



Published in final edited form as:

Med Decis Making. 2021 May ; 41(4): 453–464. doi:10.1177/0272989X21995805.

Comparison of Decision Modeling Approaches for Health Technology and Policy Evaluation

John Graves, Ph.D.^{1,*} [Associate Professor], Shawn Garbett, MS² [Assistant in Biostatistics], Zilu Zhou, MPH³ [Senior Statistical Analyst], Jonathan S. Schildcrout, Ph.D⁴ [Associate Professor], Josh Peterson, MD, MPH⁵ [Professor]

¹Department of Health Policy, Vanderbilt University School of Medicine, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 1200, Nashville, TN 37203, (615) 343 – 4875

²Vanderbilt University Medical Center, Department of Biostatistics

³Vanderbilt University Medical Center Department of Health Policy

⁴Department of Biostatistics, Vanderbilt University School of Medicine, Vanderbilt University Medical Center

⁵Vanderbilt University School of Medicine, Vanderbilt University Medical Center, Department of Biomedical Informatics

Abstract

We discuss tradeoffs and errors associated with approaches to modeling health economic decisions. Through an application in pharmacogenomic (PGx) testing to guide drug selection for individuals with a genetic variant, we assessed model accuracy, optimal decisions and computation time for an identical decision scenario modeled four ways: using (1) coupled-time differential equations [DEQ]; (2) a cohort-based discrete-time state transition model [MARKOV]; (3) an individual discrete-time state transition microsimulation model [MICROSIM]; and (4) discrete event simulation [DES]. Relative to DEQ, the Net Monetary Benefit for PGx testing (vs. a reference strategy of no testing) based on MARKOV with rate-to-probability conversions using commonly used formulas resulted in different optimal decisions. MARKOV was nearly identical to DEQ when transition probabilities were embedded using a transition intensity matrix. Among stochastic models, DES model outputs converged to DEQ with substantially fewer simulated patients (1 million) vs. MICROSIM (1 billion). Overall, properly embedded Markov models provided the most favorable mix of accuracy and run-time, but introduced additional complexity for calculating cost and quality-adjusted life year outcomes due to the inclusion of “jumpover” states after proper embedding of transition probabilities. Among stochastic models, DES offered the most favorable mix of accuracy, reliability, and speed.

Keywords

Decision modeling; pharmacogenomics; health economic methods

*Corresponding Author: john.graves@vanderbilt.edu.

Introduction

Decision analytic assessments of health policies and technologies play an important role in determining the pricing and reimbursement of health care services worldwide. These assessments draw on a range of modeling methods to inform policy and clinical decision making. Commonly used methods include decision trees, cohort and individual state transition (Markov) models, discrete event simulation, stochastic process theory models based on differential equations, and hybrid models that blend elements across approaches (e.g., Discretely Integrated Condition Event models).^{1–6}

For a given health technology assessment or research question, best practice recommends model choice based on consultation with clinical, methodological, and policy experts.^{5,7} This ensures that the model represents essential elements of the underlying disease or therapeutic processes, as well as the alternative strategies under consideration.

Our objective for this study was to compare alternative approaches to modeling health economic decisions. Specifically, we assessed model outputs, optimal decisions and model computation time for an identical decision problem modeled four ways: using a discrete time cohort state transition (MARKOV) model, an individual discrete time state transition microsimulation (MICROSIM) model, discrete event simulation (DES), and a stochastic theory-based model based on time delay differential equations (DEQ). Importantly, these models drew on an identical set of underlying parameters that should, we hypothesized, lead to equivalent (in expectation) cost and quality-adjusted life year (QALY) estimates and decision outcomes. Our results demonstrate that using commonly-applied discrete time model structure and adjustment methods, they do not.

An important contribution of our study is the exposition of key sources of modeling error and how they can affect the accuracy, reliability and resolution of model outputs. We demonstrate through an application how these errors can play out in practice. In particular, we show that adjustments must be made to discrete time individual and cohort state transition models to produce equivalent estimates as DES and DEQ models. These adjustments are necessary due to an interaction between competing events and the coarsening of continuous time into discrete time cycles—and are important in (common) situations where literature-based parameters are embedded within transition probability matrices in discrete-time Markov models. Finally, we also show that to achieve a given degree of output reliability, discrete time microsimulation models require considerably more Monte Carlo draws than a DES model. We include replication code (for the R statistical programming language) and simulation functions for all model types considered.

Background

Modeling Types

The following section provides basic information on the model types we consider. This information is not intended to be exhaustive of the details and assumptions of each method; we leave such comprehensive treatment to the literature.^{1,8,9} Rather, we highlight specific

aspects of each modeling approach that are germane to issues of output scope, resolution, variance/bias, and computational efficiency that we highlight in our results.

Discrete Time Cohort and Individual State Transition Models—State transition models (STM) are the most widely used modeling framework in health economic evaluation.³ STMs are conceptualized around a set of mutually exclusive and collectively exhaustive states (e.g., healthy, sick, dead). Models are further defined by an initial occupancy vector (i.e., a count or fraction of patients in each state at baseline) and a set of state transition rates (for a continuous time model) or probabilities at a specified cycle length (for a discrete time model). The vast majority of health economic modeling applications are based on the discrete time case,⁶ so we focus on that model type here.

In a discrete-time STM, an initial occupancy vector is modeled forward using a transition probability matrix for a defined time horizon or until all individuals enter an absorbing state (e.g., death). Values, such as health-related quality of life and costs, are assigned to each state. These state values, along with information on the amount of person-time spent in each state, are summarized (often with discounting) to produce outcome estimates on various quantities of interest (e.g., expected life expectancy, expected quality-adjusted life expectancy, and expected costs).

