A Taxonomy of Model Structures
for Economic Evaluation of Health Technologies

Authors (alphabetical):

Alan Brennan

Director of Health Economics and Decision Science
http://www.shef.ac.uk/scharr/sections/heds
School of Health and Related Research (ScHARR)
University of Sheffield
Regent Ct
30 Regent St
Sheffield S1 4DA
Tel:+44 (0)114 2220684
Fax:+44 (0)114 2724095
a.brennan@sheffield.ac.uk

Stephen E. Chick

Associate Professor of Technology Management
Member, Health Management Institute
INSEAD
Technology and Operations Management Area
Boulevard de Constance
77300 Fontainebleau France
stephen.chick@insead.edu
http://faculty.insead.edu/chick
Tel: (33) 1.60.72.41.57
Fax: (33) 1.60.74.61.79

Ruth Davies

Professor of Operational Research and Systems
Warwick Business School
University of Warwick
Coventry CV4 7AL
http://users.wbs.warwick.ac.uk/group/ors/people/academic/daviesr/contact
Tel: 44 (0)24 7652 2475
Fax: +44 (0)24 7652 4539

Model Structure Selection for Health Economic Evaluations:

Taxonomy of Options and Criteria for their Use

Summary

Models for the economic evaluation of health technologies provide valuable information to decision makers. The choice of model structure is rarely discussed in published studies and can affect the results produced. Many papers describe good modelling practice, but few describe how to choose from the many types of available models. This paper develops a new taxonomy of model structures. The horizontal axis of the taxonomy describes assumptions about the role of expected values, randomness, the heterogeneity of entities, and the degree of non-Markovian structure. Commonly used aggregate models, including decision trees and Markov models require large population numbers, homogeneous sub-groups and linear interactions. Individual models are more flexible, but may require replications with different random numbers to estimate expected values. The vertical axis describes potential interactions between the individual actors, as well as how the interactions occur through time. Models using interactions, such as system dynamics, some Markov models, and discrete event simulation are fairly uncommon in the health economics but are necessary for modelling infectious diseases and systems with constrained resources. The paper provides guidance for choosing a model, based on key requirements, including output requirements, the population size, and system complexity.

Introduction

The decision about which model structure should be chosen in a particular health economic evaluation context is only rarely discussed in published studies. Policy recommendations based on a model may depend upon the explicit and implicit assumptions of the model. Many evaluations use aggregate or 'cohort' models, which examine the proportions of the population undergoing different events with associated costs and utilities. Patient or individual level models, which sample individuals with specific attributes and follow their progress over time, have recently become more prevalent. The problem is that aggregate models adopt assumptions that may unwittingly produce inaccurate or inadequate solutions, whilst individual level models can adopt less stringent assumptions, but may be more time consuming to develop and run. Recent debates on choice of model structure have focussed on this cohort versus individual level dichotomy [1]. This paper presents a taxonomy of a wide range of available model structures, and guidance for the selection of an appropriate model.

We define a model based evaluation as a formal quantified comparison of health technologies, synthesising sources of evidence on costs and benefits, in order to identify the best option for decision makers to adopt. Interventions can include drugs, surgical procedures, and psychological therapies or much broader health system strategies such as screening policies, public health interventions or bio-terrorism defence. Outputs usually include intervention costs and a measure of morbidity and mortality such as quality of adjusted life years (QALYs). Adoption may depend upon a single measure e.g. the expected incremental cost per QALY between technologies against a threshold willingness to pay ($?$) or equivalently, expected monetary net benefit (NB=$?$*QALY – Cost). Sometimes, decision makers are also interested in

other measures around uncertainty, individual or geographic variability and trends over time.   Recent guidance recommending probabilistic sensitivity analysis [2, 3] leads to a preference for fast, efficient models with a single output measure.

Model structure is usually determined by considering the relationship between the inputs (natural history of disease, clinical pathways, evidence of interventions' effectiveness, utilities associated with health states, intervention and other costs etc.), and the output measures required by the decision maker. Practical considerations also include availability of data, the background and skill of the researcher and the type of software available. As the choice is usually made relatively early in projects, it is often hard for model developers to abandon an initial structure and start again.

There are several papers whose guidance on good modelling practice distils down to a series of principles (e.g. transparent structure, appropriate and systematic use of evidence) rather than more detailed guidance on the structure that is appropriate [4, 5, 6,7]. Sculpher et al. (2000) are slightly more detailed, recommending that model structure be as simple as possible, consistent with the stated decision problem and a theory of disease and not defined by data availability or health service inputs alone. They do not address the issue of which specific modelling techniques to use, but implicitly assume, in making their recommendations, the use of a cohort model [8]. Some researchers have directly compared alternative model structures. Karnon's economic evaluation of breast cancer showed that an individual level discrete event simulation and a cohort Markov model could be 'tuned' to produce the same results [9]. Koopman et al. have experimented with a hierarchy of models, designed to answer similar epidemiological questions [10]. Barton et al, reviewing methodology

in decision trees, cohort Markov models, and individual level models [11], suggest that the choice between these three approaches depends upon: whether pathways could be adequately represented by probability trees, whether a Markov model would require an excessive number of states, and whether interaction between patients is important.

This paper extends those articles by more systematically establishing the range of model structure options and criteria for their selection.

A Taxonomy of Model Structures

Table 1 shows the range of available approaches for health economic evaluation models, and their relationships to each other. This section describes the taxonomy from a conceptual perspective. The Appendix describes a variety of computational tools for implementing the models.

Table 1: Taxonomy of Model Structures

The rows (1 to 4) describe factors involving both time and interaction between individuals. Health economists' current approaches are largely those in the top half of the table. The models assume independence between individuals, and time may (row 2) or may not (row 1) be modelled explicitly. Models with interactions (rows 3, 4) are necessary when individuals interact (e.g., infectious disease transmission) or constraints affect individuals (e.g., finite service capacity or restricted supplies of organs for transplantation). We distinguish discrete time and continuous time models.

