Paper ID: 33

Fragile - handle with care!
Issues in the long-term term extrapolation of individual patient level data for cost-effectiveness modelling

A Manca[1,2] , P Saramago[1] MJ Sculpher[1], P Royston[2], M Parmar[2], A Copas[2]

1. *Centre for Health Economics, The University of York, UK*
2. *Clinical Trials Unit, Medical Research Council, UK*

Abstract

Background Much of the evidence informing cost effectiveness models comes from RCTs which are short in durations. However, resource allocation decisions often require long term estimates of costs and quality-adjusted survival. Thus, researchers are expected to extrapolate RCT results beyond the observed period, with or without the benefit of additional external information. Survival regression models (SRMs) can naturally be used to implement such an extrapolation. Two issues must be considered though when carrying this out. First, how well SRMs fit/predict the RCT data, and second, the extent to which their results can be used to inform decision models aimed at producing (beyond trial) long-term cost-effectiveness estimates.

Aims This paper identifies a set of desirable features SRMs should have for use in cost-effectiveness modelling. Distinguishing between *within* and *beyond*-trial prediction, the manuscript discusses the extent to which competing SRMs satisfy the proposed 'wish list'. The candidate SRMs are tested using both a micro-simulation and a real life dataset.

Methods SRMs for IPD and probabilistic decision models.

Data RITA3 trial.

Results SRMs for cost-effectiveness modelling must be able to capture the time-varying nature and shape of the baseline risk, as well as estimate the effect of covariates on the hazard rate. Standard parametric SRMs may be too rigid. The Cox model is limited by how it handles baseline hazard. The Royston-Parmar flexible parametric model provides a middle ground solution and has a series of desirable features for use in CEA. Extrapolation beyond trial follow-up for cost-effectiveness purposes, though, ideally requires mature survival data and external information.

Conclusions Beyond trial extrapolation of time-to-event data without external information must be handled with care and the uncertainty associated with extrapolation should be captured as fully as possible in models.

Correspondence to: A. Manca, Centre for Health Economics, University of York, UK. Email: am126@york.ac.uk, Tel (0)1904 321430, Fax (0)1904 321402.

1. Introduction

Economic evaluation is increasingly used to inform policy making decisions concerning the adoption of new healthcare technologies. The National Institute for Health and Clinical Excellence (NICE) for England and Wales is an example of the type of agency that requires cost-effectiveness information to issue recommendations regarding the likely value for money of the health technologies undergoing the Institute's review process.

One of the key elements of any such analyses is their time horizon. While most national guidelines are silent on this respect, some (including NICE) [1] explicitly indicate that the time horizon for estimating cost-effectiveness must be sufficiently long to reflect important differences in costs and outcomes between the health technologies being evaluated. In many cases, the appropriate time horizon for the analysis will be patients' lifetimes.

Most randomised controlled clinical trials (RCTs), though, have durations that are shorter than patients' lifetimes. To produce cost-effectiveness evidence to support healthcare policy decisions the analyst will have to use some form of modelling to *extrapolate*[a] the trial data beyond the observed follow up period [2, 3], and there may or may not be relevant additional data.. Survival analysis may naturally be used to carry out such an extrapolation. In particular, survival regression models (SRMs) provide a formal way of establishing the link between the risk of particular clinical events (and hence costs and outcome effects) and time.

This paper focuses on the methodological issues surrounding the use of SRMs for the analysis of trial-based IPD to extrapolate the trial results beyond the study follow up period with the objective to estimate the long-term cost-effectiveness of competing strategies.

It is argued that two separate but related issues must be considered when extrapolating time-to-event IPD for CEA purposes. First, how well SRMs fit/predict the RCT data, and second, the extent to which their results can be used to inform decision models aimed at producing (beyond trial) long-term cost-effectiveness estimates. Using SRMs that are unfit for extrapolation purposes or applying these models making inappropriate or untenable assumptions may introduce inaccuracy in the estimation of the time-to-event outcomes and costs, ultimately leading to biased cost-effectiveness estimates.

[a] The term 'extrapolation' has been used to describe instances where the trial data were (i) adjusted to describe clinical practice (i.e. extrapolation from study to practice); (ii) adapted to make them applicable to another country (i.e. extrapolation from country to country); (iii) analysed to estimate occurrence of costs and effects beyond the study follow up period. In this paper we use the latter definition.

The manuscript begins with a brief review of the methods, followed by an attempt to define a set of desirable features SRMs should have for use in cost-effectiveness modelling. These characteristics include, for instance, the flexibility with which these models can describe baseline hazard, capture the effect of covariates on the hazard rate, and facilitate extrapolation of baseline. Distinguishing between *within* and *beyond*-trial predictions, the paper discusses the extent to which competing SRMs satisfy the proposed 'wish list'. Candidate SRMs are then tested using both a micro-simulation and a real life dataset. The discussion section summarises the manuscript's contribution and future lines of research.

2. Methodological background

A common approach to estimate long-term costs and outcomes in economic evaluation is to develop a state-transition model [4-6] simulating the occurrence of the events of interests (e.g. patient's progression towards more severe health states). In an increasing number of models, the input parameters governing the occurrence of these events in the model are probability estimates derived using a system of risk (regression) equations applied to time-to-event data. The Framingham risk equations [7] for cardiovascular disease risk profiles – for instance – were developed using parametric survival regression techniques. The objective here was to produce a prediction model for the risk of specific competing events in this patient population (e.g. myocardial infarction, coronary heart disease (CHD), death from CHD, stroke, cardiovascular disease, and death from cardiovascular disease) conditional on a set of individual characteristics. There are various applied examples of this approach to trial-based IPD analysis for cost-effectiveness modelling (see for examples [8-12]).

Given the developing role of survival analysis in healthcare economic evaluation it seems appropriate to start with a brief review of some of the key concepts and mathematical relationships underpinning statistical modelling of time-to-event data.