A key decision when considering a STM is whether to jointly simulate the collective experience of a cohort, or to simulate individual patient trajectories via first order Monte Carlo microsimulation. Advantages of cohort STMs (often referred to as Markov models) include their simplicity, computational efficiency, and ease of use. A disadvantage is that the Markovian “memoryless” assumption complicates applications in which patient history affects transition probabilities.¹⁰ While tunnel states provide a workaround, they can result in models with large numbers of states (i.e., “state explosion”).¹⁰

Microsimulation models are useful in applications where state transition probabilities depend on a patient’s history (e.g., time since disease onset, the occurrence of previous events, or time-varying response to treatment).⁴ More broadly, microsimulation models facilitate transition probabilities that are a function of any number of attributes. Microsimulation also affords the flexibility to capture a greater scope of outputs since the model can return estimates of the entire distribution of events, rather than just expected values. This level of detail could be important for informing multi-criteria decision analyses that base decisions on a wider array of criteria beyond expected quality-adjusted life expectancy and costs.^{11–14}

The flexibility afforded by microsimulation comes at a cost, however, since the computational burden is large. As described below, these computational demands derive from three primary sources: first-order (stochastic) error; the need to simulate all time cycles for each individual, even if there is no event; and an additional increase in model output variance and bias that occurs when the precise timing of events are truncated to occur at cycle boundary points.

Discrete Event Simulation—Discrete Event Simulation (DES) can also incorporate the timing and interdependency of events.^{7,15,16} Though its origins are in industrial engineering and operations research, DES is increasingly used in health technology assessments.^{3,5,15–17}

DES models are similar to microsimulation models in that they simulate individual patient trajectories. As such, they can also be computationally demanding. One advantage, however, is that DES extends the flexibilities of microsimulation. For example, a DES model can allow the probability of some future event to depend on the time spent in a given state. In addition, DES models can more easily incorporate interdependencies in the timing of events and/or restrictions on available resources based on queuing or other constraints. An additional advantage is that DES models the timing of events and, unlike microsimulation, does not cycle through time periods when no events occur. These advantages afford DES more modeling flexibility and computational efficiency.

Differential Equations Modeling—Many decision and disease processes can be represented using stochastic process theory to deterministically solve for key outcomes.⁶ These processes can be represented in terms of differential equations (DEQs), the solution to which provides the occupancy of an underlying process at any time t . Expected values of key outcomes (e.g., average discounted QALYs and costs) are obtained by integrating this solution from baseline (i.e., $t=0$) to some specified time T . Solutions to this integration problem exist for health economic applications in which transition rates are constant over time and/or evolve as a piecewise function of age.⁶ Standard differential equation models conform to the Markovian “memoryless” property, while coupled time delay differential equations can model processes in which the solution at any time depends on previous values of the function.¹⁸

Estimates derived from the numerical solution to a DEQ deterministically solve for expected outcomes, and thus eliminate stochastic uncertainty inherent to DES and microsimulation. A downside, however, is that the scope of modeled outputs is limited: DEQ solves for average outcomes but does not provide insights into higher order moments (e.g., variance). Moreover, model specification is challenging—making even minor changes to model structure is a nontrivial task. DES and microsimulation methods, by comparison, return information on the entire distribution of events and can be constructed and modified without advanced mathematical training.

A Taxonomy of Errors in Health Economic Modeling

As discussed above, modeling choices affect computational demands and the range of analyzable outputs. Model structure also determines the types of error—that is, differences between quantities of interest based on the underlying event generation process and the model’s output—that must be accounted for. Some errors are well-known and are shared across modeling approaches. Others are more nuanced, less widely acknowledged, and specific to certain modeling strategies. Finally, depending on their size, errors may or may not affect optimal decisions—though *ex ante* it is difficult to know which errors could affect decision-making for any given application.

This section, as well as Table 1, outlines a taxonomy of modeling error. We focus on errors that can affect estimates and decisions, and further distinguish between **errors of reliability and uncertainty** (i.e., error that can be addressed by increasing the number of simulated patients or through further research) and **errors in expectation** (i.e., errors that, even with an infinite number of simulated patients or with parameters with little-to-no uncertainty, would yield biased estimates and, possibly, different optimal decisions). We discuss ways to adapt model specification and execution to reduce and/or eliminate sources of error, and cover specific details on their implementation in the methods section and appendix.

To understand our error taxonomy it is useful to conceptualize the modeled process (e.g., the progression of a disease) occurring in continuous time. Ideally, one has data on the precise timing of all events experienced by the population of interest. In that case, one can estimate all competing event rates directly (e.g., using a multi-state model of event duration outcomes). Transitions among discrete states, moreover, can be summarized in terms of transition intensities in continuous time.

In practice, models must deviate from this ideal in meaningful ways. Researchers often do not have access to comprehensive data on the population of interest. Model parameters must therefore be curated from the published literature, or may be estimated via calibration with aggregate data on disease prevalence, incidence, etc.¹⁹ Literature-based parameters can take on any number of “flavors” (e.g., estimated event rates, odds ratios, probabilities, etc.) and there is no guarantee that what is published maps directly into the preferred model structure. Therefore, parameters are often transformed (e.g., rates are converted into probabilities for the selected cycle length) before embedding them within the model.

Even when the underlying data or parameters are available, assumptions are often made over whether model parameters are fixed or evolve differentially over time and/or with specific patient attributes (e.g., race, age, gender, etc.). These questions are particularly germane to models that draw on literature-based parameters because the population data underlying these parameters often do not align with the population of interest.

We define **structural error** as attributes of models and model parameters (e.g., time-dependency of transition rates/probabilities, the number of overall states in the model) that result in model estimates that deviate from the “true” underlying event generation process. It is worth differentiating structural errors from parameter heterogeneity, or differences in model parameters across relevant population sub-groups. With heterogeneity, the overall structure of the model could accurately represent the event generation process. However, decisions could be improved by estimating model outputs and making separate decisions for each sub-population.²⁰

Model parameters are also often subject to sampling and estimation uncertainty. **Estimation errors** are inherent to all model types discussed here. Understanding the role of estimation uncertainty is the focus of a growing body of research on VOI.^{21–24} However, for the purposes of this study we sidestep structural and estimation errors because we specify an underlying event generation process in which the structure and values of the parameters are pre-specified.