The columns (A to D) separate cohort from individual level models and disentangle assumptions concerning expected values, randomness, and the heterogeneity of entities. Cohort models (columns A, B) quantify the proportion of people with common characteristics. In cohort models with randomness (column B), the Markovian property is typically assumed, meaning that the future is conditionally independent of the past, given the present. Cohort models can account for different attributes/covariates (e.g. multiple ages, weights, genders, other risk factors, stages of natural history of disease etc.) by subdividing the number of states or branches, but the number of dimensions rises exponentially (e.g. M binary attributes imply 2^M dimensions). Individual level models (columns C, D) overcome this problem by simulating the progression of each individual with different characteristics (a population of N individual patients with M attributes each requires MxN data entries). Non-Markovian distributions (column D) allow greater flexibility in modelling the timing of health-related events. Stochastic models may require many simulation runs to quantify the output mean and variance with sufficient accuracy. More complex individual level models (e.g. D4) can examine interactions both with other individuals and with the environment, including the availability of resources (e.g., doctors, beds, transplant organs). Each grid cell in the taxonomy is related to its neighbours by varying some of the basic assumptions that underlie each model.

Aggregate Models without Interaction

Decision Trees for Cohorts (A1, B1)

Decision trees (A1), as typically used in health economics, are among the most widely used aggregate level models [11,12]. A decision tree outlines decisions (the square in Figure 1), the probability or fraction of various outcomes (emanating from the circles), and the valuation of each outcome, such as a QALY, cost or net benefit measure. The mean value of a decision is computed analytically ("rollback") by summing the probability of each outcome with its value, such as E[NB | Treat] = 0.9 * 0.9 + 0.1 * 0.5 = 0.86. This reflects either a large-population assumption, so that probabilities essentially match the actual fractions of individuals with a given attribute (a law of large numbers argument), or an assumption that expected values for patients are the desired outcome (rather than a distribution or variation, assuming a risk-neutral decision maker), or both. Recursion, or looping, is not allowed, but much more complicated scenarios are possible than in Figure 1, with many decision strategies, long sequences and multiple outcomes from chance nodes.

Figure 1:A Decision Tree

The simulated decision tree cohort (B1) provides an alternative approach to estimating the mean value of each decision option. In Figure 1, E[Treat] might be estimated by Monte Carlo sampling a one million (10^6) patient cohort with a Binomial (10^6, 0.9) distribution. This simulates the number of individuals on each path, but not each individual separately. The advantage of this approach is that it can provide a measure of the variability of the number of individuals likely to be in each state.

Markov Models for Cohorts (A2, B2)

Markov models (A2) can provide a more compact representation than the decision tree when a repeated set of outcomes is possible through time A cohort-based Markov transition matrix uses a transition probability per unit time for individuals in the cohort to change to another state, with associated costs and utilities [11,13,14] (Figure 2).

Figure 2: Transition Probability Matrix for a Markov Model

This formulation assumes that the transition probabilities are constant over time. Survival in a particular state is therefore defined by a geometric distribution in discrete time. Analysts often use subtle ways to extend the generaliseability of the Markovian assumption, for example, by using different transition matrices as time progresses. If transition probabilities depend on certain attributes, then one can redefine states. For example, patient history can influence risk by defining states to include healthy with previous history of illness, healthy with no previous history of illness, etc.

It is a common misconception that analysts should use a time step defined by data availability (e.g. a year). The time-step used, however, affects the model results. With discrete time models, electing a smaller and smaller time slice is warranted until the output is no longer affected. To determine the probabilities for different time intervals for a two-state transition function, one can use an equation that relates the probability p_{ij} of an outcome, j, from state i through time t via the instantaneous hazard rate r_{ij}, namely $p_{ij} = 1 - e^{-r_{ij}t}$ [13]. Alternatively, the transition probability can

be modelled by $p_{ij} = r_{ij}t$ (both have the same first-order linear term in their Taylor series expansion).

Markov models for cohorts can be processed either analytically (A2) with expected values or with simulated random Markov model transitions (B2). The approach and rationale of Monte Carlo simulation is like that for the Simulated Decision Tree (B1).

Individual Sampling Models without Interaction (CD1, CD2)

Rather than tracking data for every compartment or path, individual sampling models (ISMs) track specific individuals. The individuals have potentially heterogeneous characteristics that affect their progression through the model, yet individuals progress through the model independently of each other and of environmental constraints. Typically, ISMs generate a large numbers of simulated patient histories and evaluate results with a sampling algorithm (Figure 3).

Figure 3 An Individual Sampling Model Algorithm

Simulated Patient Level Decision Tree models (CD1) can be used to simulate individuals through a tree's pathways, maintaining a record of the 'history' as nodes are processed. Time elapses implicitly. This approach is most useful when probabilities, QALYs, and costs depend upon a large or conceptually infinite number of values, (e.g. normally distributed costs that depend upon patient glucose levels).

The most common implementations of ISMs in health economics are Markov models that are modified to simulate individuals through each time period (CD2). One key advantage lies in modelling multiple co-morbidities which depend on multiple attributes / covariates. Examples include models of diabetes where patient co-morbidities interact and affect the outcome [15,16,17], and other work on rheumatoid arthritis [18] and osteoporosis [19]. The ISM approach can be further modified to simulate the "time to next event" rather using equal time periods. This can provide richer modelling power, computational efficiency and greater elegance in coding. For example, the Birmingham Rheumatoid Arthritis Model [20,21] considers long-term treatment for rheumatoid arthritis with strategies defined by alternative sequences of disease-modifying anti-rheumatic drugs.