Survival analysis is concerned with studying the timing of events whose occurrence is uncertain within the period of observation. For instance, a RCT comparing two oral inhalers for asthma management will be designed to assess whether the time to asthma exacerbation in the two treatment arms differs significantly and perhaps whether this is related to any of the patients' characteristics.

Let us denote *T* a positive random variable, with a given *probability distribution* representing survival times, and *t* the actual observed time to the event of interest.[b] The first concept one needs to become familiar with is the *survival function*, which is a monotone, non-increasing, differentiable function of time:

$$S(t) = \Pr(T > t) \tag{1}$$

This is typically used to estimate the probability of observing the event of interest beyond time *t* (or equivalently, the probability that there is no event before *t*). Notice that as $t \to 0$ then $S(t) \to 1$, and as $t \to +\infty$ then $S(t) \to 0$. From the survival function one can derive the probability *density function*

$$f(t) = \frac{dF(t)}{dt} = \frac{d}{dt}\{1 - S(t)\} = -S'(t) \tag{2}$$

where $F(t) = 1 - S(t)$. When *T* is continuous this function represents the (instantaneous) unconditional failure rate of event occurrence [13]. The most important concept in survival analysis is the *hazard function*. This function captures the relationship between survival function and failure times

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t + \Delta t > T > t \mid T > t)}{\Delta t} = \frac{f(t)}{S(t)} \tag{3}$$

and estimates the *hazard rate*; that is, the rate at which units fail by *t* (or at an infinitesimally small interval $[t; t + \Delta t]$) given that they survived until *t*. In this sense, the hazard rate is a conditional rate. It should be noted that the shape of the hazard function is determined by the nature of the phenomenon under study and that the hazard rate can be constant or it can change over time. Finally, the *cumulative hazard function*

$$H(t) = \int_0^t h(u)du = \int_0^t \frac{f(u)}{S(u)}du = -\int_0^t \frac{1}{S(u)}\left\{\frac{d}{du}S(u)\right\}du = -\ln\{S(t)\} \tag{4}$$

measures the total amount of risk that has been accumulated up to time *t*. In summary, using (1)-(4) one can rewrite the survival function and its density as functions of the cumulative hazard

$$\begin{aligned} S(t) &= \exp\{-H(t)\} \\ F(t) &= 1 - \exp\{-H(t)\} \\ f(t) &= h(t)\exp\{-H(t)\} \end{aligned} \tag{5}$$

[b] This can be thought as a realisation of the random variable *T*.

The class of survival models used to extrapolate trial-based cost-effectiveness data beyond the study follow-up period assume that the observed time-to-event data follow a specific parametric distribution.

The simplest parametric survival model posits that the time-to-event data follow an exponential distribution, with hazard equal to a single parameter, $h(t)=\lambda$, which implies a constant hazard with respect to time (i.e. the risk of observing the event of interest is constant over time).

The parameter λ can be made covariate-dependent. In this case the hazard function with covariates for the exponential family can be written as $h(t|\mathbf{x})=\exp[-(\beta_0+\beta_1 x_1+\ldots+\beta_n x_n)]$, where the constant β_0 represents the baseline hazard rate (i.e. $h(t)=\lambda=\exp[-(\beta_0)]$) . Notice that any increase or decrease with respect to the baseline hazard is a function of the covariates only, as can been seen by calculating the ratio of the two *hazards.* In the simple case with a binary covariate (e.g. x_1=0,1) the hazard ratio is

$$\frac{h_i(t|x_1=1)}{h_i(t|x_1=0)}=\frac{\exp[-(\beta_0+\beta_1 x_1)]}{\exp[-(\beta_0)]}=\exp[-(\beta_1)] \qquad (6)$$

which demonstrates the *proportionality of the hazards* assumption.

The simplicity of the exponential model is also its weakness when modelling survival data. Equation (6) can also be used to show that the exponential distribution is 'memoryless'; where the future is independent of the past, and the fact that an event has not happened yet, tells us nothing about how much longer it will take before it does happen. The probability of an event occurring beyond some time $t_1=(t_0+s)$ conditional on having survived t_0 is equal to

$$\Pr(T\geq t_1|T\geq t_0)=\frac{\Pr(T\geq t_1)}{\Pr(T\geq t_0)}=\frac{S(t_1)}{S(t_0)}=\frac{\exp(-\lambda t_1)}{\exp(-\lambda t_0)}=\exp[-\lambda(t_1-t_0)]=\exp(-\lambda s) \qquad (7)$$

Its simplicity may be the reason behind its widespread use in CEA. However, this choice may be justifiable in (cost-effectiveness) models where the analyst does not have access to IPD, but is perhaps more difficult to defend it when individual specific survival time information is available. In fact, the exponential may be regarded as being too rigid a parameterisation in many processes modelled in CEA. A more general formulation, again commonly used in CEA is the Weibull distribution. The hazard function here takes the form

$$h(t)=\lambda p(\lambda t)^{p-1} \qquad (8)$$

where *p* is the (ancillary) *shape* parameter estimated by the data, and λ is the *scale* parameter. When *p*<1 the hazard decreases with time, and analogously when *p*>1 the hazard increases with

time. When *p*=1 the hazard is constant over time, leading to the special case defined by the *exponential distribution*.

<<Table 1 about here>>

Once again, we can make the parameter λ in the Weibull covariate-dependent, and derive a hazard function of the form $h(t|x_{1t},....x_{nt}) = pt^{p-1}\exp(\beta_0 + \beta_1 x_{1t} + ...\beta_n x_{nt})$.

There are many more parametric distributions for time-to-event data and the more commonly used in CEA are reported in Table 1, together with the functional form of the baseline hazard and the hazard function. Following the relationships defined in (4) and (5), for each of the models in Table 1 we write the baseline hazard function , its covariate-dependent formulation and the resulting transition probabilities (assuming discrete time cycles of length *u*) for use in a state-transition cost-effectiveness model.

