M. S. thesis. Department of Industrial Engineering and Graduate Program in Operations Research, North Carolina State University, Raleigh, NC.

Davies, R., Roderick, P., and Raftery, J. 2003. The evaluation of disease prevention and treatment using simulation models. *Eur. J. Oper. Res. 150*, 53–66.

Davies, R. M., O'Keefe, R. M., and Davies, H. T. O. 1993. Simplifying the modeling of multiple activities, multiple queuing, and interruptions: A new low-level data structure. *ACM Trans. Model. Comput. Simul. 3*, 332–346.

Debrota, D. J., Dittus, R. S., Roberts, S. D., and Wilson, J. R. 1989. Visual interactive fitting of bounded Johnson distributions. *Simulation 52*, 199–205.

Eddy, D. M. 1980. *Screening for Cancer: Theory, Analysis and Design*. Prentice-Hall, Englewood Cliffs, NJ.

Fuchs, C. S., Giovannucci, E. L., Colditz, G. A., Hunter, D. J., Speizer, F. E., and Willett, W. C. 1994. A prospective study of family history and the risk of colorectal cancer. *New England J. Med. 331*, 1669–1674.

Gold, M. R., Siegel, J. E., Russell, L. B., and Weinstein, M. C. 1996. *Cost-Effectiveness in Health and Medicine*. Oxford University Press, New York, NY.

Habbema, J. D., van Oortmarssen, G. J., Lubbe, J. T., and van der Maas, P. J. 1984. The MISCAN simulation program for the evaluation of screening for disease. *Comput. Meth. Progr. Biomed. 20*, 79–93.

Harrell, C. R., Bhosh, B. K., and Bowden, R. O. 2004. *Simulation Using Promodel*. McGraw-Hill, New York, NY.

Joines, J. A. and Roberts, S. D. 1998. Object-oriented simulation. In *Handbook on Simulation*, J. Banks, Eds. John Wiley & Sons, New York, NY, 100–150.

Karnon, J. 2003. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ. 12*, 837–848.

Kelton, W. D., Sadowski, R. P., and Sturrock, D. T. 2004. *Simulation with Arena*, 3rd ed. McGraw-Hill Higher Education, Boston, MA.

L'Ecuyer, P. 1999. Good parameters and implementations for combined multiple recursive random number generators. *Oper. Res. 47*, 159–164.

L'Ecuyer, P. L., Simard, R., Chen, J. E., and Kelton, D. W. 2002. An object-oriented random-number package with many long streams and substreams. *Oper. Res. 50*, 1073–1075.

Law, A. M. and Kelton, D. W. 2000. *Simulation Modeling and Analysis*, 3rd ed. McGraw-Hill Higher Education, Boston, MA.

Liebsch, C. 2003. Simulation modeling in the absence of data. M. S. thesis. Department of Industrial Engineering and Graduate Program in Operations Research, North Carolina State University, Raleigh, NC.

Loeve, F., Boer, R., van Oortmarssen, G. J., Van Ballegooijen, M., and Habbema, J. D. 1999. The MISCAN-Colon simulation model for the evaluation of colorectal cancer screening. *Comput. Biomed. Res. 32*, 13–33.

Loeve, F., Boer, R., Zauber, A. G., van Ballegooijen, M., van Oortmarssen, G. J., Winawer, S., and Habbema, J. D. 2004. National polyp study data: Evidence for regression of adenomas. *Int. J. Cancer 111*, 633–639.

LOEVE, F., BROWN, M. L., BOER, R., VAN BALLEGOOIJEN, M., VAN OORTMARSSEN, G. J., AND HABBEMA, J. D. 2000. Endoscopic colorectal cancer screening: A cost-saving analysis. *J. Nat. Cancer Inst. 92*, 557–563.

MANDEL, J. S., BOND, J. H., CHURCH, T. R., SNOVER, D. C., BRADLEY, G. M., SCHUMAN, L. M., AND EDERER, F. 1993. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *New England J. Med. 328*, 1365–1371.

MANNING, W. G., FRYBACK, D. G., AND WEINSTEIN, M. C. 1996. Reflecting uncertainty in cost-effectiveness analysis. In *Cost-Effectiveness in Health and Medicine*, M. R. Gold, J. E. Siegel, L. B. Russell, and M. C., Weinstein, Eds. Oxford University Press, New York, NY, 247–75.

NESS, R. M., HOLMES, A. M., KLEIN, R., AND DITTUS, R. 1999. Utility valuations for outcome states of colorectal cancer. *Amer. J. Gastroent. 94*, 1650–1657.

NESS, R. M., HOLMES, A. M., KLEIN, R., AND DITTUS, R. 2000. Cost-utility of one-time colonoscopic screening for colorectal cancer at various ages. *Amer. J. Gastroent. 95*, 1800–1811.