Aggregate Models with Interaction Allowed

System Dynamic Models (A3, A4)

If individuals do interact, it is not always necessary to model each individual separately. System dynamics, a common cohort-based approach in epidemiology and operational research, models the state of the system in terms of changing, continuous variables over time. Crucially, it enables the rate of change in the system to be a function of the system's state itself (i.e., feedback). Typical examples of feedback include infectious disease outcomes, where higher levels of infection produce higher risks of further infection but also reduce the number of people in the susceptible pool, and health care service constraints, where the system performs differently when it is

full or over-capacity [22,23,24]. Factors affecting rates of change may include

behavioural influences.

Continuous time system dynamics models (A4) use ordinary differential equations

(ODE), like $dx/dt = f(x, t)$ to describe the rate of change of a state vector x. For

example, Equation 1 describes the rate of change in the number of infected

individuals i (the "state") in a population of N individuals as a function of the contact

rate (c contacts per unit time), the probability $\boldsymbol{b}$ of infection per potentially infectious

contact, the probability i/N that one of the $(N\text{-}i)$ susceptibles contacts an infected, and

the mean duration of infection D.

Equation 1 An Infection Systems Dynamics Model expressed as Ordinary Differential Equation

$$
\begin{aligned}
\text{(Equation 1)} \qquad \frac{di}{dt} \;&=\; \text{population infection rate} - \text{population recovery rate} \\[2mm]
&=\; c\boldsymbol{b}\,\frac{i}{N}(N-i) - \frac{i}{D}
\end{aligned}
$$

System dynamics models can have multiple linked ODEs to specify the full model.

The Euler-forward approximation to the continuous time ODE model results in a

discrete time finite difference equation (FDE) model (A3) approximation

$$x_{j+1} = x_j + f(x_j, t_j)\,\boldsymbol{D}t$$

where x_j is the state of the system at time $t_j = j\boldsymbol{D}t$. [25]. The facilities for interaction

are the same as for the continuous model. The solution to the differential or difference

equations is specified over time so that the output resulting from a different strategy

can be quantified. The approach therefore provides the ability to model the

deterministic interaction of groups of individuals but without modelling each

individual in the system.

System dynamics models are commonly implemented in packages with graphical

interfaces, which make them easier to build and use. There have been applications

concerned with community care, short-term psychiatric care, HIV/ AIDS [23],

smallpox preparedness [26,27,28], CJD [29] and others.

The cohort level Markov models (A2) that are evaluated using expected values are in

fact special cases of discrete time finite difference (FDE) models (A3), but without

the ability to model interactions (e.g., nonlinear dynamics from interpersonal infection

transmission or resource capacity constraints). More accurate FDE approximations to

ODEs are available [30], and partial differential equations [24] can complement ODEs

to enhance model richness (e.g., geographic effects).

Discrete State, Continuous Time Markov Models (B3, B4)

The system dynamics approach has two important assumptions – firstly, that the

change dynamics are deterministic and secondly, that fractions of individuals can

occur in specific health states (an infinitely divisible population assumption). With

large populations such assumptions may be reasonable. Continuous time Markov

chain (CTMC, grid cell B4) can address both a need to model an integer number of

individuals in each health state and the variability that can occur in such systems a. A

CTMC can model many interactions including, for example, infectious disease

dynamics and limited healthcare resources. Monte Carlo simulation is often

employed to analyze these systems, although analytical stochastic process methods

may sometimes be employed for sufficiently simple models.

A CTMC that is analogous to the system dynamics model in Equation 1 is depicted in Figure 4. Each circle is a potential state of the system (i people infected in a population of N individuals), the arrows out of a particular state represent the possible state transitions (from i people infected, a recovery changes the state to i-1 people infected, an infection changes the state to i+1 infected). The rates may depend upon the current state, so that the force of infection when there are i people infected is $\lambda_i = c\beta i(N-i)/N$, and the recovery rate is $\mu_i = i/D$.

Figure 4: Continuous-time Population Markov Model Transition Probability

The distribution of the time to the next event (infection or recovery) is exponentially distributed with mean that is one divided by the sum of the rates out of the current state ($1/(\lambda_i + \mu_i)$). The probability of a particular event is proportional to the appropriate transition rate (an infection event has probability $\lambda_i/(\lambda_i + \mu_i)$). A CTMC in general can be simulated by sampling an exponentially distributed time to next event, then sampling the specific event type using probabilities that are proportional to the individual event rates. After sampling the specific event type and updating any statistics, the rates may change for sampling the time to the next event. The basic theory to support this sampling mechanism is Theorem 1.

Theorem 1: Sampling Time to First of Many Events

Proof: See, e.g. Çinlar [31], Also see Banks *et al.*p. 222 [32].

In a context where k different events could happen (e.g., recoveries or potential

infections of many classes of ill people), Theorem 1 says that the time to the next

event can be sampled first (exponential, rate $\sum_{m=1}^{k} g_m$), and then the type of event to

occur can be sampled next (event j with probability $g_j \Big/ \sum_{m=1}^{k} g_m$).

CTMC models are often analysed stochastically by simulating realisations, or sample

paths, that describe how the state changes through time. The number of individuals in

each state is tracked, and each individual event must be processed, but there is no

need to sample the identities of individual patients to maintaining their histories.

Multiple independent realisations can be used to estimate statistics like time-varying

means, time averages, and the probability of extreme events like outbreaks. If the

number of individuals in the population gets very large, and a scaling is done so that

the fraction of individuals in each state is maintained, then plots of sample paths get

less variable and the number infected converge to a differential equation (e.g., see

[33,34] for epidemic models, and [35] for related results for queueing models). Figure 5

illustrates this by comparing the ODE infection model in Equation 1 with the

stochastic analogue in Figure 4. The approximation is poorer for small populations in

general. The upper left of Figure 5, where the possibility of infection being

eliminated in a short time is high, gives a specific example of a poor approximation.

Figure 5: Scaled realizations of ODE infection model Equation (1) and stochastic

model in Figure 4 (increasing population size N; c=10; b=0.1; D=1.2 days).

