Original Article

MDM
Medical Decision Making

Medical Decision Making
2019, Vol. 39(1) 57–73

DOI: 10.1177/0272989X18814770
journals.sagepub.com/home/mdm
SAGE

Comparing Strategies for Modeling Competing Risks in Discrete-Event Simulations: A Simulation Study and Illustration in Colorectal Cancer

Koen Degeling, Hendrik Koffijberg, Mira D. Franken, Miriam Koopman, and Maarten J. IJzerman

Abstract

Background. Different strategies toward implementing competing risks in discrete-event simulation (DES) models are available. This study aims to provide recommendations regarding modeling approaches that can be defined based on these strategies by performing a quantitative comparison of alternative modeling approaches. **Methods.** Four modeling approaches were defined: 1) event-specific distribution (ESD), 2) event-specific probability and distribution (ESPD), 3) unimodal joint distribution and regression model (UDR), and 4) multimodal joint distribution and regression model (MDR). Each modeling approach was applied to uncensored individual patient data in a simulation study and a case study in colorectal cancer. Their performance was assessed in terms of relative event incidence difference, relative absolute event incidence difference, and relative entropy of time-to-event distributions. Differences in health economic outcomes were also illustrated for the case study. **Results.** In the simulation study, the ESPD and MDR approaches outperformed the ESD and UDR approaches, in terms of both event incidence differences and relative entropy. Disease pathway and data characteristics, such as the number of competing risks and overlap between competing time-to-event distributions, substantially affected the approaches' performance. Although no considerable differences in health economic outcomes were observed, the case study showed that the ESPD approach was most sensitive to low event rates, which negatively affected performance. **Conclusions.** Based on overall performance, the recommended modeling approach for implementing competing risks in DES models is the MDR approach, which is defined according to the general strategy of selecting the time-to-event first and the corresponding event second. The ESPD approach is a less complex and equally performing alternative if sufficient observations are available for each competing event (i.e., the internal validity shows appropriate data representation).

Keywords

competing events, competing risks, discrete event simulation, individual patient data, survival analysis

Date received: December 19, 2017; accepted: October 30, 2018

The use of discrete-event simulation (DES) for evaluating health care interventions is expected to increase due to the complexity of many novel personalized treatment options.[1] In particular its ability to model dynamic pathways based on patient-level histories and patient characteristics makes DES particularly useful for representing personalized treatment processes.[2,3] Besides dynamic model structures and flexibility toward defining different levels of abstraction throughout a model, DES embodies alternative approaches for modeling the occurrence of

Corresponding Author

Maarten J. IJzerman, Health Technology and Services Research Department, Technical Medical Centre, University of Twente, P.O. Box 217, Enschede, 7500 AE, The Netherlands (m.j.ijzerman@utwente.nl).

events. In contrast to more commonly applied patient-level discrete-time state transition models (STMs),[4] DES does not require time to be defined by fixed discrete cycles but allows for parametric distributions to be directly implemented to represent patient-level time-to-event variation. Uncertainty in parametric distributions' parameters can be accounted for in probabilistic sensitivity analyses, so that both stochastic uncertainty (i.e., first-order uncertainty) and parameter uncertainty (i.e., second-order uncertainty) are reflected.[5] Although parametric distributions can also be used to populate STMs, this requires an additional discretization step, that is, evaluation of the cumulative density functions at fixed time points, to obtain discrete-time transition probabilities.

Health economic models that are informed with individual patient data (IPD) and employ DES for patient-level simulations are subject to several design choices,[6] one of which relates to the approach taken to implement competing risks. Since competing risks are present in every clinical pathway and may affect the observation of outcomes in clinical studies, it is important to appropriately represent competing risks in health economic models.[7] Modelers are provided with a high degree of flexibility with regard to selecting one or multiple strategies for implementing competing risks in DES models. According to Barton et al.,[8] 4 general strategies are available: 1) simulating times for all competing events and selecting the event that is the first to occur, 2) selecting the event to occur first and the corresponding time-to-event second, 3) selecting the time-to-event first and the corresponding event second, and 4) using discretized cyclic probabilities to resemble discrete-time STM.

According to the Professional Society for Health Economics and Outcomes Research (ISPOR) and Society for Medical Decision Making (SMDM) modeling good research practices guidelines, the recommended strategy toward implementing competing risks in DES models is Strategy 3 as proposed by Barton et al.[8]: selecting the time-to-event first, based on a joint time-to-event distribution, and then selecting the corresponding event.[6] However, a thorough quantitative comparison of the strategies available has not yet been performed, which might increase variation in approaches taken to represent competing risks in published health economic DES models. For example, recently published studies used approaches based on the strategy of using discretized cyclic probabilities,[9,10] selecting the event first and the time-to-event second,[11,12] or selecting the event that is the first to occur.[13–15] Moreover, the specification and motivation of the approach taken to implement competing risks are not always made explicit in modeling publications.[16,17] This may be due to space limitations or, alternatively, due to limited awareness of the availability of different approaches.

The objective of this study is to describe, illustrate, and compare different approaches for handling competing risks in DES models informed by uncensored IPD. The comparison will ideally lead to generalized recommendations, so modelers can make informed and deliberative decisions regarding the handling of competing risk data in DES models. To achieve this objective, approaches are compared in a simulation study to assess their accuracy in representing the incidence of competing events and corresponding event-specific time-to-event distributions, in terms of bias and relative entropy, respectively. Subsequently, the approaches are applied in a case study based on uncensored patient-level data obtained from the randomized controlled CAIRO3 trial[18] to illustrate potential differences in health economic outcomes of a cost-effectiveness analysis in colorectal cancer.

Methods

We focus on 3 general strategies for handling competing risks identified by Barton et al.[8]: 1) select the event that is the first to occur, 2) select the event first and the time-to-event second, and 3) select the time-to-event first and the event second. The strategy of using discretized cyclic probabilities was deliberately not included, as this approach would discard major advantages of using DES and resemble a STM. Moreover, this would require arbitrary decisions about the cycle length and time dependency of transition probabilities, creating the need to also assess the impact of these decisions on simulation outcomes. Based on the 3 included strategies, 4 specific modeling approaches for implementing competing risks in DES models informed by uncensored IPD were defined, which are described below in more detail. Pseudo-algorithms for data analysis and simulation according to these modeling approaches are provided in

Health Technology and Services Research Department, Technical Medical Centre, University of Twente, Enschede, The Netherlands (KD, HK, MJI); Department of Medical Oncology, University Medical Centre, Utrecht University, Utrecht, The Netherlands (MDF, MK); and Cancer Health Services Research Unit, School of Population and Global Health, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Australia (MJI); Victorian Comprehensive Cancer Centre, Melbourne, Australia (MJI).
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) received no financial support for the research, authorship, and/or publication of this article.

Box 1 Event-Specific Distribution Approach

Data Analysis (see Table 1)
1.1 For each competing event e, fit a time-to-event distribution D_e:
- Observations of patients are censored (C) at the time a competing event occurs
- Select a distribution type to represent the time-to-event T_e for competing event e
- Estimate n parameters $\beta_{e1}, \ldots, \beta_{en}$ that define distribution D_e

Simulation
1.2 Obtain time-to-events for each competing event:
- Draw a time t_e for each competing event e by performing a random draw from the corresponding distribution D_e

1.3 Select the competing event to occur:
- Select event k with the lowest time-to-event across all events (i.e., the first event to occur)

1.4 Simulate the selected event k at the corresponding time t_k

Box 2 Event-Specific Probability and Distribution Approach

Data Analysis (see Table 1)
2.1 For each competing event e, estimate the probability p_e this event occurs:
- Estimate probability p_e as the proportion of patients experiencing competing event e

2.2 For each competing risk e, fit a time-to-event distribution D_e:
- Only include observations of patients who experienced event e
- Select a distribution type to represent the time-to-event T_e for competing event e
- Estimate n parameters $\beta_{e1}, \ldots, \beta_{en}$ that define distribution D_e

Simulation
2.3 Select the competing event to occur:
- Draw a random number r from a Uniform distribution $U[0,1]$
- Select the event k to occur by comparing event probabilities p_e and random number r

2.4 Obtain a time-to-event for the selected event:
- Draw a time t for selected event k by performing a random draw from the corresponding distribution D_k

2.5 Simulate the selected event k at the corresponding time t

Box 3 Unimodal Joint Distribution and Regression Model Approach

Data Analysis (see Table 1)
3.1 For all competing events combined, fit a joint unimodal time-to-event distribution D:
- Include the observations of all patients
- Select a unimodal distribution to represent the time-to-event T for all competing events
- Estimate n parameters $\beta_1, \ldots, \beta_n$ that define distribution D

3.2 Fit a (multinomial) logistic regression model f to predict the competing event to occur:
- Estimate the parameters $\beta_1, \ldots, \beta_n$ of (multinomial) logistic regression model f that predicts the probabilities p_e of each competing event e to occur (dependent variable) based on the time-to-event t (independent variable)

Simulation
3.3 Obtain a time-to-event for the event to occur:
- Draw a time t for the event to occur by performing a random draw from the joint distribution D

3.4 Select the competing event to occur:
- Obtain probabilities p_e for each competing event to occur based on time-to-event t, using (multinomial) logistic regression model f
- Draw a random number r from a Uniform distribution $U[0,1]$
- Select the event k to occur by comparing event probabilities p_e and random number r

3.5 Simulate the selected event k at the corresponding time t

Boxes 1, 2, 3, and 4. Additional code illustrating how these approaches can be implemented in R Statistical Software[19] is provided online at www.personex.nl/research/competing-risks.