A popular time-to-event regression often used in medical research is the Cox (proportional hazards) model [14], which can be written as $h_i(t) = h_0(t)\exp(\beta_1 x_1 + ... + \beta_n x_n)$ where $h_0(t)$ is the baseline hazard function and $\exp(\beta_1 x_1 + ... + \beta_n x_n)$ the impact of covariates on the baseline hazard. While the parametric models presented so far explicitly define a functional form for $h_0(t)$, in the Cox model the form of the baseline hazard rate is assumed to be unknown[c] and is left unspecified (notice that, in fact, Cox models do not have an intercept term). The latter assumption is also its main limitation for use in cost-effectiveness modelling. While it is possible to derive an estimate of the baseline risk, as far as we are aware it is not possible to represent it in close form. The parametric models discussed so far are desirable if one has a good reason to expect the duration dependency to exhibit some particular form. The popularity of the Cox model as an alternative to a fully parametric model in clinical research rests on the fact that the distributional form of the duration times is left unspecified, although estimates of the hazard and baseline survivor estimates can be retrieved empirically [13]. Furthermore, many analyses in medicine and the social sciences are more interested in the relationship between some outcome and the covariates of interest and less focussed on the time-dependency element. However, in cost-effectiveness modelling the time-dependency of the events of interest (e.g. transition to worse health states, occurrence of clinical events) is as important as the relationship between outcomes and covariates. For instance, we may want to be able to model whether - for the cohort being modelled - the risk of experiencing the event of interest increases over time, if it

[c] The definition of semi-parametric model comes form the fact that it is only the hazard function which is assumed to follow a parametric distribution, while the baseline hazard rate is treated as a nuisance parameter in the Cox model.

does it monotonically, and whether the hazard is proportional across patients groups. The issues in using the Cox model to derive transition probabilities limit its usefulness in the model-based cost-effectiveness field.

One approach that can be used to overcome the limitations of the Cox model while still offering the possibility of explicitly modelling baseline hazard is to use the methodology developed by Royston and Parmar [15]. Here natural cubic splines[d] are used to model the baseline log cumulative hazard and log cumulative odds of failure functions (notice the model can be expressed in terms of proportional hazards and proportional odds). These models are based on a transformation of the survival function by a link function

$$g[S(t;\mathbf{z})]=g[S_0(t)]+\beta'\mathbf{z} \qquad (9)$$

where $S_0(t) = S(t;0)$ is the baseline survival function and β is a vector of parameters to be estimated for covariates $\boldsymbol{z}$. The authors use the Aranda-Ortaz link function

$$g(x;\theta) = \log\frac{x^{-\theta}-1}{\theta} \qquad (10)$$

where θ = 1 corresponds to the proportional odds model and θ → 0 to the proportional hazards model. Royston and Parmar [15] use natural cubic splines to model $g[S_0(t)]$ within the Aranda-Ordaz family of link functions. As indicated by the authors modelling the relationship between the baseline log cumulative hazard or log cumulative odds of failure and log time as linear rather than by using splines, the approach reduces to fitting Weibull or Log-logistic distributions [16]. It should be noted that smoothing $g[S_0(t)]$ on the log time scale means that the fitted function is typically gently curved or nearly linear, and is usually very smooth.

3. Survival regression for cost-effectiveness models: a 'wish list'

Since survival regression is increasingly used to inform cost-effectiveness models, it is legitimate to ask what features cost-effectiveness modellers might look for in a time-to-event regression model. Table 2 lists our proposed 'wish list'.

Perhaps one of the most desirable features is the ability to *explicitly model the (baseline) hazard*. The baseline risk for the event(s) of interest is often the backbone of a cost-effectiveness model, in the sense that it is used to describe the natural history of the disease or disease progression under existing clinical management. Prognosis with treatment (or a new treatment) is

[d] Splines are special functions defined piecewise by polynomials. Smoothing splines may be viewed as generalizations of interpolation splines where the functions are determined to minimize a weighted combination of the average squared approximation error over observed data and the roughness measure.

subsequently obtained by applying the relative treatment effect (e.g. relative risk reduction) estimates onto the relevant baseline parameter(s) in the model (e.g. probability of making a transition towards a more severe health state).

While this depends on the data and the context in which it is used, the one-parameter survival model may be too rigid when the objective is to describe patient's baseline risk and model natural history, since this model only allows constant hazard with respect to time. Slightly more flexible are the Weibull and Gompertz that allow monotone (increasing or decreasing) hazard rates. The other two-parameter models (e.g. log-logistic and log-normal distribution) are even better in this sense given they admit a non-monotone hazard rate over time. The problem with these non-monotonic hazard models is that they are non-proportional hazard models and their use would force the analyst to work using a different metric, which in itself may be acceptable but it may become a problem when synthesising the survival model results with existing evidence estimated using proportional hazard models.

Since there are many processes that do not exhibit monotone hazard with respect to time it would be desirable to use survival regression models that afford more flexibility in this dimension. The generalised Gamma (GG) is a three-parameter model that can capture non-monotone hazard and allows a flexible formulation of the baseline risk. Depending on the value of the two ancillary parameters, the GG becomes a Weibull or a log-logistic, inheriting their pro and cons. Finally, another rigidity of the one- and two-parameter models is the fact that they do not let the analyst chose the distribution with which to model the time to event data.

The Cox proportional hazard model can capture non-monotonic patterns of the hazard and it is as easy to implement as the models discussed so far. However, the main limitation of the Cox proportional hazard model is that it treats the baseline function as a nuisance parameter, leaving it unspecified. Even if it is possible to obtain empirical estimates of the baseline risk in the Cox model, this will be very fuzzy and it would not lend itself to use in cost-effectiveness modelling.