The discrete time Markov chain (DTMC) model (B3) is like the CTMC except that it is updated with finite time steps, at times *0, Dt, $2Dt$, $3Dt$*, … for some suitably chosen period *Dt*. Brookhart et al. (2002) used a DTMC to model a Cryptosporidiosis outbreak due to water treatment failure [36]. The larger the period *Dt* of the DTMC, the worse a DTMC approximation of the original CTMC process becomes, in part because individuals are allowed to have only one event affect them per time step. The error in a DTMC approximation to a CTMC can practically be controlled by shrinking the time step size and rescaling parameters appropriately) as described above for simulated Markov models (B2).

Individual Level Models with Interactions

To model interactions, analysts often intuitively think of modelling individuals as separate entities.

Individual level Markovian models with interactions (C3, C4)

Individual level Markovian models with interactions can be thought of as an extension of Markovian cohort models with interaction (B3, B4). Individual-level means that patient characteristics may be heterogeneous, and that histories may be tracked for each individual in the population. For that reason, this falls into the class of models that Koopman *et al.* calls individual event history (IEH) models.

There are at least two methods to simulate CT IEH models (C4). The first explicitly uses the Markovian assumption in Theorem 1 to simulate the time to the next event,

followed by the selection of which event occurs. This differs from the use of the

Theorem for Markovian cohort models (B4) in that a rate g_{jl} is assigned for any

event j that can occur for each individual l (rather than for each cohort). For the

epidemic model in Figure 4, if event 1 is identified with infection, then the infection

rate of susceptible individual l when there are i infected individuals is

$g_{1l} = cb(N - i)/N$. More generally, the parameters and rates may differ for each

individual to reflect heterogeneous population characteristics, and rates may also

depend upon resource constraints (e.g., the service completion rate is nonzero for at

most one patient if there is only one server). A second method is to use discrete-event

simulation (see below) with exponential distributions.

The analogous discrete time IEH model (C3) steps forward in discrete time intervals

in the same way as the DTMC model. Multiple individuals may be sampled at each

time step, but each individual experiences at most one state transition per time step.

Discrete Event Simulation (D4, D3)

Probably the most flexible of all modelling techniques, continuous time discrete event

simulation (CT, DES) (D4) describes the progress of individuals (entities), which

undergo various processes (events) that affect their characteristics and outcomes

(attributes) over time. Law and Kelton [37] present an excellent introduction. DES

permits sampling from any distribution, Markovian or not (C3). A queue structure

enables interaction to take place with constraints and between entities. The state of

the modelled system includes the current entities, their attributes, and a list of events

that can occur either at the current simulation time or that are scheduled to occur in

the future. Events change the state at discrete times, and each event is processed one by one, in a non-decreasing time order. As each event is processed, it may produce consequent additional events to be processed at the current simulated time, or at some sampled time in the future.

The nomenclature DES has been applied differently in different papers in health economics [11], epidemiology, [38,39,40] and health care delivery system design or scheduling [41]. We define DES to be consistent with the operational research literature that developed DES theory and applications for the last 50 years [37,42,43,44], and for which rigorous mathematical formalisms exist [45,46]. In a healthcare DES, individual interacting patients are usually the entities. Additional entities may include health care service system resources, such as doctors, nurses, and ambulances for transport. Examples of disease models include models of the activities and non-Markovian survival distributions related to the natural history and treatment of coronary heart disease [47], of vertical transmission of HIV [48], and of vaccine trial design [38]. While constrained resources pose no problems for most DES tools, special care may be required to model infection dynamics or multiple correlated health risks. For example, models of multiple diseases may require multiple events on the future event list per patient, as well as a need to reschedule events if there are related underlying causes. POST is a tool to facilitate such event management [49]. Custom-built programs are sometimes produced to adapt to the special needs of DES models of infection [38, 50,51].

All simulation models have sampling error for estimating mean output values (costs, QALYs, prevalence, etc.) that running multiple simulations can reduce. DES has the

additional disadvantage of potential computational slowness associated with inserting and removing events from the future event list.

DES models have discrete-time analogues (D3) that are not different in any significant methodological way to the continuous time models if the discrete time steps are small. In fact, most commercial DES packages have animation graphics that show the entities moving through the system to aid model development and checking, and these typically implement a discrete-time DES variation if and when graphics are running.

How to Decide Which Model Structure To Use

The taxonomy summarises the key structural assumptions that underlie each type of model. Table 2 gives a numbered checklist of additional key issues (I1 to I15), which can help to determine the exact approach required. The checklist follows the taxonomy structure, moving through the assumptions from the top left (the aggregate decision tree with homogeneous states, no recycling, no interaction, and a large enough population to make an integer approximation valid) across to the right, and down to the circumstances requiring a DES approach. The features of these separate issues can interact in the context of particular case studies.

Table 2 Issues and Guidance on Choice of Model Structure

Decision Makers' Requirements

The purpose of health economic evaluations is to inform decision-making. If the decision maker's perspective requires knowledge of the variability in the results rather than just mean outputs (I1) then stochastic rather than deterministic models may be required. Such requirements might include knowledge of the range of impact achievable in different geographical or organisational contexts or at different times. For example, Davies et al. (2002) show that, in screening for and eradicating Helicobacter pylori, the number of lives saved from death from gastric cancer may vary widely, meaning that some Health Authorities may incur large costs with little benefit [52]. It can be argued whether such variability should be part of the decision makers' considerations, but if it is, then the appropriate model structure is required.

Decision makers such as NICE now expect probabilistic sensitivity analysis (PSA) to quantify the uncertainty in mean outputs due to parameter uncertainty (I3). A deterministic model will provide the mean output for a given parameter set directly. In individual sampling models, the mean output for a given parameter set is estimated from (sometimes hundreds or thousands of) sampled individuals and the PSA will thus be relatively time consuming. The need for PSA should not directly drive model structure decisions but will influence them. There are interesting developments to facilitate PSA including Gaussian meta-models of individual level models [53,54] and quantifying the efficient number of individuals needed in each sample [55].