Modeling Approach 1: ESD

The ESD approach implements a cause-specific hazards model[20,21] and uses event-specific time-to-event distributions to draw random times at which the competing events would occur and subsequently selects the first event to occur to be simulated. To estimate the event-specific distributions, observations of competing events are considered censored observations because those competing events prohibit observing the event of interest.[21] Box 1 reviews how IPD is analyzed according to the ESD modeling approach and how the resulting event-specific time-to-event distributions can be implemented in simulation models. To illustrate this for the illustrative data presented in Table 1, a distribution D_e needs to be fitted for each competing event e by estimating n parameters $\beta_1, \ldots, \beta_n$ that define D_e, resulting in separate distributions for both the time-to-death and time-to-progression. When estimating the time-to-death distribution D_{Death}, the 3 patients who died (i.e., Patient 1, Patient 5, and Patient 6) are considered observations and those who progressed are considered right-censored observations (C), as these patients would have died from the initial state at some point in time if they would not have progressed.[20] In the corresponding simulation model, a time t_e to each of the competing events e needs to be drawn randomly from each time-to-event distribution D_e. Subsequently, the event that is the first to occur (i.e., the event e corresponding to the lowest drawn time-to-event t_e) is selected and will be simulated.

Modeling Approach 2: ESPD

The ESPD approach implements a subdistribution hazards model[20,21] and first selects the event to be simulated based on event-specific probabilities, and it subsequently draws the time at which that event will occur from an event-specific time-to-event distribution. Box 2 reviews how IPD is analyzed according to the ESPD modeling approach and how the resulting event-specific probabilities and time-to-event distributions can be implemented in simulation models. Event-specific probabilities p_e are defined as the cumulative incidence function limit,[20] simply representing the probability that corresponding competing event e occurs, which is 3 out of 9 for death in the exemplary data of Table 1. Notice that probabilities p_e of competing events always add up to 1 over the lifetime of a patient. The event-specific time-to-event distributions D_e are estimated solely based on observations of patients who experienced competing event of interest e. Considering the data in Table 1, this indicates that the time-to-death distribution D_{Death} is estimated based on the time-to-events of the 3 patients who died (i.e., Patient 1, Patient 5, and Patient 6). In the corresponding simulation model, a random number r needs to be compared to the event-specific probabilities p_e to select the event k that a hypothetical patient will experience. Subsequently, a time-to-event t needs to be drawn randomly from distribution D_k corresponding to the selected event k.

Modeling Approach 3: UDR

The UDR approach first selects the time at which an event will be simulated based on a joint time-to-event distribution, representing all competing events, and selects the event corresponding to the selected time-to-event using a (multinomial) logistic regression model. Multinomial logistic regression models are required for cases involving more than 2 competing risks, as standard logistic regression models can only account for binary data (i.e., 2 competing risks). When estimating the joint time-to-event distribution, the UMR approach assumes the joint time-to-event distribution to be unimodal. Box 3 reviews how IPD is analyzed according to the UDR modeling approach and how the resulting time-to-event distribution and (multinomial) logistic regression model can be implemented in simulation models. For the exemplary data presented in Table 1, this indicates that to represent time-to-event distribution D, for example, a single Weibull distribution[22] is estimated based on all time-to-event observations. Next, logistic regression model f,[23] predicting the type of event (dependent variable), needs to be estimated based on the time-to-event (independent variable). In the corresponding simulation model, a random time-to-event t needs to be drawn from the joint distribution D, which then can be used together with a random number r to select the corresponding event k to occur using probabilities p_e obtained from the logistic regression model f.

Modeling Approach 4: MDR

The MDR approach is similar to the UDR approach, except for the fact that the MDR approach does not assume the time-to-event data to be unimodal distributed but allows the joint time-to-event distribution MD to be multimodal. This implies that, for example, a phase type[24] or mixture[25] distribution can be estimated to represent the patient-level variation in time-to-event values. Box 4 reviews how IPD is analyzed according to the MDR modeling approach and how the resulting time-to-event distribution and (multinomial) logistic regression model can be implemented in simulation models. Except for the type of distribution used to represent the time-to-event data, the data analysis and simulation processes

Table 1 Hypothetical Patient-Level Time-to-Event Data Used to Illustrate the Modeling Approaches Included in This Study[a]

Observed Data			ESD Approach		ESPD Approach				UDR Approach		MDR Approach	
Patient ID	Time-to-Event	Observed Event	Time to Death	Time to Progression	Event Is Death	Event Is Progression	Time to Death	Time to Progression	Time-to-Event	Observed Event	Time-to-Event	Observed Event
1	154	Death	154	C	Yes	No	154	—	154	Death	154	Death
2	351	Progression	C	351	No	Yes	—	351	351	Progression	351	Progression
3	294	Progression	C	294	No	Yes	—	294	294	Progression	294	Progression
4	310	Progression	C	310	No	Yes	—	310	310	Progression	310	Progression
5	257	Death	257	C	Yes	No	257	—	257	Death	257	Death
6	211	Death	211	C	Yes	No	211	—	211	Death	211	Death
7	438	Progression	C	438	No	Yes	—	438	438	Progression	438	Progression
8	80	Progression	C	80	No	Yes	—	80	80	Progression	80	Progression
9	347	Progression	C	347	No	Yes	—	347	347	Progression	347	Progression
Example of data analysis output			$T_{Death} \sim D(\beta_{Death1}, \beta_{Death2})$	$T_{Progression} \sim D(\beta_{Progression1}, \beta_{Progression2})$	P(Death)	P(Progression)	$T \mid Death \sim D(\beta_{Death1}, \beta_{Death2})$	$T \mid Progression \sim D(\beta_{Death1}, \beta_{Death2})$	$T \sim D(\beta_1, \beta_2)$	Event $\mid$ T ~ $f(\beta_1, \ldots, \beta_n)$	$T \sim MD(\beta_1, \ldots, \beta_n)$	Event $\mid$ T ~ $f(\beta_1, \ldots, \beta_n)$

ESD, event-specific distribution; ESPD, event-specific probability and distribution; MDR, multimodal joint distribution and regression model; UDR, unimodal joint distribution and regression model.

a. The data concern 9 hypothetical time-to-event observations of patients who were subject to 2 competing risks: progression and death. C = censored at the time of observing the competing event; T represents a random variable to describe a time-to-event; D represents a certain unimodal time-to-event distribution (e.g., Weibull or Gamma) defined by parameters β_1 and β_2; P(. . .) represents the probability that the corresponding competing event will occur; *f* represents a (multinomial) regression model defined by parameters $\beta_1, \ldots, \beta_n$ predicting the probability of the competing events to occur (dependent variable) based on the time-to-event T (independent variable); MD represents a multimodal distribution defined by n parameters $\beta_1, \ldots, \beta_n$.