The Royston and Parmar (RP) model on the other hand allows greater flexibility when the objective is to model baseline risk, thanks to the use of spline functions. It can deal with non-monotone hazard and can be formulated using either the proportional hazard or the proportional odds metric, offering greater range of choices to the analyst. Another advantage is that the RP approach does not require prior specification of the distribution for the time to event variable.

With the exception of the Royston and Parmar models - which can currently be fitted only in STATA using the user-written command `stpm2` - all the other models listed in Table 2 can be estimated using most standard statistical software (e.g. R, STATA, SAS).

Perhaps the most important feature we may be looking for in these models is the extent to which they can be used to extrapolate the observed data beyond the trial follow up period. There are two issues that need to be considered when selecting an appropriate SRM for extrapolation purposes. The first is how well the model fits the observed data, and the second is how much trust we can place in using the best data-fitting SRM for extrapolation purposes. In fact, analysts often assume that the best model to represent the within-trial data is also the best model to extrapolate these beyond the trial follow up period. The problem is that while there are relatively well established rules for choosing the best fitting survival regression model for the within-trial analysis, it is not clear how we can establish if the model can be trusted for out-of-sample predictions without using external data as a benchmark. Without external data the possibility of dramatic changes in covariate effects or baseline hazard beyond the trial period cannot be excluded. Mindful of this, in the absence of external data and in the impossibility of discriminating between competing models some authors suggested to use statistical model-averaging.[17] However, unless the time-to-event data being analysed are mature, this solution is only appropriate for a within-trial analysis.

There are some examples where analysts have carried out extrapolation for cost-effectiveness purposes using external information. For instance, Demiris and Sharples [18] use age and sex matched population average data from the UK Government Actuary's Department (GAD) to extrapolate the observed survival of individuals at risk of sudden cardiac death with amiodarone *versus* implantable cardioverter defibrillator (ICD). In their analysis of the multinational EUROPA trial, Briggs *et al* [10] modelled non-disease mortality using ONS standard life tables for non-cardiovascular mortality using a cause elimination approach, rather than using the trial data. Cardiovascular mortality was modelled from the EUROPA trial data and extrapolated beyond trial using an exponential model. Epstein *et al [12]* used the EVAR1 trial data to estimate the long term cost and quality-adjusted survival of patients undergoing endovascular versus open repair for abdominal aortic aneurysm. Consistent with the approach used by Briggs *et al* [10], mortality from non-cardiovascular causes in the EVAR1 population was estimated from age- and sex-specific population life tables, adjusted to exclude deaths from cardiovascular causes. The rate of non-fatal cardiovascular events and aneurysm-related procedures was estimated from EVAR1. A promising method, yet to be used in cost-effectiveness modelling to achieve extrapolation beyond

trial duration, is 'relative survival'. Relative survival estimates the mortality rate for patients with the condition of interest after correcting for estimated mortality from all other causes. Nelson *et al* [19], for instance, use the Royston and Parmar flexible parametric survival model in combination with relative survival methods to extrapolate cardiac mortality data observed in the coronary unit of the Leicester Royal Infirmary, modelling expected mortality using rates from the England and Wales population data matched on age, sex and period of hospitalization.

The relative survival method relies on the existence of long-term time-to-event external data onto which to anchor the extrapolation of the trial results. The use of these methods is feasible in some clinical areas, like cancer and cardiovascular disease, where long term registers or population-based data on mortality exist. Their applicability however is limited in the absence of long term external data. One interim way forward could be the use of appropriately elicited expert opinion[20, 21]. Furthermore, cost-effectiveness models often need to consider the time-to-occurrence of a non-fatal event (e.g. asthma episode) or time-to-transitions to health states associated with poor health-related quality of life and high costs (e.g. terminally ill). The use of relative survival methods (and for that matter any other adjustment requiring external data) in these cases is limited by the availability of routinely collected information on the occurrence of these 'events of interest'.

4. Motivating example – RITA3

In this section, the use of SRMs for cost-effectiveness modelling is illustrated using a specific case study: the Intervention Trial of unstable Angina (RITA3). A decision analytic model based on the RITA3 trial has been already reported [22]. The analysis from the original paper was replicated in the current work with a view to test the impact of using alternative SRMs to estimate the probability (and sampling uncertainty) of time to primary composite endpoint on the decision model. To facilitate comparison of results all the other assumptions surrounding the original analysis were maintained, including model specification and choice of covariates for the regression models used to estimate the other parameters in the decision model.

5.1 Rationale and model structure

Briefly, a state-transition model was used to extrapolate the RITA3 results to patients' lifetimes' horizon. A series of risk equations were estimated to determine the rates of cardiovascular death or non-fatal MI during the index hospitalisation and for the remainder of the trial follow-up period. These estimates of effectiveness were then incorporated into the cost-effectiveness model, which had two parts: a short-term decision tree (instantaneous in time) and a long-term state-transition structure. The main purpose of the short-term tree was to reflect the distribution

of the trial cohort over the initial health states the prognosis of which was to be described by the state-transition model. The short and long-term models represent respectively the index hospitalisation and the post-index hospitalisation, respectively. Costs and QALYs were determined for the index hospitalisation and for each state in the long-term Markov structure. The model structure is shown in Figure 1. The box [MI/CVD] in the figure indicates that a composite event has occurred during a cycle and does not represent a formal health state since patients are then assigned to either a fatal or non-fatal state based on a separate calculation.

5.2 Survival analysis to estimate transition probabilities for the cost-effectiveness model

In the original analysis to estimate the risk of the combined endpoint of cardiovascular death or MI during the trial period, a time-to-event Weibull proportional hazards model was employed with the starting time set at hospital discharge (Equation 2 in Figure 1). In extrapolating beyond the period of trial follow-up (5 years), a conservative assumption of no continued treatment effect from the early interventional strategy was made.