The choice of states and risk factors and the identification of their relationship to each other should normally precede the choice of model structure. There are occasions, however, when the decision makers require a very quick answer to a problem and there is a temptation to provide a model with a simple cohort structure without testing

the underlying assumptions. At the other extreme, some decision makers require a model that can be adapted over time to examine different populations and include emerging evidence on new risk factors or treatments. In these circumstances, the analyst should consider the use of an individual sampling model and, in particular, discrete event simulation which is more flexible [I4].

Sub-division into states

In all models the states represent the natural history of disease, treatments and their effects. In cohort models the states are often subdivided according to risk factors and may be further subdivided into short time periods to approximate non-Markovian distributions [I6]. The representation of risk factors in a cohort model leads to the danger of an explosion in the number of states, discussed previously [I7]. It is clearly desirable to keep the number of states as small as possible [8]. Risk factors may be combined if they have little impact on the results of interest to the decision maker.

The states in a cohort model are assumed to represent a homogenous population and the implicit use of averages by deterministic models may cause statistical bias in estimates of the mean outputs. For example, if the heart attack risk (per unit time) increases with age, then the average of the heart attack rates in a population is higher than the heart attack rate at the average age. The larger the age groups, the worse this problem will be. This is a consequence of Jensen's inequality, which states that $E[f(X)] \geq f(E[X])$ when f() is a convex function (heart attack risk) and X is a random variable (a covariate like age)[31,56]. More generally, if risk factors are not linearly related to the covariates, there may be some bias in QALY and cost measures. Individual level modelling can cope with a non-linear interaction between the risk and

outcome [I5], for example the relationship between the level of cholesterol and the risk of coronary heart disease [57].

Resource constraints

Another example of the inappropriate use of expected values is in the presence of variability and resource constraints [I14]. Deterministic models that use average service times and mean demand rates can under-estimate or completely miss queuing delay effects. If the event times are non-Markovian, then the queuing effects may differ even more (Pollacek-Khintchine formula) [56]. If health outcomes or delivery costs are linked to delays before intervention, stochastic models are required. Where there are explicit resource constraints and non-Markovian time to event distributions [I15], discrete event simulation may be the best choice. Non-Markovian distributions are recommended as the most natural way to model process improvements in health care delivery, such as reductions in the variability in patient arrival and service delivery times [58].

Model run time and population size

Aggregate models sometimes need thousands or millions of states, possibly to cope both with the number of co-variates ("the curse of dimensionality") and to 'fix' non-Markovian effects. The resulting models may lack transparency and be slow to run. Individual sampling models can solve these problems, but on the other hand, need time consuming multiple replications to get good estimates of means. Discrete event simulation may be particularly slow because of the need to maintain a future events

list. Variance reduction methods are available to reduce the number of replications, and hence the time needed [37].

The time taken by individual sampling models will increase with the size of the population to be modelled whereas deterministic cohort models presume a law of large numbers effect. For large populations therefore, providing the non-linearity assumptions are met, a cohort model may be an appropriate choice.

It may be important, however, that the population remain large in individual states. For example, in infectious disease models, the stochastic effects of local die-out of infection in states with small numbers of individuals may results in very different dynamics from deterministic models [59]. For diseases that are transmitted primarily through smaller social structures (e.g. local families and local neighbourhood transmission), then stochastic models may be preferred since the implicit large population assumptions for deterministic models do not apply.

Discussion

This new taxonomy grid extends the palette of modelling approaches beyond those commonly used in health economics. This paper describes the underlying theory linking each approach and criteria for selection.

It is the responsibility of model developers, rather than policy makers, to select the most appropriate modelling approach. It is particularly important to think creatively about the techniques to be used where there are interactions and large uncertainties.

This is because different modelling approaches can produce very different results; one example is the comparison of the different smallpox models [26,27,28,50].

The taxonomy grid is incomplete in the sense that further combinations of approaches and assumptions are also practised. A DES tool can track counts of individuals in various states rather than maintain data on each individual. A recent model of microbial drinking water infection combined ODE / CTMC models (for microbe concentration) with a stochastic ODE (for infection outcomes in humans) rather than using an unacceptably slow stochastic individual level model[60]. Mixed DES/ODE models can assess the interacting effects of system capacity, service delays, screening quality, and health outcomes [61].

Multiple modelling approaches can be used for model validation. For example, Chick et al. (2000) did this as part of a study to show how long-term partnership concurrency can strongly influence the prevalence of sexually transmitted diseases [51]. Output from differential equation, CTMC and DES models were compared to insure that the DES was coded correctly, then the DES explored how mixing patterns affected STD prevalence. Such exercises are an important part of 'structural sensitivity analysis', which can sometimes be ignored in favour of probabilistic sensitivity analysis on parameters only. It may be valuable not to define the 'best' model but rather to establish whether decisions would be different if different structures are used.

The influence of model development time on model structure selection will remain. It is often related to the experience of the modeller, the complexity of the system and the

23

software available. If the decision problem allows, simple models with a small number of states are quicker to develop and run and easy to understand. However, where multiple risk factors and non-Markovian delays are present, an individual level model may be both quicker to code and more transparent to the decision maker. Policy makers also need to be sensitive to the time required to provide appropriate models for the purpose required.

National decision makers such as NICE and equivalent bodies in other countries contribute heavily to methodology development, but there are also decision makers at international and more local levels, who utilise evidence on costs and benefits. NICE are interested in mean cost per QALY gained and PSA, which can appear to encourage researchers to use cohort model structures, but modellers need to be aware that PSA should not drive a decision to use an inappropriate structure. For other decision makers more detailed models may be needed to provide disaggregated output, such as cost estimates over time and variability between different geographical regions.

Finally, it is apparent that there is a great need for a wider awareness of the variety of techniques and their applicability. Many of these techniques are in the domain of Operational Research, epidemiology, and statistics. Increased training, communication and 'buying in' will help the Health Economics discipline. This is a high priority because the use of modelling and evidence synthesis for evaluating health technologies has developed substantially over the last 10 years. The use of the full range of techniques set out in the taxonomy grid and the criteria for appropriate model selection is the route by which the discipline will develop.