Even when a parameter's "true" value is known, the realization and timing of events among otherwise identical patients may vary. This stochastic (first-order) variation is inherent to modeling types that utilize Monte Carlo methods to simulate individual patient trajectories. For these methods, a series of simulations, each with fixed simulated patient size M , will produce varying estimates of model outputs. Absent any other errors, this series of simulations will yield a set of estimates that center around the outcome's expected value—and as M increases this set of estimates will converge towards this value. But as long as M is finite, any given model simulation result will deviate from the expected value. Mirroring the literature, we classify this type of error as **stochastic error**.²⁰ Cohort STMs and DEQ modeling do not suffer from stochastic error because expected values of model outcomes are deterministic solutions from the model.

Stochastic error can never be eliminated (doing so would require simulating an infinite population) but it can be reduced. Best practice recommends a sufficient number of Monte Carlo draws to reduce stochastic error to within an acceptable tolerance level²⁰—though as we show below, the number of simulated patients needed to achieve a fixed tolerance level is significantly higher with microsimulation than with DES.

Models that coarsen the timing of events into discrete cycles (e.g., cohort-based Markov models and discrete-time microsimulation models) can create additional errors. **Integration error** occurs when the model accumulates events at time cycle boundaries (either at the beginning or the end) rather than at any point in continuous time. Applications in health economics and health technology assessment frequently rely on half-cycle corrections or the life-table method, but other methods (e.g., those based on Simpson's rule) can achieve an even greater degree of accuracy.²⁵

Discrete time models (e.g., MARKOV, MICROSIM) exhibit **embedding error** if transition probabilities are not specified to align with the underlying continuous time process. These errors are most likely to occur in models that rely on literature-based parameters that are transformed and embedded using widely-used conversion formulas (e.g., conversion of event rates to probabilities using $P = 1 - e^{-rt}$, where r is the rate and t is the time cycle). These errors are also more likely to occur in circumstances where events are clustered near each other (e.g., time-varying probabilities of complications or events after a procedure, drug initiation, etc.).

Embedding errors arise because if not applied properly, standard conversion formulas do not separately accommodate more than one competing event.^{6,26–29} This observation is relevant because in many applications, the modeled process is conceptualized around multiple competing events or sequences of events occurring in continuous time (e.g., progression to or among various disease states or to an absorbing death state).

If probabilities are not embedded properly, standard rate-to-probability conversion processes effectively rule out the possibility of two or more events occurring in the same cycle. Suppose that the underlying process progresses along three states: $A \rightarrow B \rightarrow C$. In our application below, for example, healthy (A) individuals can develop a chronic disease (B) requiring a daily maintenance drug that can result in an adverse event (C). When specified to

approximate the underlying continuous time process, the probabilities in a transition matrix should reflect the possibility of both B and C occurring in the same cycle. That is, from state A there are two event sequences that can happen within a single cycle: (1) $A \rightarrow B$; and (2) $A \rightarrow B \rightarrow C$.

Accounting for these event sequences requires an embedded transition matrix that includes a non-zero $A \rightarrow C$ transition probability—that is, a transition that is seemingly inconsistent with the underlying disease progression process. This need arises because the simple conversion of the rate of B (r_B) to the probability of B (i.e., $P(B) = 1 - e^{-r_B t}$) returns the marginal probability of event B occurring within the cycle. This marginal probability is the sum or union of *all* the probabilities of experiencing the event, i.e., $P(B) = P(A \rightarrow B) + P(A \rightarrow B \rightarrow C)$.ⁱ Thus, embedding the marginal probability as the probability of $A \rightarrow B$ transitions in a cycle will both overstate the correct $A \rightarrow B$ transition probability and rule out any $A \rightarrow B \rightarrow C$ transitions within a cycle.

This implicit assumption (i.e., that there are no $A \rightarrow B \rightarrow C$ transitions within the time cycle) means that an improperly embedded transition matrix no longer approximates a continuous time process. More technically, it means that the process may no longer be expressed in terms of an underlying generator matrix of continuous transition intensities.³⁰ Specifying a transition matrix consistent with an underlying generator matrix is important not only for accurate modeling of the underlying continuous time process, but also for changing Markov cycle lengths.^{26,31} Otherwise, matrix transformation routines (e.g., based on unit roots from eigendecomposition of a specified transition probability matrix) can result in transformed transition matrices with negative probabilities and/or without unique solutions.^{26,31}

Approaches to addressing embedding error complicate the structure and execution of discrete time cohort and microsimulation models.ⁱⁱ If transitions can first be specified (or estimated) as rates and summarized in a transition intensity matrix, then properly embedding a transition probability matrix is straightforward: one simply takes the matrix exponential of the transition intensity matrix scaled by the time cycle duration. The result is the solution to Kolmogorov's forward equation, and this matrix can be used as the transition probability matrix for the model.^{27,29} Note, however, that this transformation process requires the inclusion of non-Markovian event “accumulators” in the transition matrix. These accumulators ensure that model outputs (e.g., average QALYs and costs) account for events that occur during “jump-over” states that arise from the embedding process (i.e., the model must be structured to account for any cost or utility implications of event B for the $A \rightarrow C$ transitions in the example above). The appendix and replication code to this study demonstrates how to incorporate non-Markovian accumulators within a transition probability matrix.

ⁱThe same logic still holds if events B and C were simply competing and not sequential events, in which case the marginal probability of event B reflects $P(A \rightarrow B)$ and $P(A \rightarrow B \& A \rightarrow C)$.

ⁱⁱOne tempting solution for embedding error is to reduce the cycle length so that the likelihood of multiple transitions within a single cycle approaches zero. Shortening the cycle length carries additional computational costs. More importantly, a unique solution to a time step transformation still requires a generator matrix of continuous transition intensities as specified above. Otherwise, converting (marginal) literature-based probabilities to rates, and then using these rates to change the time step (e.g., using eigendecompositions) can result in nonunique solutions and/or even negative probabilities.³¹ Thus, when starting from a transition probability matrix (as opposed to a transition intensity matrix), shortening the time step does not fundamentally address embedding error issues.