Box 4 Multimodal Joint Distribution and Regression Model Approach

Data Analysis (see Table 1)
4.1 For all competing events combined, fit a joint multimodal time-to-event distribution *MD*:
- Include the observations of all patients
- Select a multimodal distribution to represent the time-to-event T for all competing events
- Estimate n parameters $\beta_1, \ldots, \beta_n$ that define distribution *MD*

4.2 Fit a (multinomial) logistic regression model f to predict the competing event to occur:
- Estimate the parameters $\beta_1, \ldots, \beta_n$ of (multinomial) logistic regression model f that predicts the probabilities p_e of each competing event e to occur (dependent variable) based on the time-to-event t (independent variable)

Simulation
4.3 Obtain a time-to-event for the event to occur:
- Draw a time t for the event to occur by performing a random draw from the joint distribution *MD*

4.4 Select the competing event to occur:
- Obtain probabilities p_e for each competing event to occur based on time-to-event t, using (multinomial) logistic regression model f
- Draw a random number r from a Uniform distribution $U[0,1]$
- Select the event k to occur by comparing event probabilities p_e and random number r

4.5 Simulate the selected event k at the corresponding time t

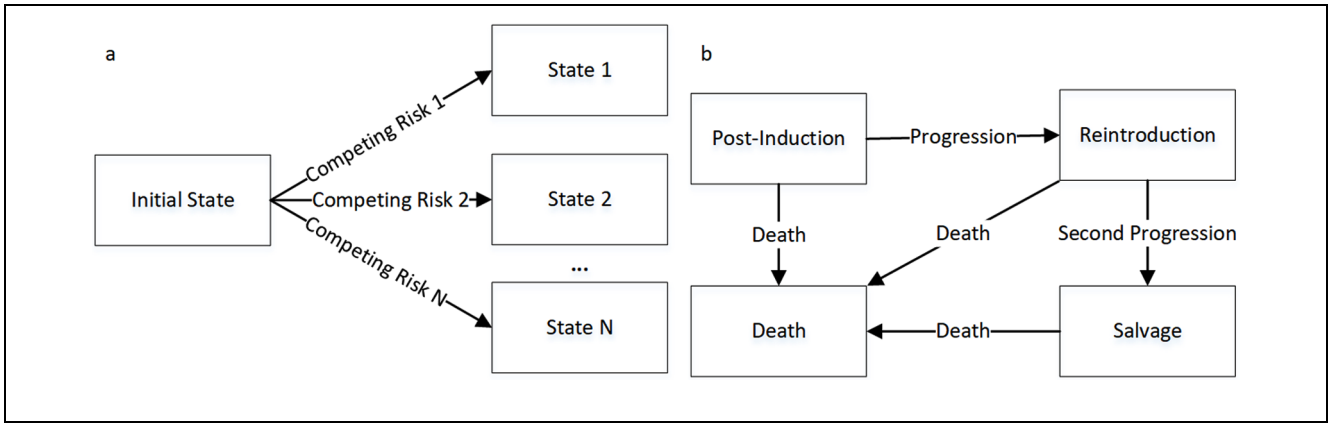

Figure 1 Overview of (a) the structure of the discrete-event simulation (DES) model used in the simulation study and (b) the structure of the case study DES model.

are the same as those for the UDR modeling approach illustrated in Box 3.

Simulation Study to Compare the Performance of the Modeling Approaches

A simulation study was performed to compare the accuracy of the approaches and assess whether data and disease pathway characteristics, such as the number of competing risks, affect the performance. This simulation study included the analysis and simulation of the incidence and time-to-event distributions of competing risks from an initial health state (Figure 1a). As illustrated in the simulation study overview presented in Figure 2 and as reviewed in Box 5, the simulation study was run for $i = 9$ different patient populations P_i, which were simulated according to unique combinations of the number of competing risks (i.e., 2, 3, or 4 competing risks) and the degree of overlap between the corresponding competing time-to-event distributions (i.e., low $\sim$10%, medium $\sim$50%, and high $\sim$90% overlap). Figure 3 illustrates the different levels of overlap for a population defined by 3 competing risks. Details on the exact population definitions and calculation of overlap between distributions are provided in Supplementary Materials 1.1 and 1.2, respectively. In addition, the simulation study was performed for different hypothetical trial arm sample sizes n_{sample} (n_{sample} = 50, 100, 200, 500) to assess sample size impact on the performance of the approaches, resulting in a total of 36 unique scenarios.

A total of $j = 10{,}000$ simulation runs were performed for each unique population P_i and sample size n_{sample}

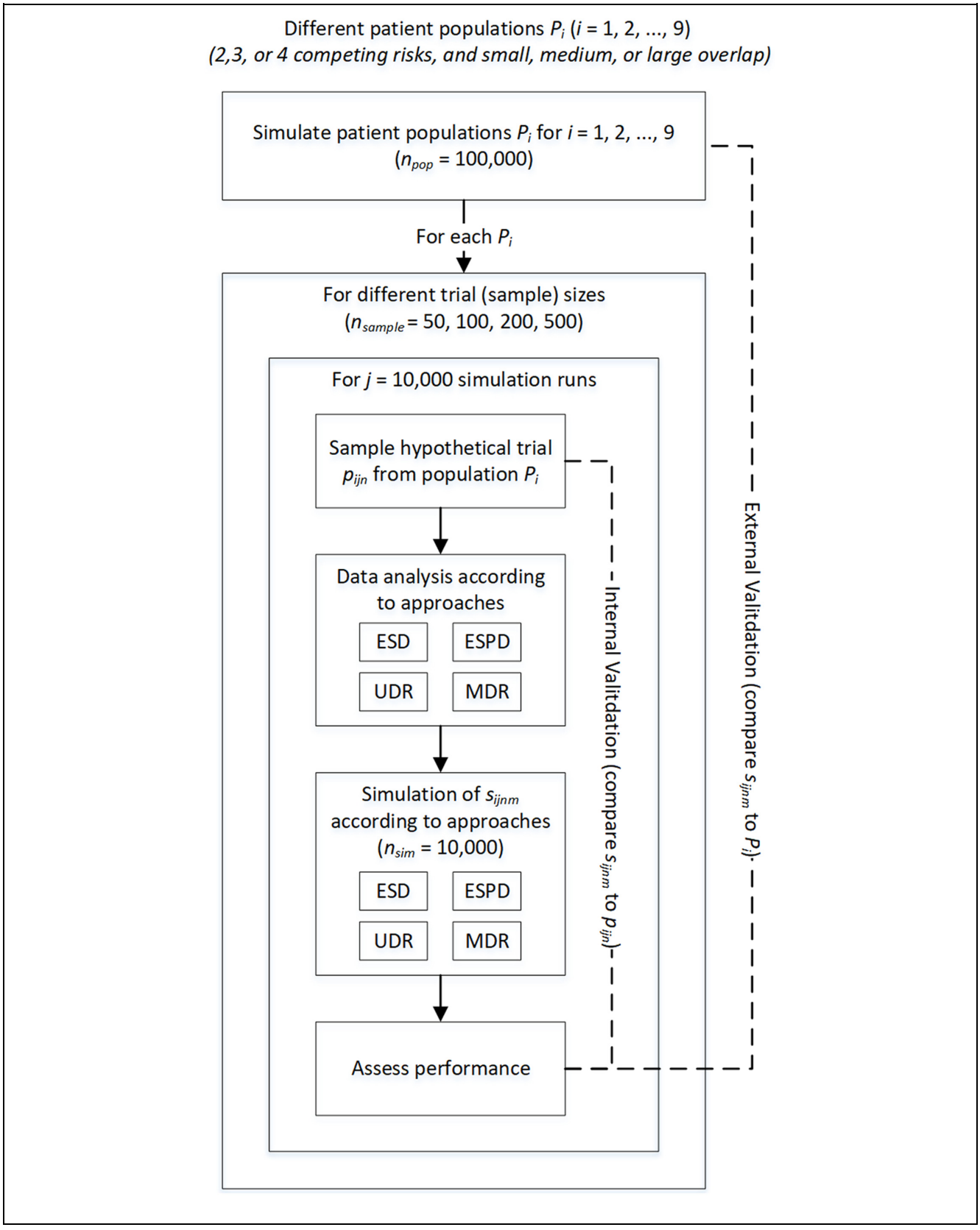

Figure 2 Overview of the simulation study. ESD, event-specific distribution; ESPD, event-specific probability and distribution; MDR, multimodal joint distribution and regression model; UDR, unimodal joint distribution and regression model.