Focussing on equation 2 only we explore the impact of using the following models discussed in Section 3: *Weibull, Exponential, Gompertz, Generalized Gamma and Royston-Parmar models*. Table 3 compares the SRMs examined in terms of their deviance statistic. It suggests that the Royston and Parmar flexible parametric model (with 2 knots = 3 degrees of freedom) is slightly more efficient than the Weibull (original model) and the Generalized Gamma specifications.

Yearly probabilities of experiencing a composite event (for risk group 1) in each treatment group for the 5 years duration of the trial have been estimated using the Royston and Parmar model and, for comparison, the Weibull regression. This was done using the formulas reported in Table 1. Figure 2 plots these time-dependent transition probability estimates. There is little difference between the predicted transition probabilities during the first year after index admission. At the beginning of year 2 the Weibull model predicts monotonically decreasing probabilities of experiencing the composite event until year 5, while the Royston-Parmar model suggests that this probability decreases during year 2 but then increases ever so slightly during year 3 reaching a *plateau* thereafter.

Figure 3 plots the trial-wide estimated survival functions using the Weibull and the Royston-Parmar method, comparing these against the Kaplan-Meier curves. It can be seen that although there is little to choose between the Weibull and Royston-Parmar approach, due to its additional flexibility the latter approximates much more accurately the observed occurrence of composite event in the trial.

5. Discussion

Survival regression methods are a key tool for estimating input parameters for cost-effectiveness models. We proposed a set of desirable features the analyst may want to look for in a SRM when assessing the performance of competing models. The choice of the appropriate model is not only relevant for *within* trial analyses and predictions of observed time to event data (e.g. it is important to make sure the model accurately reflects the trial data), but also – and some would say especially important - for long term *extrapolation* of the trial results. The latter is an important objective of many cost-effectiveness models.

While acknowledging that this practice may be dictated by lack of patient level data, we invite the reader to consider the risks of uncritically use simple parameterisations to represent time-to-event processes in cost-effectiveness models. We argued that the Royston-Parmar parametric survival regression is a more flexible approach to model the hazard function, providing a better understanding of the impact of the baseline hazard through the approximation of a desired baseline distribution function. The RP method is particularly useful for extrapolation purposes since it allows out-of-sample prediction to be based on the hazard estimates obtained for the time period in which the survival data are more mature, rather than using the entire observation period like most standard methods. In the RITA 3 trial survival data the flexible parametric proportional hazards modelling provided a slightly superior fit to the trial data over the Weibull modelling strategy implemented in the original study. Further examination of the competing SRMs for use in cost-effectiveness modelling is needed, particularly in situations where the time-to-event data display different characteristics (e.g. non-proportional hazards).

With respect to the issue of extrapolation it is argued though that no matter how good the survival model used to analyse the *within* trial data, it is of essence to be able to use external data when extrapolating the trial results. *Beyond* trial extrapolation of time-to-event data without external information must be handled with care and the uncertainty associated with extrapolation should be captured as fully as possible in models.

Issues for discussion

Is the proposed wish list comprehensive of the features health economists typically look for in a survival regression model?

How useful will a micro-simulation exercise be as opposed to have more case studies?

References

[1] International Society for Pharmacoeconomics and Outcomes Research. Pharmacoeconomic Guidelines Around the World - comparative table. [web] 2009 [cited 2009 9 April]; Available from: http://www.ispor.org/PEguidelines/index.asp

[2] Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in economic evaluation: an unavoidable fact of life. Health Economics. 1997;6(3):217-27.

[3] Sculpher MJ, Claxton K, Drummond MF, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Economics. 2006;15(7):677-87.

[4] Briggs AH, Sculpher MJ. An Introduction to Markov Modelling for Economic Evaluation. PharmacoEconomics. 1998;13(4):397-409.

[5] Sonnenberg FA, Beck JR. Markov models in medical decision making: A practical guide. Medical Decision Making. 1993;13(4):322-38.

[6] Kuntz KM, Weinstein MC. Life expectancy biases in clinical decision modeling. Med Decis Making. 1995 Apr-Jun;15(2):158-69.

[7] Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J. 1991 121(1 Pt 2):293-8.

[8] Lindgren P, Buxton M, Kahan T, Poulter NR, Dahlof B, Sever PS, et al. The Lifetime Cost Effectiveness of Amlodipine-Based Therapy Plus Atorvastatin Compared with Atenolol Plus Atorvastatin, Amlodipine-Based Therapy Alone and Atenolol-Based Therapy Alone: Results from ASCOT1. Pharmacoeconomics. 2009;27(3):221-30.

[9] Lindgren P, Geborek P, Kobelt G. Modeling the cost-effectiveness of treatment of rheumatoid arthritis with rituximab using registry data from Southern Sweden. Int J Technol Assess Health Care. 2009 Spring;25(2):181-9.

[10] Briggs A, Mihaylova B, Sculpher M, Hall A, Wolstenholme J, Simoons M, et al. Cost effectiveness of perindopril in reducing cardiovascular events in patients with stable coronary artery disease using data from the EUROPA study. Heart. 2007 Sep;93(9):1081-6.

[11] Briggs AH, Bousquet J, Wallace MV, Busse WW, Clark TJ, Pedersen SE, et al. Cost-effectiveness of asthma control: an economic appraisal of the GOAL study. Allergy. 2006 May;61(5):531-6.

[12] Epstein DM, Sculpher MJ, Manca A, Michaels J, Thompson SG, Brown LC, et al. Modelling the long-term cost-effectiveness of endovascular or open repair for abdominal aortic aneurysm. Br J Surg. 2008 Feb;95(2):183-90.

[13] Box-Steffensmeier JM, Jones BS. Event history modelling: a guide for social scientists. Cambridge: Cambridge University Press 2004.