Appendix

The online appendix at http://faculty.insead.edu/chick/papers/HTAModel-AppTable.html summarizes implementation issues for the models in the taxonomy, and provides links to software reviews and vendors (the inclusion of a product or vendor is not intended to imply endorsement, and vice versa).
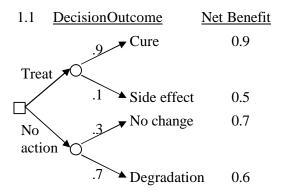
Tables and Figures

1.1   DecisionOutcome   Net Benefit

Figure 1:A Decision Tree

Figure 2: Transition Probability Matrix for a Markov Model

From\To:	Healthy	Ill
Healthy	0.95	0.05
Ill	0.25	0.75

Figure 3 An Individual Sampling Model Algorithm

For $i=1, 2, …, n$

  Sample the patient attributes (e.g. risk factors) $\boldsymbol{q}_i$

  Determine the path of patient i through the model whose
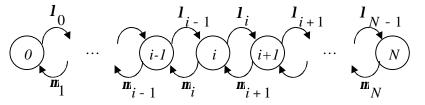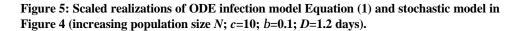   probabilities may depend upon $\boldsymbol{q}_i$

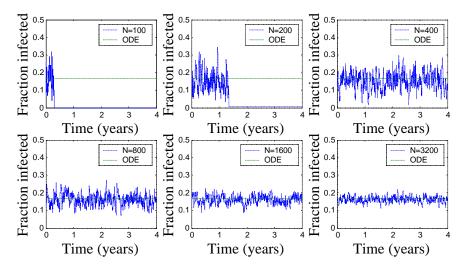 Determine cost c_i and utility (QALY) u_i for individual i

End For

Estimate the mean cost and utility E[(C U)] by $\hat{\boldsymbol{a}}_n = \dfrac{1}{n}\sum_{i=1}^{n}(c_i, u_i)$

Figure 4: Continuous-time Population Markov Model Transition Probability

Figure 5: Scaled realizations of ODE infection model Equation (1) and stochastic model in Figure 4 (increasing population size *N*; *c*=10; *b*=0.1; *D*=1.2 days).

Theorem 1: Sampling Time to First of Many Events

Theorem 1 : If $T_1, T_2, ..., T_k$ are independent random variables for the times to events 1, 2,…, k with exponential distribution $T_j \sim$ exponential(rate g_j), and mean $E[T_j] = 1/g_j$, for $j=1, 2, ...k,$

then the time to first event $T_{\min} = \min\{T_1, T_2, ..., T_k\}$ has exponential (rate $\sum_{m=1}^{k} g_m$) distribution, and the probability that $T_j = T_{\min}$ is $g_j \bigg/ \sum_{m=1}^{k} g_m$.

Table 1: Taxonomy of Model Structures

			A	B	C	D
			Cohort/Aggregate Level/Counts		Individual Level	
			Expected value, Continuous state, Deterministic	Markovian, Discrete State, Stochastic	Markovian, Discrete State, Individuals	Non-Markovian, Discrete-State, Individuals
1	No Interaction Allowed	Untimed	Decision Tree Rollback	Simulated Decision Tree (SDT)	Individual Sampling Model (ISM): Simulated Patient-Level Decision Tree (SPLDT)	
2		Timed	Markov Model (Evaluated Deterministically)	Simulated Markov Model (SMM)	Individual Sampling Model (ISM): Simulated Patient-Level Markov Model (SPLMM) (variations as in quadrant below for patient level models with interaction)	
3	Interaction Allowed	Discrete Time	System Dynamics (Finite Difference Equations, FDE)	Discrete Time Markov Chain Model (DTMC)	Discrete-Time Individual Event History Model (DT, IEH)	Discrete Individual Simulation (DT, DES)
4		Continuous Time	System Dynamics (Ordinary Differential Equations, ODE)	Continuous Time Markov Chain Model (CTMC)	Continuous Time Individual Event History Model (CT, IEH)	Discrete Event Simulation (CT, DES)

Table 2 Issues and Guidance on Choice of Model Structure

	Issue	Example	Choice of model
I1	Does the decision maker require knowledge of variability to inform the decision?	Effects of intervention are small and variable over time	Need for stochastic output (columns B to D)
I2	Is the decision maker uncertain about which sub-groups are relevant and likely to change his/her mind?	Decision maker may want to sub-divide the risk groups or test new interventions	Individual level models are more flexible to further covariates or changed assumptions (columns C and D)
I3	Is Probabilistic Sensitivity Analysis (PSA) required?	Decision maker uses cost-effectiveness acceptability curves or expected value of information	Deterministic model may be preferred (column A) but need for PSA should not drive model structure decisions
I4	Do individual risk factors affect outcome in a non-linear fashion?	Effects of age, history of disease, co-morbidity	Need to subdivide states in an aggregate model. Need to consider individual level modelling if the number is large. (columns C and D)
I5	Do covariates have multiple effects, which cause interaction?	Co-morbidities in diabetes affect renal failure and retinopathy	Individual level modelling likely to be necessary. (columns C and D)
I6	Are times in states non-Markovian?	Poor survival after an operation, moving from one age group to another, length of stay in hospital	Need to use "fixes" in Markovian models or use non-Markovian models (columns D)
I7	Is the dimensionality too great for a cohort approach?	Large number of risk factors and /or subdivision of states to get over non-Markovian effects	Individual level modelling likely to be necessary. (columns C and D)
I8	Do states 'recycle'?	Recurrence of same illness. E.g. heart attack, stop responding to drugs	Decision tree approach is probably not appropriate (rows 2 to 4)
I9	Is phasing or timing of events decisions important?	In smokers, if lung cancer occurs before bronchitis, then patient may die before bronchitis occurs	Possible to have different branches in the decision tree but Markov model or simulation may be necessary. (rows 2 to 4)
I10	Is there interaction directly between patients?	Infectious disease models	Models with interaction (rows 3, 4)
I11	Is there interaction due to constrained resources?	Models with resource constraints	Models with interaction (rows 3, 4)
I12	Could many events occur in one time unit?	Disaster, outbreak of infection, risk of co-morbidities (e.g. diabetes)	Need for small time intervals or continuous time models (row 4)
I13	Are interactions occurring in small populations?	Use in hospital catchments area rather than nationally	Need to consider individual level modelling because of the inaccuracies in using fractions of individuals (columns C, D, rows 3, 4)
I14	Are there delays in response	Rapid treatment with	Need for stochastic output and