Finally, microsimulation models also suffer from **truncation error**, which introduces variance and a small amount of bias to the modeled estimates. This error occurs because information on the precise timing of an event occurring in continuous time is lost when events times are “rounded up” or “rounded down” to the boundary point in the cycle. These rounding errors, when aggregated across individual simulated patients in a microsimulation, distort both the mean and the variance of outcome estimates relative to what would be estimated in a continuous time model (e.g., DES) that can specify each individual’s event timing to any degree of precision.³²ⁱⁱⁱ

Recall from the discussion above that relative to DES, microsimulation models require a greater number of simulated patients to achieve the same tolerance level. This requirement derives, in part, from the increase in variance from truncation error.^{iv} To reduce the influence of truncation error, modelers can increase the number of simulated patients or can reduce the cycle length (e.g., from 1 year to 1 day) to allow for more precise timing of events. Both of these strategies carry additional computational demands, however.

Methods

Application: Pharmacogenomics

Our modeling application draws on ongoing work in pharmacogenomics (PGx). PGx involves the use of genetic testing to guide drug selection and/or dosing based on associations between genetic phenotypes and drug metabolism. All model parameters (shown in Table 2) were selected to both approximate a common PGx scenario and to produce results that highlight differences across modeling approaches.

We focus on a common PGx decision problem: whether to utilize testing to guide drug selection for individuals with a genetic variant that can affect metabolism of a chronic disease medication. Specifically, we consider a population of healthy 40-year-old women at risk of developing an indication (10% incidence rate over 10 years) for a health condition with a standard maintenance medication therapy, but where there is a more expensive and more effective pharmacogenomic alternative available to individuals with a variant.

We assume all who develop the chronic condition incur a one-time treatment cost of \$10,000, a transient (one-year) 0.05 utility decrement, and are placed on daily maintenance medication (\$0.25 per day) for life. Individuals on this medication are at risk of an adverse event that affects 25% of individuals in the first year they are on the drug. The adverse event has a 10% case fatality rate, carries a one-time \$20,000 cost among decedents and, among the survivors, a one-time \$25,000 cost and 0.1 utility decrement for life.

ⁱⁱⁱIntuitively, integration error occurs when events are *moved* from their precise time to the beginning or end of the cycle. Truncation error arises because information on the precise timing of events is lost when all events in the cycle are truncated to the cycle boundary point.

^{iv}In the Appendix, we demonstrate how truncation error can affect model output variance by constructing a straw-man DES model that truncates DES event times at a time interval (year). Because DES models model the precise timing of events, a comparison of a standard DES to a truncated DES model (holding all other parameters and random seeds fixed) can isolate the contribution of truncation error. The plot of model estimate convergence shows clearly that the variance of model outputs is considerably greater with the truncated DES model.

We compare the above reference case (“no testing”) scenario to an alternative scenario (“PGx”) in which all individuals who develop the condition receive a (\$200) genetic test as part of their initial diagnosis. Patients who test positive for the variant (20% population prevalence) are placed on a more expensive (\$3/day) alternative maintenance medication for the remainder of their life. This medication lowers the risk of the adverse event (relative risk 0.8). Individuals without the variant do not benefit from taking the pharmacogenomic alternative (relative risk 1.0), so remain on the standard (\$0.25/day) therapy.

We model a 40-year time horizon with background mortality based on age- and gender-based mortality as summarized using a Gompertz model fit to U.S. life-table data.¹⁸ We assumed that all events were not recurrent. All costs are provided in \$2019. We also assume a standard discount rate of 3%. Key model outputs included average discounted QALYs and costs, the incremental cost-effectiveness ratio (ICER) comparing the PGx vs. no testing strategies, and Net Monetary Benefit (NMB) based on a willingness-to-pay (WTP) of \$100,000 per QALY.

Simulation Models

We developed four models: (1) a differential equations [DEQ] model, which we used as the reference model for comparison with other models since it deterministically solved for average cost and QALY outcomes; (2) a cohort-based discrete-time state transition model [MARKOV]; (3) an individual discrete time state transition model [MICROSIM]; and (4) a discrete event simulation model [DES]. For our primary results the MARKOV and MICROSIM models utilized 10 million simulated patients. For the MARKOV and MICROSIM models we also produced results based on three cycle lengths: a one-year cycle (1Y), a monthly (1M) cycle, and (where feasible) a daily cycle (1D).

For MICROSIM and MARKOV, we modeled each model-cycle-length combination using two approaches: (1) by converting rates to probabilities using a standard conversion formula, i.e., $P(t) = 1 - e^{-rt}$, where r is the rate and t is the time cycle; and (2) by embedding transition probabilities by taking the matrix exponential of the transition intensity matrix (G) as defined by the rates as summarized above and in Table 2, i.e., $P(t) = e^{tG}$. We augmented the embedded models with non-Markovian accumulators to ensure that utility and cost estimates accounted for competing within-cycle events. We label results based on this embedded transition intensity matrix using -EMB in the model output.

Stochastic Convergence and Model Run Time

We investigated stochastic convergence of estimates of the NMB for the DES and MICROSIM with deterministic NMB estimates from the DEQ and MARKOV-based models. We did so by varying the number of Monte Carlo draws from one thousand to 1 billion patients, though we also considered simulation sizes up to 100 billion patients using a bootstrapping (with replacement) procedure.

We assessed model run time by re-estimating each model 100 times using different random seeds. From these results we measured key moments (mean, variance) in the run-time distribution. For this benchmarking procedure we simulated ten million patients for the DES

and MICROSIM models. To ensure fair comparison, runtime benchmarking was run on a high-performance computing cluster requesting 1 CPU and 120G memory. All models were run using the R statistical software program (version 4.0.2).

Results

Baseline Results

Based on the DEQ model, PGx testing yielded an ICER of \$103,212/QALY vs. a reference no-testing strategy (Table 3). A Markov model with an annual cycle duration and transition probabilities converted using $P(t) = 1 - e^{-rt}$ resulted in an ICER of \$91,929/QALY. Therefore, at a willingness-to-pay threshold of \$100,000/QALY, different decisions on the optimal treatment strategy would be made depending on the choice of a Markov vs. DEQ model. Only at successively shorter cycle durations (e.g., 1 month and 1 day with standard probability conversion) did Markov model results resemble those obtained using DEQ. For example, a Markov model with a daily time cycle yielded an ICER of \$103,173/QALY.