[14] Cox DA. Regression models and life tables. Journal of the Royal Statistical Society - Series B. 1972;34:187-220.

[15] Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med. 2002 Aug 15;21(15):2175-97.

[16] Royston P, Parmar MK. Flexible parametric alternatives to the Cox model, and more. The Stata Journal. 2001;1(1):1-28.

[17] Jackson CH, Thompson SG, Sharples LD. Accounting for uncertainty in health economic decision models by using model averaging. J R Stat Soc Ser A Stat Soc. 2009 Apr;172(2):383-404.

[18] Demiris N, Sharples LD. Bayesian evidence synthesis to extrapolate survival estimates in cost-effectiveness studies. Stat Med. 2006 Jun 15;25(11):1960-75.

[19] Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. Stat Med. 2007 Dec 30;26(30):5486-98.

[20] O'Hagan A, Buck CA, Daneshkhah A, Eiser JR, al. e. Uncertain Judgements: Eliciting Experts' Probabilities. Chichester, UK: John Wiley & Sons Ltd 2006.

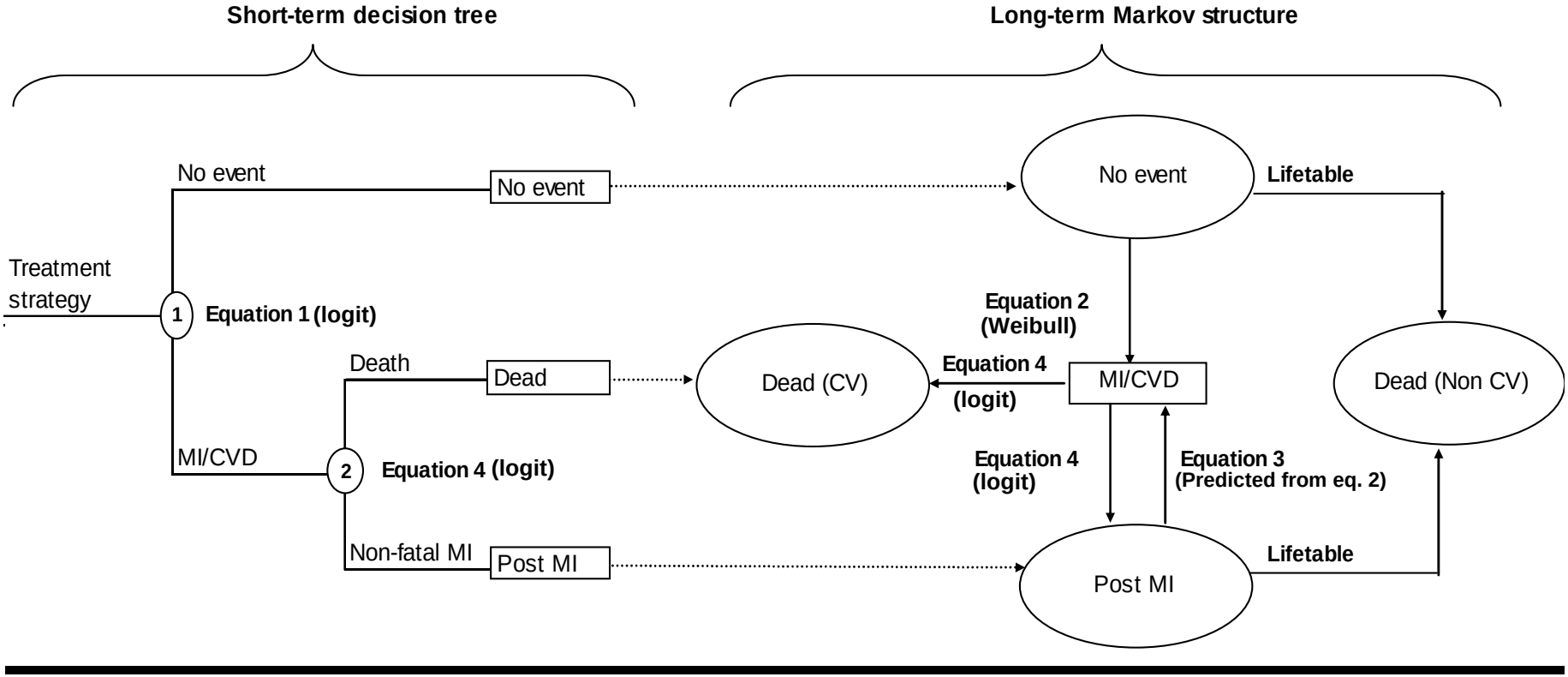
[21] Ayyub BM. Elicitation of Expert Opinions for Uncertainty and Risks. Boca Raton, FL, USA: CRC Press 2001.

[22] Henriksson M, Epstein DM, Palmer SJ, Sculpher MJ, Clayton TC, Pocock SJ, et al. The cost-effectiveness of an early interventional strategy in non-ST-elevation acute coronary syndrome based on the RITA 3 trial. Heart. 2008 Jun;94(6):717-23.