| | due to resource constraints which then affect cost or health outcome | angioplasty and stents after a myocardial infarction | interaction (columns C, D, rows 3, 4) |
| I15 | Is there non-linearity in system performance when inherent variability occurs? | A marginal change in parameters produces a non-linear change in the system ITU is suddenly full and newly arriving patients must transfer elsewhere | DES useful |

References

[1] Griffin, Claxton, Hawkins, Sculpher. Probabilistic Sensitivity Analysis and computationally expensive models: Necessary and required? Health Economists Study Group Oxford January 2005

2 National Institute for Clinical Excellence (NICE). Guide to the Methods of Technology Appraisal. NICE: London, 2004.

3 Claxton, Sculpher, McCabe, Briggs, Akehurst, Buxton, Brazier, O'Hagan. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Economics 14: 339–347 (2005)

4 Halpern M.T., Luce B.R. Brown R.E. Genest B.. Health and Economic Outcomes Modeling Practices: A Suggested Framework. Value in Health 1998; 1(2):131-147.

[5] Chilcott J, A Brennan, A Booth, J Karnon and P Tappenden , The role of modelling in planning and prioritising clinical trials. Health Technology Assessment 2003; Vol 7: number 23

[6] Weinstein, M.C., O'Brien, B., Hornberger, J., Jackson, J., Johannesson, M., McCabe, C., and Luce, B.R. Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value in Health 2003; 6(1):9-17

[7] Z Philips, L Ginnelly, M Sculpher, K Claxton, S Golder, R Riemsma, N Woolacott and J Glanville. Review of guidelines for good practice in decision-analytic modelling in health technology assessment Health Technology Assessment 2004; Vol 8: number 36

[8] Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. Pharmacoeconomics 2000; 17: 461–477

[9] Karnon, J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Economics 2003; 12: 837-848.

[10] Koopman, J.S., Jacquez G., Chick, S.E. Integrating Discrete and Continuous Population Modeling Strategies, In Population Health and Aging: Strengthening the Dialog between Epidemiology and Demography, M. Weinstein, A. Hermalin and M.A. Stoto, Eds. Annals of the New York Academy of Sciences, 2001. 954: 268-294.

[11] Barton, P. Bryan, S. Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. Journal of Health Services Research & Policy 2004; 9(2): 110-118.

[12] Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence (NICE). Lancet 2002; 360:711–715.

[13] Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Medical Decision Making 1993; 13: 322–338

[14] Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics 1998; 13: 397–409.

[15] Eastman, R.C. Javitt J.C., Herman W.H., Dasbach E.J., Zbrozeh A.S., Dong F. et al. Model of complications of NIDDM. Model construction and assumptions, Diabetes Care, 1997; 20: 725-24.

[16]. Chilcott J.B, Whitby S.M. Moore R., Clinical Impact and Health Economic Consequences of Posttransplant Type 2 Diabetes Mellitus, Transplantation Proceedings, 33 (Suppl 5A), 32S–39S (2001)

17 Brown JB, Palmer AJ, Bisgaard P, Chan W, Pedula K, Russell A: The Mt. Hood challenge: cross-testing two diabetes simulation models. Diabetes Res Clin Prac 50 (Suppl. 3):S57-S64, 2000

[18] Brennan A, Bansback A, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. Rheumatology 2004;43:62–72.

[19] Stevenson M.D., Brazier J.E., Calvert N.W., Lloyd-Jones M., Oakley J., Kanis J.A. Description of an individual patient methodology for calculating the cost-effectiveness of treatments for osteoporosis in women. Journal of Operational Research Society 2005: 56: 214-221.

[20] Barton P., Jobanputra P., Wilson J., Bryan S. and Burls, A. The use of modeling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technology Assessment. 2004; 8(11): 1-104.

[21] Clark W., Jobanputra P. Barton P. Burls A. The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis Health Technology Assessment 2004; 8(18):1–118

[22] Jacquez, J.A., Compartmental analysis in biology and medicine - 3rd Ed., BioMedWare, Ann Arbor, MI 1996.

[23] Dangerfield, B.C., System dynamics applications to European health care issues, Journal of the Operational Research Society, 1999; 50: 345-353.

[24] Diekmann, O. and Heesterbeek, J.A. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, Wiley: 2000.

[25] Luenberger, D.G. Introduction to Dynamic Systems: Theory, Models and Applications, John Wiley and Sons, Inc. New York, 1979.

[26] Kaplan E.H., Craft D.L., Wein L. M. Emergency response to a smallpox attack: The case for mass vaccination. Proc. National Academy of Sciences. 2002; 99(16): 10935-10940.

[27] Koopman J.S. Controlling smallpox. Science 2002; 298: 1342-1344.

[28] Kaplan E.H., Wein L.M. Smallpox bioterror response. Science 2003; 300: 1503.

[29] Bennett P Hare A, Townsend J. Assessing the risk of vCJD transmission via surgery: models for uncertainty and complexity. Journal of the Operational Research Society 2005; 56(2):202-213.

[30] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, Cambridge, 1992.