O'HAGAN, A., BUCK, C. E., DANESHKHAH, A., EISER, R., GARTHWAITE, P. H., JENKINSON, D. J., OAKLEY, J. E., AND RAKOW, T. 2006. *Uncertain Judgements: Eliciting Expert Probabilities*. John Wiley & Sons, New York, NY.

PARIENTE, A., MILAN, C., LAFON, J., AND FAIVRE, J. 1998. Colonoscopic screening in first-degree relatives of patients with 'sporadic' colorectal cancer: A case-control study. *Gastroent. 115*, 7–12.

PIGNONE, M., RUSSELL, L., AND WAGNER, J. 2005. *Economic Models of Colorectal Cancer Screening in Average-Risk Adults: Workshop Summary*. The National Academies Press, Washington, DC.

PIGNONE, M., SAHA, S., HOERGER, T., AND MANDELBLATT, J. 2002. Cost-effectiveness analyses of colorectal cancer screening: A systematic review for the U.S. Preventive Services Task Force. *Ann. Intern. Med. 137*, 96–104.

ROBERTS, S. D. 2004. Simulation input modeling without data. In *Proceedings of the 6th Annual Simulation Solutions Conference* (Orlando, FL; Institute of Industrial Engineers, Atlanta, GA). 1–34.

SCHRUBEN, L. 1983. Simulation modeling with Event Graphs. *Commun. ACM 26*, 957–963.

SMITH, R. A., COKKINDES, V., AND EYRE, H. J. 2004. American cancer society guidelines for the early detection of cancer. *Ca: Cancer J. Clinicians 54*, 41–52.

ST. JOHN, D. J., MCDERMOTT, F. T., HOPPER, J. L., DEBNEY, E. A., JOHNSON, W. R., AND HUGHES, E. S. R. 1993. Cancer risk in relatives of patients with common colorectal cancer. *Ann. Intern. Med. 118*, 785–790.

STEVENSON, M. D., BRAZIER, J. E., CALVERT, N. W., LLOYD-JONES, M., OAKLEY, J. E., AND KANIS, J. A. 2005. Description of an individual patient methodology for calculating the cost-effectiveness of treatments for osteoporosis in women. *J. Oper. Res. Soc. 56*, 214–221.

TAFAZZOLI, A. 2004. A comparison of screening methods for colorectal cancer. M. S. thesis. Department of Industrial Engineering and Graduate Program in Operations Research, North Carolina State University, Raleigh, NC.

TAFAZZOLI, A., ROBERTS, S. D., NESS, R. M., AND DITTUS, R. S. 2005. A comparison of screening methods for colorectal cancer using simulation modeling. In *Proceedings of the 2005 Winter Simulation Conference* (Orlando, FL), M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds. IEEE Computer Society Press, Los Alamitos, CA, 2236–2245.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES. 2004. SEER cancer incidence public-use database, 1973–2001; mortality data 1969–2001. U. S. Department of Health and Human Services, Washington, DC.

VERNON, S. W. 1997. Participation in colorectal cancer screening: A review. *J. Nat. Cancer Inst. 89*, 1406–1422.

WEINSTEIN, M. C., O'BRIEN, B., HORNBERGER, J., JACKSON, J., JOHANNESSON, M. M. C., AND LUCE, B. R. 2003. Principles of good practice for decision analytic modeling in health-care evaluation: Report of the ISPOR task force on good research practices—Modeling studies. *Val. Health 6*, 9–17.

WEST, D. H. 1979. Updating mean and variance estimates: An improved method. *Commun. ACM 22*, 532–535.

WILMOTH, J. R. 2003. Berkeley mortality database, university of California, Berkeley, Berkeley, CA. Go online to `http://www.demog.berkeley.edu/wilmoth/mortality`.

WINAWER, S., FLETCHER, R. H., REX, D., BOND, J., BURT, R., FERRUCCI, J. T., GANIATS, T. G., LEVIN, T., WOOLF, S. H., JOHNSON, D., KIRK, L., LITIN, S., SIMMANG, C., AND GASTROINTESTINAL CONSORTIUM PANEL. 2003. Colorectal cancer screening and surveillance: Clinical guidelines and rationale—update based on new evidence. *Gastroenterology 124*, 544–560.

WINAWER, S. J., ZAUBER, A. G., GERDES, H., O'BRIEN, M. J., GOTTLIEB, L. S., STERNBERG, S. S., BOND, J. H., SCHAPIRO, M., AND PANISH, J. F. 1996. Risk of colorectal cancer in the families of patients with adenomatous polyps. National Polyp Study Workgroup. *New England J. Med. 334*, 82–87.

WINAWER, S. J., ZAUBER, A. G., HO, M. N., O'BRIEN, M. J., GOTTLIEB, L. S., STERNBERG, S. S., STEWART, E. T., BOND, J. H., SCHAPIRO, M., PANISH, J. F., ET AL. 1993. The National Polyp Study. *Eur. J. Cancer Prevent. 2*, Suppl 2:83–7.

Received September 2006; revised January 2007, February 2007, March 2007; accepted March 2007