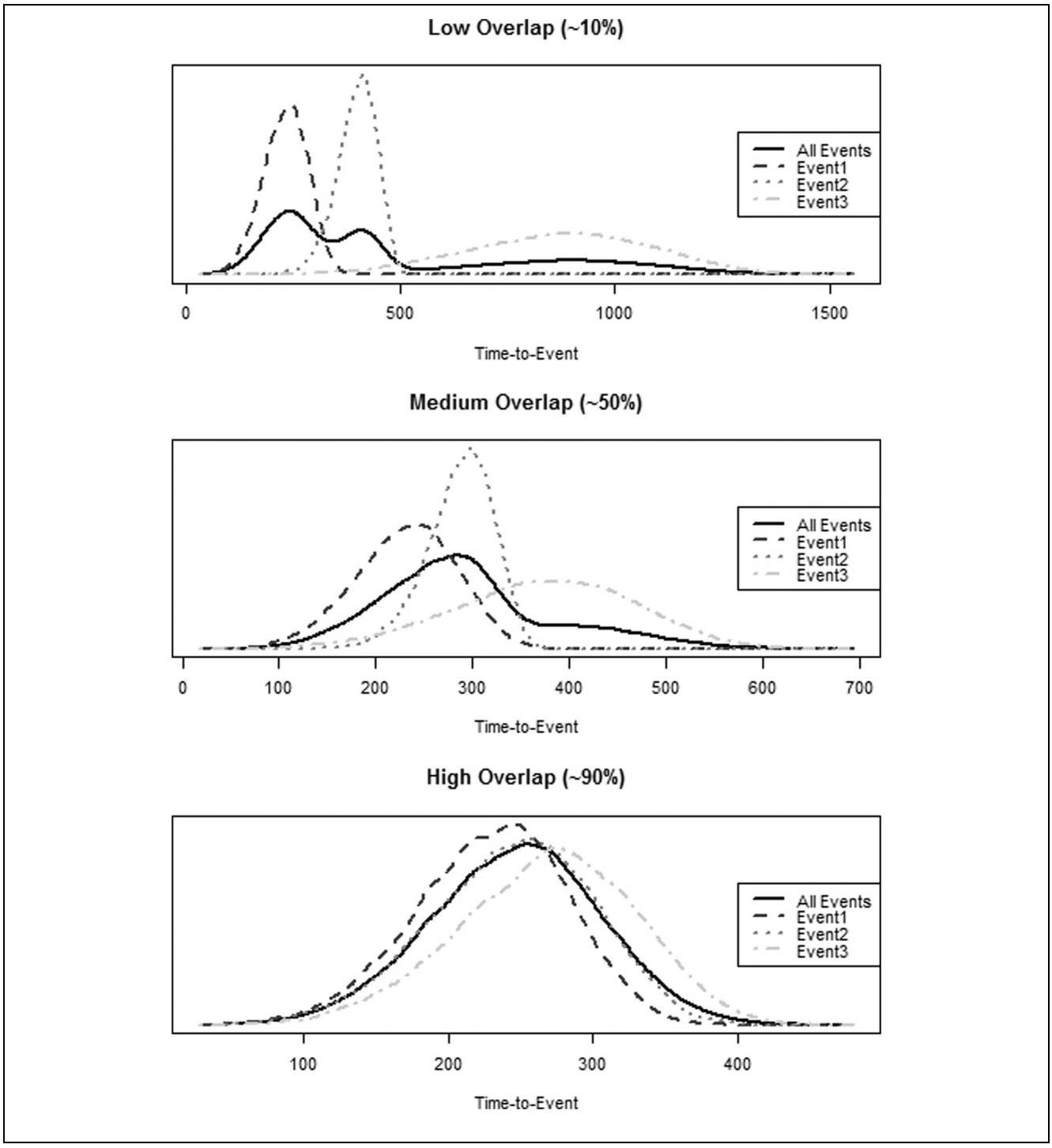

Figure 3 Illustration of the different levels of overlap between competing time-to-event distributions used in the simulation study.

combination. In each of these runs, a hypothetical trial arm sample p_{ijn} of the applicable sample size n_{sample} was randomly sampled from the corresponding population P_i. Next, the hypothetical trial sample p_{ijn} was analyzed according to the $m = 1, 2, 3, 4$ modeling approaches, based on which the incidence and time-to-event

distributions were simulated for $n_{sim} = 10{,}000$ new patients to obtain simulation sample s_{ijnm}. Finally, the performance of the approaches was assessed by comparing the event incidence and time-to-event distributions in these newly simulated patients s_{ijnm} to those in the population P_i (i.e., external validation) and hypothetical trial sampled from this population p_{ijn} (i.e., internal validation). Regarding the incidence of events, the bias in terms of relative incidence difference (%) and relative absolute incidence difference (%) of the approaches in s_{ijnm} compared to the population P_i and trial sample p_{ijn} were assessed. The performance with regard to the simulated time-to-event distributions was obtained by comparing the simulated event-specific distributions in s_{ijnm} to those of the population P_i and trial sample p_{ijn} based on the relative entropy (i.e., the Kullback-Leibler divergence).[26] The relative entropy is a measure of the difference between 2 probability distributions, for which lower values indicate a better performance. To summarize the relative absolute incidence difference and Kullback-Leibler divergence for each of the 36 scenarios (i.e., combinations of P_i and n_{sample}), event-specific performance outcomes were weighted according to event incidences in the population. Event-specific relative incidence differences were not weighted to obtain summarized performance measures, because weighing relative incidence differences based on the incidence does not result in meaningful outcomes (i.e., outcomes of approximately zero).

The simulation study was performed in R Statistical Software version 3.3.2.[19] All time-to-event data were simulated and analyzed using Weibull distributions[22] to rule out potential bias due to mismatching distributions. Weibull distributions were selected to represent patient-level time-to-event variation in the simulation study, because these distributions are commonly used in survival analysis and accurately represent the IPD of the case study. Univariate Weibull distributions were estimated using the *fitdistrplus* package[27] for the ESD, ESPD, and UDR approaches. For the MDR approach, the *mixtools* package[28] was used to estimate Weibull mixture distributions, providing the parameter estimates of the ESPD approach as starting values to increase the likelihood of convergence. If the algorithm for estimating a mixture distribution did not converge in a specific simulation run, the parameter estimates of the ESPD approach were used to define the corresponding mixture distribution. The *nnet*[29] package was used to estimate (multinomial) logistic regression models. The Kullback-Leibler divergence was determined using the *flexmix* package.[30–32]

Illustration of Competing Risks Modeling Approaches in Colorectal Cancer

To assess the modeling approaches' potential performance and impact on health economic outcomes in real-word scenarios, a case study was performed based on an anonymized data set from the randomized phase 3 CAIRO3 study (NCT00442637) of the Dutch Colorectal Cancer Group. The CAIRO3 study randomized 558 metastatic colorectal cancer patients with stable disease or better after 6 cycles of capecitabine, oxaliplatin, and bevacizumab (CAPOX-B) induction therapy to either capecitabine and bevacizumab (CAP-B) maintenance treatment (intervention) or observation (control) until progression of disease.[18] For both the maintenance and observation strategy, CAPOX-B treatment was to be reinduced upon progression and continued until second progression (PFS2), the primary end point of the study. The original health economic evaluation of the CAIRO3 study was based on a discrete-time cohort STM and has been published elsewhere.[33]

A previously developed DES model[5] was implemented in R Statistical Software version 3.3.2,[19] according to the structure of the CAIRO3 study: postinduction, reintroduction, salvage, and death (Figure 1b). Model state postinduction refers to observation (control) or CAP-B maintenance treatment (intervention) after 6 cycles of CAPOX-B. Reintroduction of CAPOX-B refers to modeling state reintroduction for both treatment strategies (i.e., observation and CAP-B maintenance treatment). If patients progress from the reintroduction state (i.e., the cancer no [longer] responds to reintroduction of CAPOX-B), salvage therapy (i.e., alternative treatment options) is provided, which refers to the salvage state. As illustrated in Figure 1b, patients were subject to competing risks progression and death while in the postinduction state and to second progression and death in the reintroduction state. The original DES model handled these competing risks according to the ESPD approach; here, 3 alternative versions of the model were created to reflect the other approaches (i.e., ESD, UDR, and MDR). For the case study, all modeling approaches were implemented using the same R packages as were used for the simulation study.

Stochastic uncertainty (i.e., patient-level variation) in health state durations was accounted for by using Weibull (mixture) distributions. Parameter uncertainty in probabilities, parametric distributions, and regression models used to reflect time-to-event evidence according to the 4 approaches was accounted for by bootstrapping the CAIRO3 data set,[5] averting potential bias by applying all approaches to the exact same bootstrap samples.

Box 5 Simulation Study (Also See Figure 2)

5.1 Simulate different patient populations according to the number of competing risks and time-to-event distribution overlap:
- Define patient population i as P_i for $i = 1, 2, \ldots, 9$
- Each P_i represents a combination of 2, 3, or 4 competing risks and a small, medium, or large overlap of time-to-event distributions (see Figure 3)
- Simulate $n_{pop} = 100{,}000$ hypothetical patients to define P_i according to the number of competing risks and time-to-event distribution overlap

5.2 For each population P_i ($i = 1, 2, \ldots, 9$):
5.3 For different trial sizes $n_{sample} = 50, 100, 200, 500$:
5.4 For $j = 10{,}000$ simulation runs:
5.5 Sample a hypothetical trial of size n_{sample} from patient population P_i:
5 • Define p_{ijn} as the jth trial sample of size n_{sample} of population P_i
5.6 Analyze p_{ijn} according to each modeling approach (see Boxes 1, 2, 3, and 4)
5.7 Simulate a sample of size $n_{sim} = 10{,}000$ patients according to each modeling approach (see Boxes 1, 2, 3, and 4):
- Define s_{ijnm} as the jth simulation sample according to modeling approach m ($m = 1, 2, 3, 4$) based on a sample p_{ijn} of size n_{sample} and population P_i

5.8 Assess the performance of each modeling approach:
- Calculate the relative incidence difference, relative absolute incidence difference, and relative entropy based on internal and external validation
- Internal validation: compare simulation samples s_{ijnm} for $m = 1, 2, 3, 4$ to trial sample p_{ijn}
- External validation: compare simulation samples s_{ijnm} for $m = 1, 2, 3, 4$ to population P_i

For the other model parameters, parameter uncertainty was accounted for as in the original health economic evaluation of the CAIRO3 study.[33] Probabilistic sensitivity analysis was performed based on 10,000 runs of 10,000 patients per treatment strategy in each run.