[31] Çinlar E. Introduction to Stochastic Processes. Prentice-Hall. Englewood Cliffs, NJ. 1975.

[32] Banks J., Carson J.S., Nelson B.N. Discrete-event system simulation. 2nd edition. Prentice Hall, Inc. Upper Saddle River, NJ. 1995.

[33] Kurtz, T.G. 1971. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. Journal of Applied Probability, 1971; 8: 344-356.

[34] Pollett, P.K. 1990. On a model for interference between searching insect parasites. Journal of the Australian Mathematical Society, Series B, 1990; 31: 133-150.

[35] Whitt, W. Stochastic-Process Limits. Springer. 2001

[36] Brookhart M.A., Hubbard A.E., van der Laan M.E., Colford J.M., Eisenberg J.N.S., Statistical estimation of parameters in a disease transmission model: an analysis of a Cryptosporidium outbreak, Statistics in Medicine, 2002; 21(23):3627-38.

[37] Law, A.M., Kelton, W.D. Simulation Modeling & Analysis, 3rd Ed., McGraw-Hill, New York, 2000.

[38] Adams, A.L., Barth-Jones, D.C., Chick S.E., Koopman, J.S. Simulations to Evaluate HIV Vaccine Trial Methods, Simulation 1998; 71(4): 228-241.

[39] Davies R, Roderick P. Raftery J. The evaluation of disease prevention and treatment using simulation models. European Journal of Operational Research. 2003; 150(1): 53-66.

[40] Fone D. Hollinghurst S. Temple M., Round A. Lester N., Weightman A., Roberts K., Coyle E., Bevan G., Palmer S., Systematic review of the use and value of computer simulation modelling in population health and health care delivery, J. Public Health Medicine, 2003; 25(4) 325-335.

[41] Jun, J.B., Jacobson, S.H., Swisher, J.R. Applications of Discrete Event Simulation in Health Care Clinics: A Survey, Journal of the Operational Research Society, 1999; 50(2): 109-123

42 Tocher K.D. The Art of Simulation. The English Universities Press Ltd 1963:9-17.

[43] Fetter RB and Thompson JD (1965). The simulation of hospital systems. Operations Research 13: 689-711.

[44] Banks J. Handbook of Simulation, John Wiley & Sons, Inc., New York, 1996.

[45] Glynn, P. On the Role of Generalized Semi-Markov Processes in Simulation Output Analysis. In Proceedings of the 1983 Winter Simulation Conference, J. R. Wilson, J. O. Henriksen and S. D. Roberts, Eds. IEEE Inc. Pscataway NJ. 1983: 38-42.

[46] Yücesan E., L. W. Schruben. Structural and behavioral equivalence of simulation models. ACM Trans. Modeling and Computer Simulation. 1992; 2: 82–103.

[47] Cooper K, Davies R, Roderick P, Chase D and Raftery J. The development of a simulation model for the treatment of coronary artery disease. Health Care Management Science, 2002; 5: 259-267.

[48] Rauner M.S. Using Simulation for AIDS Policy Modeling: Benefits for HIV/AIDS Prevention Policy Makers in Vienna, Austria. Health Care Management Science. 2002; 5(2): 121-134.

[49] Davies R, O'Keefe RM, Davies HTO. Simplifying the modeling of multiple activities, multiple queueing, and interruptions, a new low level data structure. ACM Transactions on Modeling and Computer Simulation. 1993; 3(2): 332-46.

[50] Halloran E.M., Longini I.M., Azhar N., Yang Y. Containing bioterrorist smallpox. Science 2002; 298: 1428-1432.

[51] Chick, S.E., Adams, A., Koopman, J.S., Analysis and Simulation of a Stochastic, Discrete-Individual Model of STD Transmission with Partnership Concurrency, Mathematical Biosciences 2000; 166(1):45-68

[52] Davies R, Roderick P., Raftery J., Crabbe D., Patel P., Goddard J.R. A simulation to evaluate screening for helicobacter pylori infection in the prevention of peptic ulcers and gastric cancers: Health Care Management Science, 2002; 249-258

[53] Oakley, J. (2002). Value of information for complex cost-effectiveness models. Research Report No. 533/02 Department of Probability and Statistics, University of Sheffield.

[54] Stevenson M.D et al Gaussian process modelling in conjunction with individual patient simulation modelling. A case study describing the calculation of cost-effectiveness ratios for the treatment for established osteoporosis. Medical Decision Making 2004 24: 89-100.

[55] O'Hagan, A and Stevenson, MD (2005). Monte Carlo probabilistic sensitivity analysis for patient level simulation models. In preparation.

[56] Ross, S.M. Stochastic Processes, 2nd edition, Wiley, New York, 1996.

[57] Babad H, Sanderson C, Naidoo B, White I, Wang D. The development of a simulation model for modelling primary prevention strategies for coronary heart disease. *Health Care Management Science* 2002;**2**:69-74.

[58] NHS Modernization Agency, 10 High Impact Changes for Service Improvement and Delivery: A Guide for NHS Leaders, Leicester, 2004. [Accessed online 4 March 2005, http://www.content.modern.nhs.uk/cmsWISE/HIC/HIC+Intro.htm]

[59] Koopman, J.S., Chick, S.E., Riolo, C.S., Simon, C.P., Jacquez G. Stochastic effects on endemic infection levels of disseminating versus local contacts. Mathematical Biosciences 2002;180:49-71.

[60] Chick, S.E., Soorapanth, S., Koopman, J.S., Microbial Risk Assessment for Drinking Water. In Operations Research and Health Care: Handbook of Methods and Applications, Brandeau, M.L., Sainfort, F., and W.P. Pierskalla, Eds., Kluwer Academic Publishers, 467-494, 2004.

[61] Günes, E.D., Chick, S.E., Aksin, O.Z. Breast Cancer Screening: Trade-offs in Planning and Service Provision, Health Care Management Science 2004; 7(4): 291-303.