Markov models embedded by exponentiating the transition intensity matrix yielded ICERs comparable to the DEQ for all time cycle durations (Table 3). For example, with a yearly time cycle MARKOV-EMB produced an ICER of \$103,181 vs. \$103,212 using DEQ.

Among stochastic models with 10 million simulated patients, only DES yielded ICER and NMB results similar to the DEQ and MARKOV-EMB models. The DES produced an ICER of \$103,345 while the microsimulation models—including those that used embedded transition probability matrices—yielded ICERs ranging from -\$15,925 (i.e., PGx strategy dominated) to \$28,199.

The Appendix further investigates our primary results by constructing a cost-effectiveness acceptability frontier based on a probabilistic sensitivity analysis (PSA). This PSA is based on attaching uncertainty distributions to each model parameter, and re-estimating each model type on a K=1,000 sized PSA sample.

Stochastic Model Convergence

Among the Monte Carlo-based models considered, convergence with deterministic outcomes from the DEQ was achieved with considerably fewer simulated patients for DES as compared with MICROSIM (Figure 1). Only with short (1-day) cycle durations and/or embedding corrections and 1 billion or more simulated patients did the MICROSIM models consistently estimate similar outcomes as the DEQ or MARKOV-EMB models. By contrast, DES models with 1 million patients or more reliably produced average NMB outcomes similar to the DEQ and MARKOV-EMB models.

Computation Time

Computation time differed across approaches. Deterministic solutions based on DEQ (median runtime 3.6sec) and MARKOV with a yearly cycle (median=0.01sec using standard conversion formulas and 1.1sec for the embedded model based on the transition intensity matrix) had the fastest runtimes. By comparison, MARKOV with a daily cycle had a median

runtime of 8.7 minutes. Among Monte-Carlo-based models, the DES had a median run time of 48.2 seconds for 10 million patients, while MICROSIM with an annual cycle duration had a median run-time of 19.0 minutes. Due to computational constraints we did not test run-time for a MICROSIM with a daily cycle length.^v

Discussion

Modeling choices affect the scope, accuracy and reliability of decision outcomes in health economic assessments. By comparing outputs and optimal decisions across four commonly used modeling approaches, we show that care must be taken when structuring and executing models.

Given the predominance of discrete time cohort state transition (Markov) models in health technology assessments,³ one noteworthy observation from our study is that embedding transition matrices using commonly-used rate-to-probability conversion formulas can decouple a Markov model from the continuous time process it aims to represent. In essence, an improperly-embedded discrete time model will rule out events and event sequences from occurring within a time step that would, with some probability, occur in continuous time. While this is certainly not a new observation,^{6,26–29} our results demonstrate that improper embedding can meaningfully affect decision outcomes.

Our results also highlight a telltale sign of embedding problems: outcomes change as shorter time cycles are used. While it may be tempting to simply adopt a shorter time cycle to reduce the likelihood of two competing events occurring within the time step, it is worth emphasizing that a unique solution to time-step transformations requires that the transition matrix be expressible in terms of a generator matrix of transition intensities—precisely the matrix needed to properly embed the original transition matrix in the first place.³¹ In other words, simple transformations of rates to probabilities for a shorter time cycle does not guarantee a unique or even correct solution.

Our study was also the first to run a “horse-race” between Monte Carlo-based modeling approaches. In this case, DES demonstrated clear advantages. DES was both computationally more efficient and resulted in decision outcomes that were nearly identical (with a sufficient number of simulated patients) to DEQ and embedded Markov models. By contrast, microsimulation required substantially more simulated patients to yield reliable results and, when it did converge, resulted in slightly different average outcomes (though no difference in optimal decisions) even when using embedded transition matrices; we explore several reasons why in the Appendix. It is worth noting, however, that our application employed standard Monte-Carlo-based sampling methods. Depending on a study’s objectives, outcome variance for estimates of the incremental NMB for one strategy vs. another can be further reduced using additional variance reduction techniques.^{32,33}

^vTo provide context for how long a daily MICROSIM would take to complete, a cohort-based Markov model with a daily cycle had a median 522sec of runtime, as compared with 0.00697sec for the same model with an annual cycle length—nearly 75,000 times longer. Using this ratio to scale the yearly MICROSIM result (19.0 minutes) implies a model run time of 23,758 hours.

To conclude, it is worth offering guidance based on our study that can aid future modeling decisions for other researchers. For standard health economic assessments, Markov models offer good balance between speed and accuracy. However, if competing events or event sequences are clustered near each other (e.g., complications after a procedure or prescription), modelers should take great care to ensure that embedding biases do not materially affect decision outcomes. Proper embedding complicates model structure since non-Markovian accumulators must be included to capture the cost and utility changes from sojourns implied by “jumpover” states. However, a correctly embedded transition matrix also means that model results will not change if longer time cycles are used. This could be important in applications that aim to utilize computationally intensive PSAs and VOI methods to understand model sensitivity and guide future work. That is, rather than a (computationally demanding) monthly cycle, a yearly cycle could be used to conduct VOI analyses based on a properly embedded Markov model. Finally, our results make clear that if the modeling scenario calls for the use of a Monte-Carlo model, DES models avoid many of the pitfalls that complicate microsimulation and yield a favorable mix of accuracy, reliability, and speed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health economics* 2006;15(12):1295–1310. [PubMed: 16941543]
2. Caro JJ. Discretely integrated condition event (DICE) simulation for pharmacoeconomics. *Pharmacoeconomics* 2016;34(7):665–672. [PubMed: 26961779]
3. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health economics* 2003;12(10):837–848. [PubMed: 14508868]
4. Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic Microsimulation Models for Health Outcomes: A Review. *Med Decis Making* 2011;31(1):10–8. [PubMed: 20484091]
5. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment. *Pharmacoeconomics* 2008;26(2):131–148. [PubMed: 18198933]
6. Caro JJ, Möller J. Advantages and disadvantages of discrete-event simulation for health economic analyses. *Expert Review of Pharmacoeconomics & Outcomes Research* 2016;16(3):327–9. [PubMed: 26967022]
7. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes* Oxford university press; 2015.
8. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. *Cost-effectiveness in health and medicine* Oxford University Press; 2016.
9. Siebert U, Alagoz O, Bayoumi AM, et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–3. *Med Decis Making* 2012;32(5):690–700. [PubMed: 22990084]
10. Baltussen R, Marsh K, Thokala P, et al. Multicriteria Decision Analysis to Support HTA Agencies: Benefits, Limitations, and the Way Forward. *Value in Health [Internet]* 2019 [cited 2019 Oct 16];0(0). Available from: [https://www.valueinhealthjournal.com/article/S1098-3015\(19\)32358-7/abstract](https://www.valueinhealthjournal.com/article/S1098-3015(19)32358-7/abstract)
11. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost effectiveness and resource allocation* 2006;4(1):14. [PubMed: 16923181]