Clinical relevant subgroup analyses were performed to illustrate potential sample size impact on modeling outcomes for the different approaches. A total of 8 subgroups with sample sizes ranging from 50 to 410 were defined according to patient characteristics that were found relevant in the evaluation of the CAIRO3 study,[18] that is, treatment response (stable disease [SD] v. complete or partial response [CR/PR]) and stage of disease (synchronous v. metachronous) (Supplementary Materials 1.3). This stratification strategy resulted in subgroups for which events were observed only once, or not all, which resembles the personalized medicine context.

For all subgroups, the accuracy of the approaches was compared based on the internal validity in terms of relative incidence difference (%), relative absolute incidence difference (%), and Kullback-Leibler divergence over all probabilistic sensitivity analysis runs. To obtain these performance measures, simulations according to the approaches were compared to the bootstrap sample based on which of the approaches' parameters had been estimated. Incremental cost-effectiveness planes and cost-effectiveness acceptability curves presented the health economic outcomes.

Results

Simulation Study to Compare the Performance of the Modeling Approaches

Mean weighted results for the relative absolute incidence difference and relative entropy over all runs of the simulation study are presented in Table 2. Detailed event-specific results for the relative absolute incidence difference and relative incidence difference are not presented in this article to enhance readability but are available in Supplementary Materials 2.

The bias in terms of relative incidence difference is substantially better for the ESPD and MDR approaches compared to the ESD and UDR approaches. Using the ESD or UDR approach results in a substantial amount of bias for higher numbers of competing risks and lower overlap between their corresponding time-to-event distributions. For example, for the population defined by 4 competing risks and low overlapping distributions, the ESD and UDR approaches yield up to approximately 60% and 50% bias, respectively, whereas the ESPD and MDR approaches yield up to approximately 5% bias. For all approaches, bias is lower when simulation outcomes are compared to the hypothetical trial (i.e., internal validation) rather than the simulated population (i.e., external validation). Furthermore, overall performance is better for lower numbers of competing risks, higher overlap between time-to-event distributions, and larger sample sizes, although performance of the ESD and UDR

Table 2 Mean Accuracy of the Approaches in Terms of Relative Absolute Incidence Difference and Relative Entropy Over All Simulation Study Runs

			Incidence of Events: Relative Absolute Incidence Difference (%)									Time-to-Event Distributions: Kullback-Leibler Divergence								
			External Validation					Internal Validation				External Validation					Internal Validation			
Number of Events	Distribution Overlap (%)	Sample Size	Trial v. Population	ESD v. Population	ESPD v. Population	UDR v. Population	MDR v. Population	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial	Trial v. Population	ESD v. Population	ESPD v. Population	UDR v. Population	MDR v. Population	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial
2	10	50	11.5	14.2	11.6	12.3	11.5	7.4	0.8	6.5	1.2	0.073	0.237	0.044	0.117	0.064	0.227	0.041	0.115	0.045
2	10	100	7.6	10.2	7.7	8.2	7.6	6.9	0.8	4.3	1.0	0.038	0.224	0.020	0.101	0.029	0.211	0.024	0.093	0.024
2	10	200	5.7	8.5	5.7	6.1	5.7	6.8	0.8	3.4	0.9	0.021	0.218	0.010	0.095	0.014	0.206	0.015	0.085	0.014
2	10	500	3.4	7.5	3.6	4.1	3.6	6.8	0.8	3.0	0.8	0.011	0.215	0.005	0.092	0.007	0.206	0.008	0.084	0.008
2	50	50	10.8	11.0	10.9	11.0	10.8	1.5	0.8	2.0	0.9	0.073	0.035	0.043	0.045	0.054	0.055	0.042	0.061	0.050
2	50	100	7.7	7.9	7.8	8.0	7.8	1.3	0.8	1.7	0.9	0.039	0.021	0.020	0.030	0.025	0.034	0.025	0.041	0.028
2	50	200	5.5	5.7	5.6	5.5	5.5	1.2	0.8	1.5	0.8	0.022	0.014	0.010	0.023	0.013	0.022	0.015	0.030	0.017
2	50	500	3.5	3.7	3.6	3.7	3.5	1.2	0.7	1.4	0.8	0.011	0.010	0.005	0.019	0.007	0.014	0.008	0.023	0.010
2	90	50	10.9	10.9	10.9	10.9	10.8	0.9	0.8	0.8	0.8	0.069	0.035	0.045	0.034	0.045	0.049	0.038	0.048	0.048
2	90	100	7.8	7.8	7.8	7.8	7.8	0.8	0.8	0.8	0.8	0.037	0.019	0.021	0.018	0.021	0.029	0.023	0.028	0.028
2	90	200	5.7	5.7	5.7	5.7	5.7	0.8	0.8	0.8	0.8	0.020	0.012	0.010	0.010	0.011	0.019	0.014	0.018	0.017
2	90	500	3.5	3.6	3.6	3.5	3.5	0.8	0.8	0.8	0.8	0.010	0.007	0.005	0.006	0.006	0.012	0.008	0.010	0.010
3	10	50	16.0	19.5	16.0	21.6	16.0	9.8	1.2	22.0	1.8	0.095	0.623	0.062	0.224	0.092	0.616	0.048	0.228	0.054
3	10	100	10.9	14.7	11.0	20.5	11.0	9.6	1.1	21.2	1.4	0.047	0.611	0.028	0.206	0.043	0.593	0.028	0.202	0.030
3	10	200	8.0	11.9	8.1	20.0	8.0	9.5	1.1	20.1	1.2	0.026	0.607	0.013	0.198	0.021	0.586	0.017	0.189	0.018
3	10	500	5.0	10.1	5.1	18.8	5.1	9.3	1.1	19.0	1.2	0.012	0.605	0.006	0.197	0.011	0.587	0.009	0.187	0.010
3	50	50	14.9	16.1	15.0	14.7	14.9	4.1	1.1	7.3	1.4	0.104	0.304	0.068	0.153	0.124	0.312	0.054	0.173	0.107
3	50	100	11.6	12.6	11.7	12.1	11.7	3.9	1.2	6.9	1.2	0.053	0.292	0.030	0.135	0.081	0.288	0.032	0.145	0.077
3	50	200	8.1	9.0	8.2	9.5	8.2	3.8	1.1	6.6	1.2	0.029	0.288	0.015	0.127	0.065	0.277	0.019	0.128	0.063
3	50	500	4.9	6.0	5.0	7.4	5.1	3.7	1.1	6.5	1.1	0.014	0.285	0.007	0.123	0.057	0.275	0.010	0.123	0.056
3	90	50	15.0	15.1	15.1	15.1	15.2	1.4	1.1	1.2	1.2	0.109	0.060	0.070	0.056	0.077	0.085	0.053	0.076	0.077
3	90	100	10.9	11.1	11.0	11.0	11.1	1.3	1.1	1.1	1.1	0.053	0.038	0.031	0.031	0.036	0.054	0.031	0.045	0.043
3	90	200	8.0	8.1	8.1	8.0	8.0	1.2	1.1	1.2	1.1	0.028	0.030	0.015	0.021	0.021	0.039	0.018	0.030	0.028
3	90	500	4.9	5.0	4.9	5.0	5.0	1.2	1.1	1.1	1.1	0.014	0.025	0.007	0.016	0.014	0.029	0.010	0.020	0.018
4	10	50	18.2	22.6	18.3	29.7	18.3	12.3	1.4	30.5	2.1	0.115	0.664	0.084	0.229	0.117	0.658	0.053	0.255	0.057
4	10	100	13.0	18.3	13.2	28.3	13.1	12.1	1.4	28.2	1.7	0.057	0.654	0.035	0.220	0.053	0.629	0.032	0.217	0.033
4	10	200	9.2	15.4	9.3	27.1	9.3	11.8	1.3	26.9	1.5	0.028	0.649	0.016	0.215	0.025	0.624	0.018	0.206	0.018
4	10	500	5.8	13.1	5.9	26.2	5.9	11.7	1.3	25.9	1.4	0.013	0.648	0.007	0.213	0.012	0.627	0.009	0.202	0.011
4	50	50	17.2	19.8	17.3	21.0	17.3	9.8	1.3	17.1	1.7	0.145	0.448	0.102	0.211	0.176	0.452	0.066	0.226	0.118
4	50	100	12.4	15.9	12.6	17.8	12.5	9.1	1.3	15.4	1.5	0.072	0.430	0.043	0.186	0.096	0.419	0.040	0.194	0.081
4	50	200	9.4	13.3	9.5	16.1	9.5	8.9	1.3	14.4	1.4	0.036	0.425	0.020	0.173	0.066	0.406	0.022	0.172	0.060
4	50	500	5.9	10.6	6.0	15.0	6.1	8.7	1.3	14.3	1.4	0.017	0.423	0.009	0.167	0.052	0.406	0.012	0.166	0.050
4	90	50	17.8	17.7	18.0	18.1	17.9	2.8	1.3	2.0	1.4	0.146	0.094	0.103	0.080	0.116	0.128	0.066	0.103	0.102
4	90	100	12.8	12.9	12.9	12.8	12.9	2.2	1.3	1.5	1.4	0.071	0.068	0.043	0.048	0.053	0.084	0.038	0.065	0.058
4	90	200	9.2	9.5	9.3	9.3	9.4	2.0	1.4	1.4	1.3	0.037	0.057	0.021	0.036	0.031	0.064	0.022	0.045	0.038
4	90	500	5.9	6.3	6.1	6.1	6.1	1.9	1.3	1.4	1.3	0.017	0.050	0.009	0.028	0.020	0.053	0.012	0.033	0.024