List of Tables

List of Figures

Figure 1. Model structure of the cost-effectiveness analysis of the RITA 3 trial (MI=myocardial infarction, CV=cardiovascular, CVD=cardiovascular death)



Source: Henriksson *et al.* (Heart 2008)

Figure 2: Yearly probabilities for risk group 1 and for each year of the 5 year tunnel state of the RITA 3 trial using the Weibull and the Royston-Parmar models

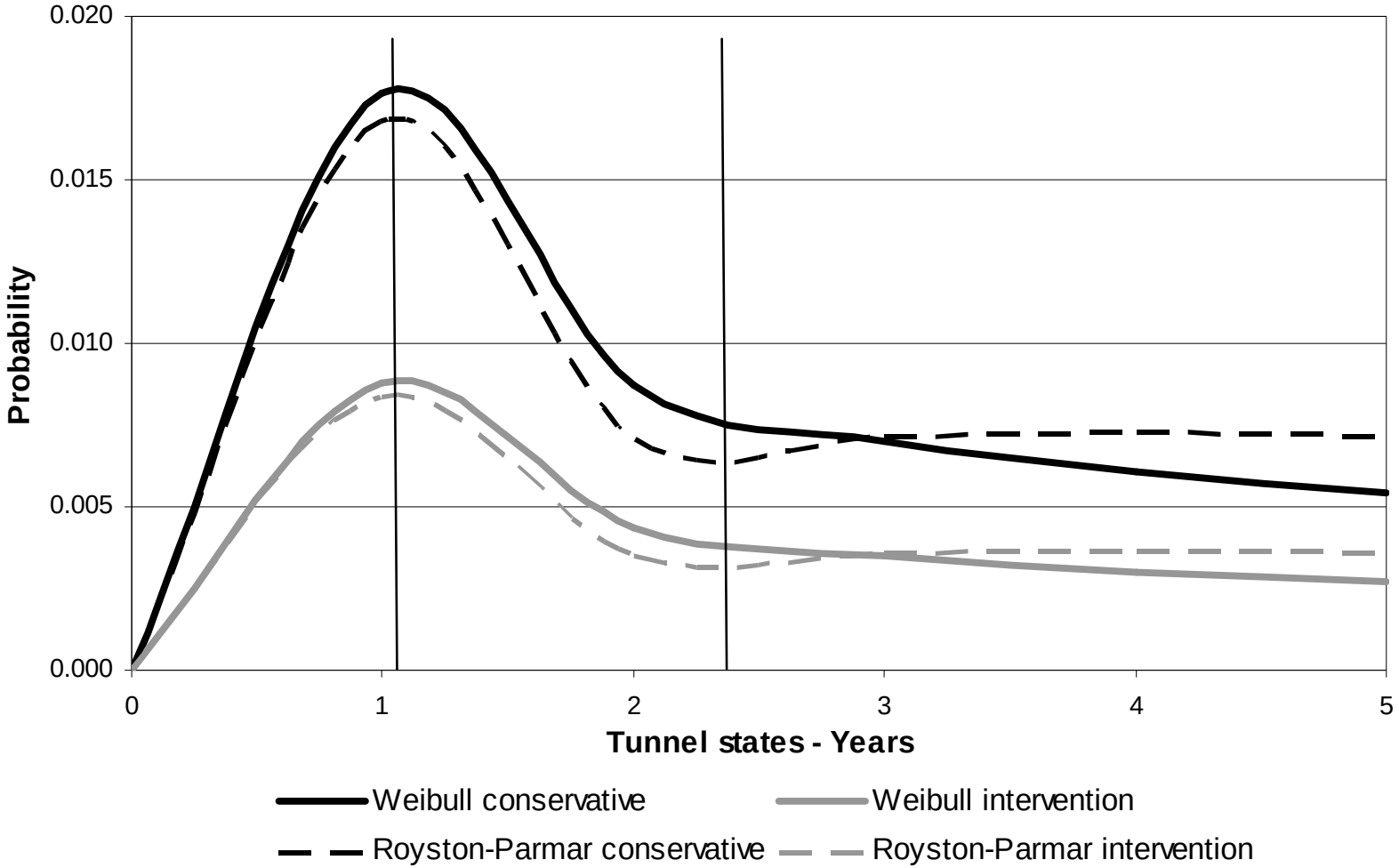


Figure 3. Estimated survival of cardiovascular death or myocardial infarction from hospital discharge until end of the trial – Weibull regression vs Royston-Parmar vs Kaplan Meier