12. Peacock S, Mitton C, Bate A, McCoy B, Donaldson C. Overcoming barriers to priority setting using interdisciplinary methods. *Health Policy* 2009;92(2–3):124–32. [PubMed: 19346024]
13. Thokala P, Duenas A. Multiple criteria decision analysis for health technology assessment. *Value in Health* 2012;15(8):1172–1181. [PubMed: 23244821]
14. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–4. *Medical decision making* 2012;32(5):701–711. [PubMed: 22990085]
15. Standfield L, Comans T, Scuffham P. Markov modeling and discrete event simulation in health care: a systematic comparison. *International journal of technology assessment in health care* 2014;30(2):165–172. [PubMed: 24774101]
16. Jacobson SH, Hall SN, Swisher JR. Discrete-event simulation of health care systems. In: *Patient flow: Reducing delay in healthcare delivery* Springer; 2006. p. 211–252.
17. Graves JA, Garbett S, Zhou Z, Peterson J. *The Value of Pharmacogenomic Information* 2018;
18. Briggs A *Decision Modelling for Health Economic Evaluation* 1 edition. Oxford: Oxford University Press, USA; 2006.
19. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. *Medical decision making* 2012;32(5):722–732. [PubMed: 22990087]
20. Campbell JD, McQueen RB, Libby AM, Spackman DE, Carlson JJ, Briggs A. Cost-effectiveness uncertainty analysis methods: a comparison of one-way sensitivity, analysis of covariance, and expected value of partial perfect information. *Medical Decision Making* 2015;35(5):596–607. [PubMed: 25349188]
21. Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research. *Pharmacoeconomics* 2006;24(11):1055–1068. [PubMed: 17067191]
22. Jalal H, Goldhaber-Fiebert JD, Kuntz KM. Computing expected value of partial sample information from probabilistic sensitivity analysis using linear regression metamodeling. *Medical Decision Making* 2015;35(5):584–595. [PubMed: 25840900]
23. Jalal H, Alarid-Escudero F. A Gaussian Approximation Approach for Value of Information Analysis. *Medical Decision Making* 2017;0272989X1771562.
24. Barendregt JJ. The half-cycle correction: banish rather than explain it. *Medical Decision Making* 2009;29(4):500–502. [PubMed: 19571330]
25. Jahn B, Kurzhthaler C, Chhatwal J, et al. Alternative Conversion Methods for Transition Probabilities in State-Transition Models: Validity and Impact on Comparative Effectiveness and Cost-Effectiveness. *Medical Decision Making* 2019;0272989X19851095.
26. Jones E, Epstein D, García-Mochón L. A procedure for deriving formulas to convert transition rates to probabilities for multistate Markov models. *Medical Decision Making* 2017;37(7):779–789. [PubMed: 28379779]
27. O'Mahony JF, Newall AT, van Rosmalen J. Dealing with Time in Health Economic Evaluation: Methodological Issues and Recommendations for Practice. *Pharmacoeconomics* 2015;33(12):1255–68. [PubMed: 26105525]
28. Welton NJ, Ades AE. Estimation of Markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. *Medical Decision Making* 2005;25(6):633–645. [PubMed: 16282214]
29. Iskandar R A theoretical foundation for state-transition cohort models in health decision analysis. *PLOS ONE* 2018;13(12):e0205543. [PubMed: 30533043]
30. Chhatwal J, Jayasuriya S, Elbasha EH. Changing cycle lengths in state-transition models: challenges and solutions. *Medical Decision Making* 2016;36(8):952–964. [PubMed: 27369084]
31. Tricker AR. Effects of rounding on the moments of a probability distribution. *Journal of the Royal Statistical Society: Series D (The Statistician)* 1984;33(4):381–390.
32. Shechter SM, Schaefer AJ, Braithwaite RS, Roberts MS. Increasing the Efficiency of Monte Carlo Cohort Simulations with Variance Reduction Techniques. *Med Decis Making* 2006;26(5):550–3. [PubMed: 16997930]

33. Stout NK, Goldie SJ. Keeping the noise down: common random numbers for disease simulation modeling. *Health Care Manage Sci* 2008;11(4):399–406.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