ESD, event-specific distribution; ESPD, event-specific probability and distribution; MDR, multimodal joint distribution and regression model; UDR, unimodal joint distribution and regression model.

approaches is much more sensitive to changes in these data and disease pathway characteristics.

Although the bias in terms of relative absolute event incidence difference shows less extreme outcomes for the ESD and UDR approaches, these approaches are again outperformed by the ESPD and MDR approaches. Furthermore, overall performance is better with regard to the hypothetical trial (i.e., internal validation) than the simulated population (i.e., external validation). Also, with respect to the bias in terms of relative absolute incidence difference, mainly the ESD and UDR approaches benefit from lower numbers of competing risks, higher overlap between the competing time-to-event distributions, and higher sample sizes.

The performance in terms of relative entropy shows the same trends in differences between approaches. The ESPD and MDR approaches strongly outperform their ESD and UDR equivalents, although the Kullback-Leibler divergence also shows that the ESPD approach slightly outperforms the MDR approach. Once more, the performance in terms of internal validity is better compared to the external validity, and especially the performance of the ESD and UDR approaches benefits from lower numbers of competing risks, higher overlap between the competing time-to-event distributions, and higher sample sizes.

Illustration of Competing Risks Modeling Approaches in Colorectal Cancer

Mean results for the relative incidence difference and relative entropy for the case study over all runs of the probabilistic sensitivity analysis are presented in Table 3. Results for the relative absolute incidence difference are not presented in the article but are available in Supplementary Materials 1.4. The relative absolute incidence differences showed negligible differences compared to the absolute value of the relative incidence differences, indicating that the approaches underestimated or overestimated the event incidences systematically.

The internal validation of the approaches shows similar trends for the case study as for the simulation study. The ESPD and MDR approaches overall yield slightly better relative incidence differences and relative absolute incidence differences. For example, for the cohort analysis (Subgroup 0), the mean relative incidence difference in the probability of progression from the postinduction state for the control group is 1.9%, 0.0%, 0.9%, and 0.4% for the ESD, ESPD, UDR, and MDR approaches, respectively. Interestingly, the results also suggest that the performance in terms of bias for the ESPD approach is more sensitive to low event rates compared to the other approaches, which is illustrated by the mean relative incidence differences for the intervention group of Subgroup 6. Only 2 of 17 patients in this subgroup died during the reintroduction state (Supplementary Materials 1.3), resulting in a mean relative incidence difference in the probability of progression from the reintroduction state of 0.1%, 5.1%, 0.0%, and 0.1% for the ESD, ESPD, UDR, and MDR approaches, respectively. With regard to the relative entropy, the ESPD and MDR approaches generally outperformed the ESD and UDR approaches. Comparing the ESPD and MDR approaches, they both alternately outperformed the other, making it difficult to state which is the best-performing approach overall.

The extent to which health economic modeling outcomes are affected by differences in performance is illustrated by cost-effectiveness planes for selected subgroup analyses in Figure 4 and in cost-effectiveness planes and cost-effectiveness acceptability curves for all subgroup analyses in Supplementary Materials 1.5 and 1.6, respectively. The cost-effectiveness planes show that cost-effectiveness point estimates are similar for large sample sizes (e.g., $n = 300$ or larger), illustrated by the overlapping points representing the incremental cost-effectiveness estimate for the different approaches. For analysis Subgroup 0 (i.e., based on the complete patient cohort), incremental cost-effectiveness ratios are €177,317, €166,997, €172,811, and €169,024 per quality-adjusted life year (QALY) gained, which translates to a net monetary benefit of –€26,601, –€26,656, –€26,579, and –€26,503 at a willingness to pay of €20,000 per QALY for the ESD, ESPD, UDR, and MDR approaches, respectively. However, the point estimates of the incremental costs and effects show small differences between approaches for smaller sample sizes, which is illustrated by the results for Subgroup 4 in Figure 4, for example. For this subgroup analysis, incremental cost-effectiveness ratios are €494,517, €342,726, €420,222, and €322,343 per QALY gained, which translates to a net monetary benefit of –€29,854, –€29,966, –€29,195, and –€29,939 at a willingness to pay of €20,000 per QALY for the ESD, ESPD, UDR, and MDR approaches, respectively. With regard to the uncertainty surrounding the point estimates, represented by the confidence ellipses, small differences are observed for certain subgroup analyses (e.g., Subgroup 4). In agreement, the cost-effectiveness acceptability curves show modest differences between approaches (Supplementary Materials 1.6), of which the magnitude increases when sample sizes decrease.

Table 3 Mean Accuracy of the Approaches in Terms of Relative Incidence Difference and Relative Entropy Over All Probabilistic Sensitivity Analysis Runs for the Case Study

					Incidence of Events[a]: Relative Incidence Difference (%)								Time-to-Event Distributions[b]: Kullback-Leibler Divergence															
					Postinduction: Progression				Reintroduction: Progression				Postinduction: Progression				Postinduction: Death				Reintroduction: Progression				Reintroduction: Death			
Subgroup Number	Treatment Response	Stage of Disease	Treatment Strategy[c]	Subgroup Size	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial	ESD v. Trial	ESPD v. Trial	UDR v. Trial	MDR v. Trial
0	—	—	0	279	1.9	0.0	0.9	0.4	–1.7	–1.6	–1.6	–1.6	0.075	0.017	0.066	0.017	0.815	0.112	0.694	0.087	0.027	0.022	0.024	0.018	0.058	0.047	0.068	0.074
			1	279	1.4	0.0	0.2	0.3	–3.2	–3.0	–3.1	–3.0	0.012	0.012	0.011	0.012	0.110	0.040	0.048	0.039	0.075	0.073	0.063	0.056	0.060	0.027	0.043	0.056
1	SD	—	0	95	—	—	—	—	–1.9	–1.7	–1.7	–1.6	0.023	0.023	0.023	0.023	—	—	—	—	0.024	0.033	0.020	0.021	0.150	0.094	0.154	0.172
			1	96	0.2	0.0	0.1	0.1	–3.4	–3.4	–3.5	–3.4	0.027	0.026	0.027	0.027	0.365	0.252	0.366	0.355	0.035	0.057	0.025	0.028	0.155	0.106	0.142	0.147
2	CR/PR	—	0	184	2.6	0.0	0.4	0.4	–1.6	–1.5	–1.5	–1.5	0.097	0.022	0.083	0.022	0.692	0.108	0.628	0.085	0.048	0.037	0.043	0.035	0.110	0.112	0.121	0.116
			1	183	2.0	0.0	0.0	0.4	–3.0	–2.9	–2.9	–2.8	0.018	0.014	0.012	0.014	0.141	0.071	0.085	0.069	0.091	0.086	0.080	0.072	0.084	0.051	0.056	0.069
3	—	Synch.	0	191	1.1	0.1	0.3	0.3	–1.4	–1.2	–1.2	–1.2	0.075	0.022	0.067	0.022	0.732	0.170	0.550	0.246	0.026	0.023	0.024	0.021	0.087	0.069	0.098	0.101
			1	219	1.3	0.0	0.2	0.3	–2.9	–2.8	–2.8	–2.8	0.012	0.009	0.009	0.009	0.127	0.038	0.051	0.039	0.047	0.031	0.038	0.027	0.059	0.029	0.047	0.061
4	—	Meta.	0	88	4.2	0.2	2.9	0.3	–2.6	–2.6	–2.7	–2.6	0.075	0.017	0.065	0.017	0.321	0.145	1.227	0.145	0.042	0.048	0.036	0.025	0.104	0.084	0.091	0.099
			1	59	1.6	0.1	0.2	0.1	–2.1	–2.6	–2.6	–2.6	0.050	0.052	0.052	0.051	0.216	0.155	0.209	0.166	0.183	0.181	0.187	0.191	0.198	0.112	0.180	0.158
5	SD	Synch.	0	67	—	—	—	—	–1.6	–1.4	–1.3	–1.2	0.028	0.028	0.028	0.028	—	—	—	—	0.025	0.034	0.024	0.024	0.319	0.219	0.317	0.358
			1	74	0.3	0.1	0.1	0.1	–3.1	–3.0	–3.1	–3.0	0.021	0.021	0.021	0.021	0.332	0.210	0.306	0.292	0.033	0.050	0.025	0.023	0.237	0.197	0.231	0.242
6	SD	Meta.	0	28	—	—	—	—	–2.6	–2.6	–4.5	–3.4	0.059	0.059	0.059	0.059	—	—	—	—	0.131	0.124	0.106	0.115	0.220	0.200	0.242	0.200
			1	22	0.1	5.1	0.0	0.1	1.5	2.3	–1.0	0.3	0.079	0.075	0.078	0.077	—	—	—	—	0.125	0.149	0.150	0.160	0.280	0.199	0.256	0.214
7	CR/PR	Synch.	0	124	1.5	0.1	0.2	0.4	–1.2	–1.1	–1.1	–1.1	0.102	0.029	0.090	0.030	0.620	0.160	0.472	0.225	0.049	0.048	0.049	0.047	0.162	0.157	0.170	0.166
			1	145	1.9	1.9	0.0	0.0	0.4	–2.8	–2.7	–2.7	0.017	0.015	0.012	0.015	0.147	0.061	0.080	0.059	0.068	0.047	0.058	0.045	0.089	0.056	0.061	0.076
8	CR/PR	Meta.	0	61	5.7	0.3	3.1	0.3	–2.5	–2.6	–2.8	–2.5	0.098	0.026	0.081	0.026	0.227	0.143	1.203	0.143	0.053	0.049	0.044	0.027	0.149	0.092	0.116	0.128
			1	37	2.5	0.2	0.4	0.3	–2.5	–1.9	–3.1	–3.0	0.066	0.057	0.063	0.057	0.239	0.182	0.255	0.185	0.223	0.221	0.231	0.233	0.232	0.141	0.214	0.200