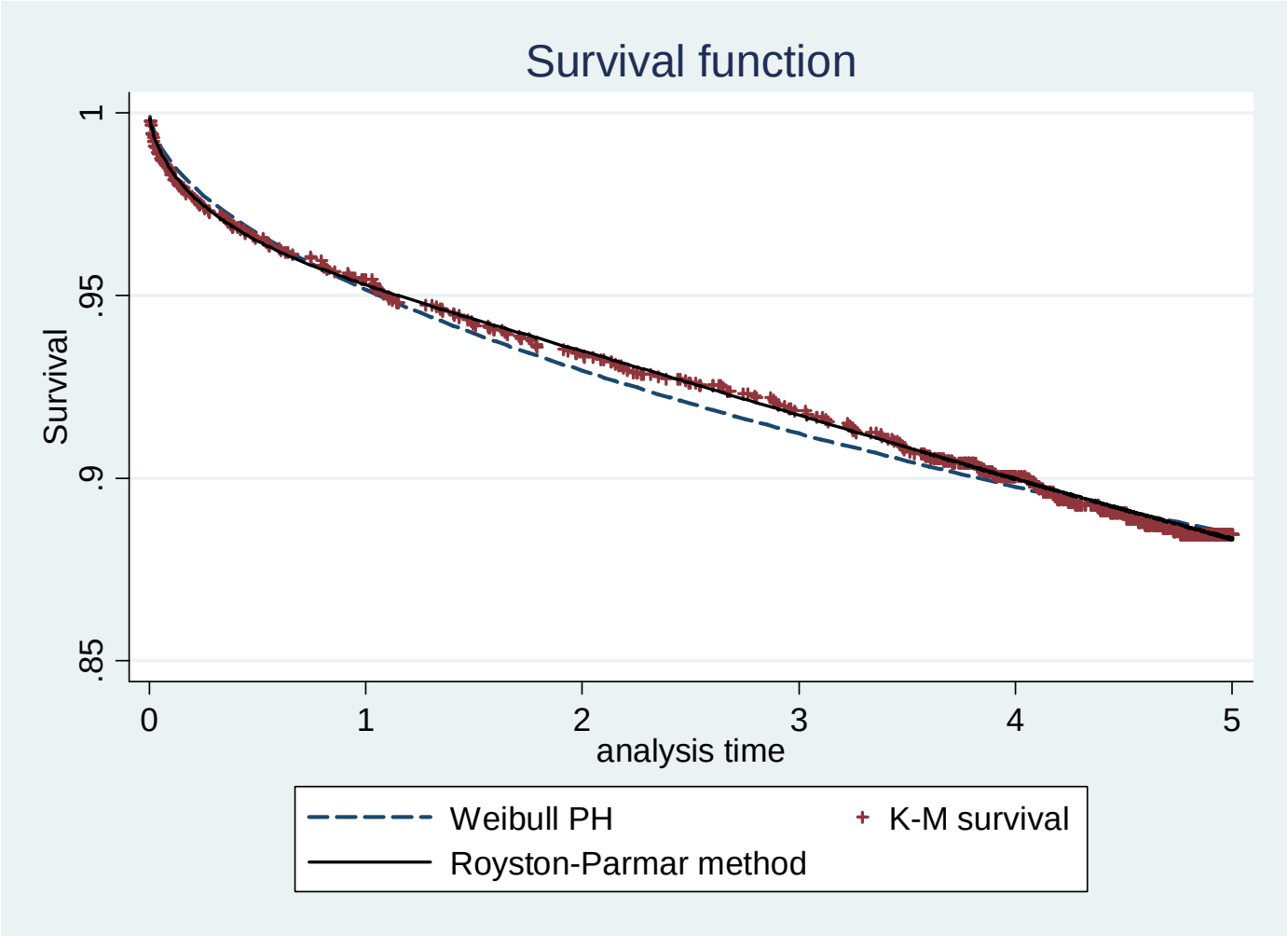


Table 1: Functional forms of baseline hazard, hazard function and derived transition probabilities for the most common survival models used in cost-effectiveness models[1]

	Baseline hazard	Hazard function	Transition probability[2]
One parameter models			
Exponential	$\exp(-\beta_0)$	$\exp(-\beta_0)\cdot\exp(-(\beta_1 x_1+\ldots+\beta_n x_n))$	$tp(t)=1-\exp\left[H(t-u)-H(t)\right]=1-\exp\left[\lambda(t-u)-\lambda t\right]=1-\exp(-\lambda u)$
Two-parameter models			
Weibull	$pt^{p-1}\exp(\beta_0)$	$pt^{p-1}\exp(-\beta_0)\cdot\exp(-(\beta_1 x_1+\ldots+\beta_n x_n))$	$tp(t)=1-\exp\left[\lambda(t-u)^p-\lambda t^p\right]$
Gompertz	$\exp(\theta t)\exp(\beta_0)$	$\exp(\theta t)\exp(\beta_0)\exp(\beta_1 x_1+\ldots+\beta_n x_n)$	$tp(t)=1-\exp\left[\frac{\lambda\{\exp(\theta(t-u))-1\}}{\theta}-\frac{\lambda\{\exp(\theta(t))-1\}}{\theta}\right]$
Three-parameter models			
Generalised Gamma	4	To be completed	4
Others			
Cox PH model	*undefined*	$h_0(t)\cdot\exp(\beta_1 x_1+\ldots+\beta_n x_n)$	*cannot be estimated*
Royston-Parmar[5]	$s(x;\gamma)=\gamma_0+\gamma_1 x+\gamma_2 v_1(x)+\gamma_3 v_2(x)$	$h(t\mid x)=\frac{\partial H(t\mid x)}{\partial t}=\frac{\lambda t^{\gamma_1}\exp\left[\gamma_2 z_1(\ln t)+\gamma_3 z_2(\ln t)\right]}{\partial t}$	$tp(t)=1-\exp\left[\lambda(t-u)^{\gamma_1}\cdot\exp(\gamma_2 z_1)\cdot\exp(\gamma_3 z_2)-\lambda t^{\gamma_1}\cdot\exp(\gamma_2 z_1)\cdot\exp(\gamma_3 z_2)\right]$

1. Formulas are expressed on the proportional hazard metric; 2. Where λ is the exponential of the linear predictor of covariates in the survival regression model, i.e. $\lambda=\exp(-(\beta_0+\beta_1 x_1+\ldots+\beta_n x_n))$; 3. Where $\gamma=\exp(-\beta_0)$; 4. It is not obvious how to write the Hazard function in closed form, since the c.d.f. involves an incomplete gamma integral. 5. Using restricted cubic splines, with 2 knots.

Table 2: Competing survival models against a proposed wish list of desirable characteristics for use in cost-effectiveness models

	Extrapolation	Explicit baseline	Metric	Flexible hazard	Distribution	Off-the-shelf
One parameter						
Exponential	Yes	Very Rigid	PH, AFT	Monotone	No	Yes
Two parameters						
Weibull	Yes	Rigid	PH, AFT	Monotone	No	Yes
Gompertz	Yes	Rigid	PH	Monotone	No	Yes
Log-logistic	Yes	Rigid	Not PH	Non-monotone	No	Yes
Log-Normal	Yes	Rigid	Not PH	Non-monotone	No	Yes
Three parameters						
G. Gamma	Yes	Flexible	As above	Non-monotone	Yes	Yes
Others						
Cox PH	No	Unspecified	PH	Non-monotone	Yes	Yes
Royston-Parmar	Yes	Very Flexible	PH, PO	Non-monotone	Yes	STATA

Table 3: Model fit

Model	Deviance
Weibull model (original framework) PH	**1715.4**
Exponential model PH	**1790.1**
Gompertz model PH	**1762.5**
Generalized Gamma model AFT	**1714.0**
Royston-Parmar model PH	**1708.3**

Note: This table compares the competing SRMs in terms of their deviance statistic. The deviance is a measure of model fit, indicating how well the model represents the observed data (lower deviance = better/more efficient model).