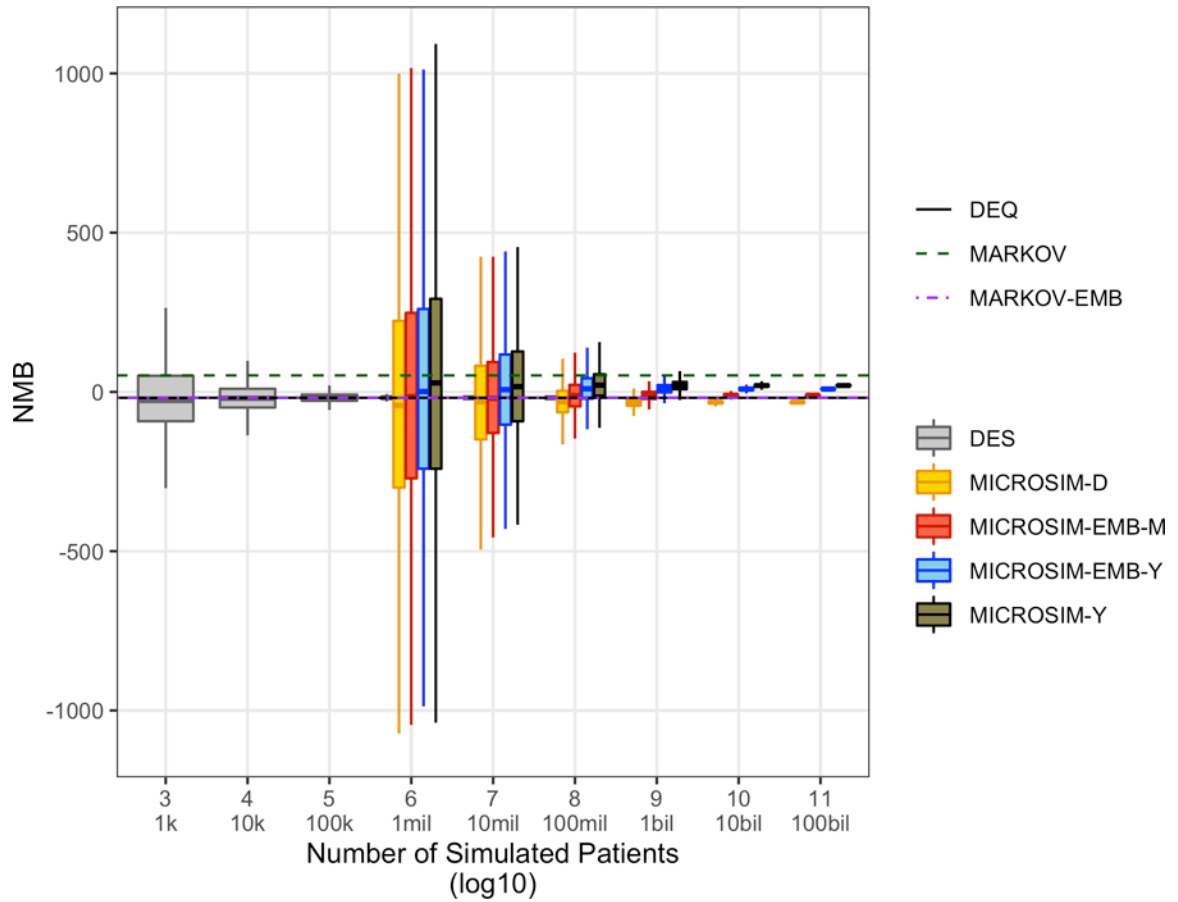


Figure 1.

Convergence of net monetary benefit outcome for stochastic models to deterministic model outputs. Embedded microsimulation with monthly and daily time cycles not simulated due to computational constraints (i.e., for some simulated patient sizes, model runtimes of > 1 year in high-performance computing environment). DEQ, differential equations model; DES, discrete event simulation; MARKOV, discrete-time (1-year cycle) Markov model with rate-to-probability adjustment based on $P(t) = 1 - e^{-rt}$, where r is the rate and t is the time step; MARKOV-EMB, embedded discrete-time (1-year cycle) Markov model with probability matrix based on $P(t) = e^{tG}$, where t is the time step and G is a transition intensity matrix; MICROSIM-[D/M/Y], discrete-time (daily/monthly/yearly time cycle) microsimulation model with rate-to-probability adjustment based on $P(t) = 1 - e^{-rt}$, where r is the rate and t is the time step; MICROSIM-EMB-[M/Y], embedded discrete-time [monthly/yearly time cycle] microsimulation model with probability matrix based on $P(t) = e^{tG}$, where t is the time step and G is a transition intensity matrix; NMB, net monetary benefit (willingness to pay \$100,000 per quality-adjusted life year).