CR/PR, complete or partial response; ESD, event-specific distribution; ESPD, event-specific probability and distribution; MDR, multimodal joint distribution and regression model; Meta., metachronous; SD, stable disease; Synch., synchronous; UDR, unimodal joint distribution and regression model.

a. A long dash (—) indicates that the corresponding probabilities were either 0 or 1 (i.e., one of the competing events was not observed, making a comparison of the relative incidence irrelevant).

b. A long dash (—) indicates that insufficient observations were available in the original trial data to calculate the Kullback-Leibler divergence.

c. Treatment strategy: 0) observation strategy, 1) maintenance treatment.

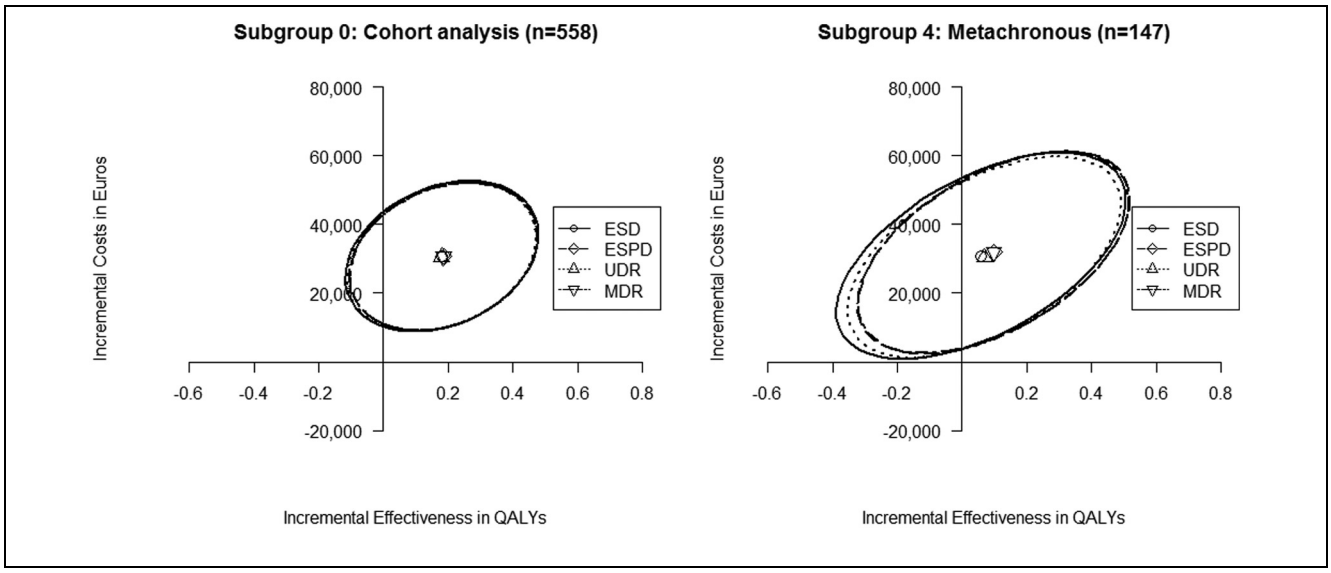

Figure 4 Cost-effectiveness planes based on the probabilistic sensitivity analysis for selected subgroup analyses. ESD, event-specific distribution; ESPD, event-specific probability and distribution; MDR, multimodal joint distribution and regression model; UDR, unimodal joint distribution and regression model.

Discussion

This article provides a thorough quantitative comparison of strategies available for implementing competing risks in DES models informed by uncensored IPD and, thereby, contributes to the general, already existing guidance.[6] According to these general modeling good research practices guidelines, the preferred modeling approach is based on the strategy of selecting the time-to-event first, from a joint time-to-event distribution, and then to select the corresponding event, which corresponds to the UDR and MDR approaches illustrated in this article. Although the results for the UDR approach question this recommendation, the results for the MDR approach, which attempts to reflect the true nature of the data by allowing for multimodal joint distributions, support this recommendation. Interestingly, the results also show that the ESPD approach, which might be considered more straightforward for the novice, provides an attractive alternative to the rather complex MDR approach.

The superior performance of the ESPD and MDR approaches can be explained by the assumptions underlying the ESD and UDR approaches. For the ESD approach, observations of competing events are considered censored observations when estimating the event-specific time-to-event distributions. Given that the analysis of censored time-to-event data is more complex compared to uncensored data, the resulting increase in uncertainty is likely to negatively affect performance. The main assumption for the UDR approach is that the joint time-to-event data, representing all competing risks, is unimodal distributed. Consequently, the (negative) impact of this assumption is illustrated in the simulation study, as the performance of the UDR approach deteriorates stronger compared to the performance of the ESPD and MDR approaches when overlap between time-to-event distributions decreases.

Based on the performance in the simulation study and internal validity in the case study, the ESPD and MDR approaches are expected to yield more accurate health economic outcomes. However, relatively small differences in cost-effectiveness point estimates and magnitude of the uncertainty surrounding these estimates between approaches are observed for the case study, which can be explained by the characteristics of the CAIRO3 data. First of all, the maximum number of competing risks for a specific state in the CAIRO3 model is 2 (Figure 1b). For health economic models in which the number of competing risks is higher, more substantial differences in health economic outcomes are expected, as the simulation study shows that differences in the approaches' performance increase according to higher numbers of competing risks. Second, the CAIRO3 study represents a relatively large sample size (i.e., $n = 558$), which contributes to the overall performance of the approaches. Third, time-to-event distributions for the case study substantially overlap (e.g., 81% for the reintroduction state). Larger differences in health economic outcomes are

expected for models informed by IPD comprising lower overlap between competing time-to-event distributions. Finally, health economic models defined by more health states may yield more substantial differences in health economic outcomes due to a cumulative effect of subsequent competing risks including states.

Based on the performance in the simulation study, either the ESDP or MDR approach would be recommended for implementing competing risks in DES models. However, the case study illustrates that the ESPD approach might be more sensitive to low event rates in practice. Contrarily, if it is infeasible to apply the MDR approach (e.g., due to difficulties with regard to estimating a joint multimodal distribution that appropriately represents the observed IPD), the ESPD approach is a legitimate alternative. Regarding the complexity of implementing the approaches, the flexible multimodal distributions and multinomial regression models required for the MDR approach are more difficult to estimate and implement compared to the event-specific probabilities and distributions of the ESPD approach. Consequently, the ESPD approach would be preferred over its MDR equivalent if the IPD comprises sufficient observations for each competing event. However, a general definition of "sufficient observations" cannot be stated and modelers should, therefore, ensure that applying the ESPD approach is appropriate by internal validation. If the internal validity of the ESPD approach indicates its use might be inappropriate (e.g., due to substantial differences in event incidence), both the ESPD and MDR approaches should be applied and their internal validity compared to select the best-performing modeling approach. If time does not allow for such a comparison, modelers are advised to apply the MDR approach. Except if there are evident reasons for using the ESD or UDR approach, the use of either of these approaches is advised against.