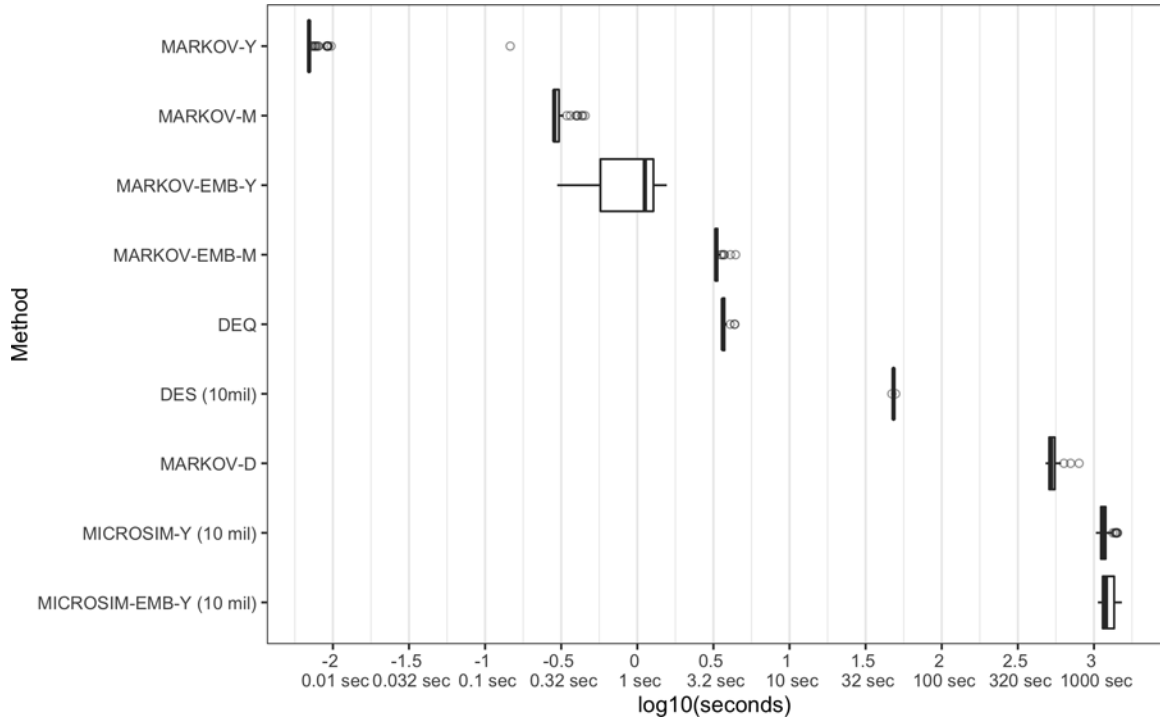


Figure 2. Model runtime comparison. Each whisker plot shows the distribution of runtime based on 100 model runs using the same random seed (for stochastic models). For Monte Carlo-based models, number of simulated patients is provided in parentheses. DEQ, differential equations model; DES, discrete event simulation; MARKOV, discrete-time (1-year cycle) Markov model with rate-to-probability adjustment based on $P(t) = 1 - e^{-rt}$, where r is the rate and t is the time step; MARKOV-EMB, embedded discrete-time (1-year cycle) Markov model with probability matrix based on $P(t) = e^{tG}$, where t is the time step and G is a transition intensity matrix; MICROSIM-[D/M/Y], discrete-time (daily/monthly/yearly time cycle) microsimulation model with rate-to-probability adjustment based on $P(t) = 1 - e^{-rt}$, where r is the rate and t is the time step; MICROSIM-EMB-M/Y), embedded discrete-time [monthly/yearly time cycle) microsimulation model with probability matrix based on $P(t) = e^{tG}$, where t is the time step and G is a transition intensity matrix; NMB, net monetary benefit (willingness to pay \$100,000 per qualityadjusted life year).

Table 1

Typology of Errors

	Brief Description	Deterministic Models			
		Differential Equations [DEQ]	Markov Cohort [MRKCHRT]	Individual Microsimulation [MICROSIM]	Discrete Event Simulation [DES]
Errors in Expectation	Errors that result in bias.				
Structural	Attributes of models and model parameters (e.g., time-dependency of transition rates/probabilities, the number of overall states in the model) result in model estimates that deviate from the "true" underlying event generation process.	X	X	X	X
Integration	Model accumulates events at time cycle boundaries (either at the beginning or the end) rather than at any point in continuous time.		X	X	
Embedding	Transition probabilities do not align with an underlying continuous time process.		X	X	
Errors of Reliability	Errors that can be addressed via additional Monte Carlo simulations or through further population sampling or research.				
Estimation and Sampling	Model parameters are estimated and/or sampled with uncertainty in the population of interest.	X	X	X	X
Stochastic	The realization and timing of events among otherwise identical patients varies.			X	X
Truncation /	Information on the precise timing of an event occurring in continuous time is lost when events times are "rounded up" or "rounded down" to the boundary point in the cycle.			X	

Notes:

¹ Truncation error can also result in a small amount of bias (see Tricker 1984)

Table 2**Model Parameters**

Description	Parameter Name	Parameter Value
Discount rate	disc	0.03
DES simulations to perform (1000s)	n	1000
Diff Eq Time step for DEQ approach	resolution	0.019
Time horizon (years) of simulation	horizon	40
Willingness to pay threshold used for NMB (\$1000s / QALY)	wtp	100
Shape Parameter from Gompertz model of secular death for 40yr female fit using 2012 U.S. Social Security mortality data.	shape	0.101
Rate Parameter from Gompertz model of secular death for 40yr female fit using 2012 U.S. Social Security mortality data.	rate	0.001
Population prevalence of genetic variant.	p_g	0.2
Condition indication percentage	r_a_pct	10
Condition indication time period (years)	r_a_dur	10
Probability of ordering the genetic test.	p_o	1
Probability of death from adverse drug event.	p_bd	0.1
Adverse drug event percentage	r_b_pct	25
Adverse drug event time period (years)	r_b_dur	1
Relative risk of adverse drug event among patients with genetic variant present and alternative therapy prescribed.	rr_b	0.8
Initial cost at indication (\$1000s).	c_a	10
Cost of standard drug therapy.	c_tx	0.25
Cost of pharmacogenomic alternative drug therapy.	c_alt	3
Cost of adverse drug event survival (\$1000s).	c_bs	25
Cost of adverse drug event case fatality (\$1000s).	c_bd	20
Cost of genetic test.	c_t	200
Disutility of developing condition.	d_a	0.05
Disutility duration (years) after developing condition.	d_at	1
Lifetime disutility of adverse drug event among survivors.	d_b	0.15

DES = Discrete Event Simulation

NMB = Net Monetary Benefit

QALY = Quality-adjusted life years

Table 3

Outcomes by Model Type

	ICER: PGx Testing vs. No Testing	NMB: PGx Testing vs. No Testing \$100k/QALY
Deterministic Models		
Differential Equations	103,212	-18.4
Markov Cohort - Basic Rate-to-Probability Adjustment ¹		
Daily Cycle	103,173	-18.2
Monthly Cycle	102,175	-12.6
Yearly Cycle	91,929	51.9
Markov Cohort - Embedded ²		
Daily Cycle	103,206	-18.3
Monthly Cycle	103,158	-18.1
Yearly Cycle	103,181	-18.2
Stochastic Models (M = 10 million simulated patients)		
Discrete Event Simulation	103,345	-19.1
Microsimulation - Yearly Cycle	(15,925)	-84.0
Microsimulation - Daily Cycle	2,286	25.0
Embedded Microsimulation - Yearly Cycle	(4,880)	-296.0
Embedded Microsimulation - Monthly Cycle	28199	-158

Notes:

¹ Basic rate-to-probability adjustment based on $P(t) = 1 - e^{-rt}$, where r is the rate and t is the time step.² Embedded probability matrix based on $P(t) = e^{tG}$, where t is the time step and G is a transition intensity matrix.

ICER = Incremental Cost-Effectiveness Ratio

NMB = Net Monetary Benefit

Negative ICERs indicate the PGx strategy is dominated but are shown to demonstrate magnitude of differences across modeling approaches.

ICER and NMB values will not match those constructed manually due to rounding in the reported average costs and QALY estimates.