All methods presented in and findings of this study are applicable to data sets in which a competing event is observed for each patient (i.e., the data are uncensored). Although clinical trial data may often be right censored due to limited time horizons, certain clinical trials do capture all (competing) events of interest for each patient. Examples of clinical contexts for which uncensored clinical trials are generally feasible include nausea and vomiting after surgery and metastatic cancers, as was illustrated in the current case study. In addition, retrospective data sets (e.g., from registries) are potentially uncensored for specific cohorts of patients. Performance of the modeling approaches may be different for censored IPD (e.g., because all modeling approaches will include survival analysis of censored data), whereas this only applies to the ESD approach for uncensored IPD or because a different implementation of the ESPD approach is required to calculate event-specific probabilities if a competing event is not observed for each patient. Consequently, the current focus on uncensored IPD is relevant due to a lack of interchangeability of the modeling approaches' implementation and performance for censored and uncensored IPD. Furthermore, this highlights a need for comparing modeling approaches toward implementing competing risks in DES models informed by censored IPD, which is part of a subsequent study.

Besides general limitations relating to the external validity of simulation studies and single-case studies, this study has certain additional limitations. Findings and conclusions presented in this article apply to DES modeling studies informed by uncensored IPD and for which a decision on the approach taken to implement competing risks needs to be made. If aggregated data (e.g., from literature) are used to populate a model, this evidence is structured according to a specific approach and can only be implemented accordingly. In addition, Weibull distributions were used in the simulation study to simulate and represent patient-level time-to-event variation, which allowed unbiased comparison of the modeling approaches. On the other hand, this design choice limits the generalizability of the results, as in practice, underlying distributions are unknown and performance of the approaches may vary, depending on their flexibility to describe the actually observed data. Using different types of distributions, such as Gamma or lognormal distributions, therefore, may affect the performance of the modeling approaches. Finally, mixture models are used to reflect the multimodality induced by competing underlying time-to-event distributions in the joint distributions for the MDR approach, whereas alternative strategies are available, such as phase-type distributions.

Conclusion

Substantial differences were observed in the accuracy of the competing risks modeling approaches in terms of event incidence and time-to-event distributions. The performance of the approaches is strongly conditional upon various data and disease pathway characteristics, such as the numbers of competing risks and overlap between competing time-to-event distributions. Performance dissimilarities between approaches, however, did not result in substantially different health economic outcomes in this specific case study. The recommended modeling

approach for implementing competing risks in DES models informed by uncensored IPD is the MDR approach, which is based on the strategy of selecting the time-to-event first, based on a joint multimodal time-to-event distribution, and then selecting the corresponding event using probabilities obtained from a (multinomial) logistic regression model. However, if sufficient observations of all competing events are available, or if use of the MDR approach is infeasible, use of the less complex ESPD approach, which is based on the strategy of selecting the event to occur first and the corresponding time-to-event second, is also appropriate.

Supplementary Material

Supplementary material for this article is available on the *Medical Decision Making* Web site at http://journals.sagepub.com/home/mdm.

References

1. Caro JJ, Moller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? *Value Health*. 2010;13(8):1056–60.
2. Karnon J, Haji Ali Afzali H. When to use discrete event simulation (DES) for the economic evaluation of health technologies? A review and critique of the costs and benefits of DES. *Pharmacoeconomics*. 2014;32(6):547–58.
3. Standfield L, Comans T, Scuffham P. Markov modeling and discrete event simulation in health care: a systematic comparison. *Int J Technol Assess Health Care*. 2014;30(2):165–72.
4. Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–3. *Med Decis Making*. 2012;32(5):690–700.
5. Degeling K, IJzerman MJ, Koopman M, Koffijberg H. Accounting for parameter uncertainty in the definition of parametric distributions used to describe individual patient variation in health economic models. *BMC Med Res Methodol*. 2017;17(1):170.
6. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–4. *Med Decis Making*. 2012;32(5):701–11.
7. Caro JJ, Moller J. Decision-analytic models: current methodological challenges. *Pharmacoeconomics*. 2014;32(10):943–50.
8. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. *Health Technol Assess*. 2004;8(11):104.
9. Jiang Y, Ni W. Expected lifetime numbers, risks, and burden of osteoporotic fractures for 50-year old Chinese women: a discrete event simulation incorporating FRAX. *J Bone Miner Metab*. 2016;34(6):714–22.
10. Erenay FS, Alagoz O, Banerjee R, Said A, Cima RR. Cost-effectiveness of alternative colonoscopy surveillance strategies to mitigate metachronous colorectal cancer incidence. *Cancer*. 2016;122(16):2560–70.
11. Wang HI, Smith A, Aas E, et al. Treatment cost and life expectancy of diffuse large B-cell lymphoma (DLBCL): a discrete event simulation model on a UK population-based observational cohort. *Eur J Health Econ*. 2017;18(2):255–67.
12. Igarashi A, Goto R, Suwa K, Yoshikawa R, Ward AJ, Moller J. Cost-effectiveness analysis of smoking cessation interventions in Japan using a discrete-event simulation. *Appl Health Econ Health Policy*. 2016;14(1):77–87.
13. Davis S, Martyn-St James M, Sanderson J, et al. A systematic review and economic evaluation of bisphosphonates for the prevention of fragility fractures. *Health Technol Assess*. 2016;20(78):1–406.
14. Standfield LB, Comans TA, Scuffham PA. An empirical comparison of Markov cohort modeling and discrete event simulation in a capacity-constrained health care setting. *Eur J Health Econ*. 2017;18(1):33–47.
15. Thanh NX, Ezekowitz JA, Tran DT, Kaul P. Cost effectiveness of eplerenone for the treatment of systolic heart failure with mild symptoms in Alberta, Canada. *Am J Cardiovasc Drugs*. 2016;16(5):365–76.
16. Bessen T, Keefe DM, Karnon J. Does one size fit all? Cost utility analyses of alternative mammographic follow-up schedules, by risk of recurrence. *Int J Technol Assess Health Care*. 2015;31(5):281–8.
17. Crossan C, Lord J, Ryan R, Nherera L, Marshall T. Cost effectiveness of case-finding strategies for primary prevention of cardiovascular disease: a modelling study. *Br J Gen Pract*. 2017;67(654):e67–e77.
18. Simkens LHJ, van Tinteren H, May A, et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet*. 2015;385(9980):1843–52.
19. R Core Team. R: a language and environment for statistical computing. Version 3.3.2 [software]. Available from: https://www.r-project.org/
20. Pintilie M. *Competing Risks: A Practical Perspective*. Hoboken, NJ: Wiley; 2006.
21. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation*. 2016;133(6):601–9.
22. Law AM. *Simulation Modeling and Analysis*. 4th ed. Singapore: McGraw-Hill Higher Education; 2007.
23. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. Hoboken, NJ: John Wiley; 2000.
24. Altiok T. On the phase-type approximations of general distributions. *IIE Trans*. 1985;17(2):110–6.
25. Bordes L, Chauveau D. Stochastic EM algorithms for parametric and semiparametric mixture models for right-censored lifetime data. *Comput Stat*. 2016;31(4):1513–38.

26. Cover TM, Thomas JA. *Elements of Information Theory*. 2nd ed. Hoboken, NJ: John Wiley; 2006.
27. Delignette-Muller ML, Dutang C. Fitdistrplus: an R package for fitting distributions. *J Stat Softw*. 2015;64(4):1–34.
28. Benaglia T, Chauveau D, Hunter DR, Young D. mixtools: an R package for analyzing finite mixture models. *J Stat Softw*. 2009;32(6):1–29.
29. Venables WN, Ripley BD. *Modern Applied Statistics with S*. 4th ed. New York, NY: Springer; 2002.
30. Grün B, Leisch F. Fitting finite mixtures of generalized linear regressions in R. *Comput Stat Data Anal*. 2007;51(11): 5247–52.
31. Grün B, Leisch F. FlexMix version 2: finite mixtures with concomitant variables and varying and constant parameters. *J Stat Softw*. 2008;28(4):35.
32. Leisch F. FlexMix: a general framework for finite mixture models and latent class regression in R. *J Stat Softw*. 2004;11(8):18.
33. Franken MD, van Rooijen EM, May AM, et al. Cost-effectiveness of capecitabine and bevacizumab maintenance treatment after first-line induction treatment in metastatic colorectal cancer. *Eur J Cancer*. 2017;75:204–12.