Cancer Epidemiology 57 (2018) 60–67

Contents lists available at ScienceDirect

Cancer Epidemiology

journal homepage: www.elsevier.com/locate/canep

# Matching the model with the evidence: comparing discrete event simulation and state-transition modeling for time-to-event predictions in a cost-effectiveness analysis of treatment in metastatic colorectal cancer patients

Koen Degeling[a,1], Mira D. Franken[b,1], Anne M. May[c], Martijn G.H. van Oijen[d], Miriam Koopman[b], Cornelis J.A. Punt[d], Maarten J. IJzerman[a], Hendrik Koffijberg[a,c,*]

[a] Health Technology and Services Research Department, Technical Medical Centre, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
[b] Department of Medical Oncology, University Medical Centre, Utrecht University, Huispost B02.225, P.O. Box 85500, 3508 GA Utrecht, The Netherlands
[c] Julius Center for Health Sciences and Primary Care, University Medical Centre, Utrecht University, Huispost STR 6.131, P.O. Box 85500, 3508 GA Utrecht, The Netherlands
[d] Amsterdam UMC, University of Amsterdam, Department of Medical Oncology, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands

ARTICLE INFO

*Keywords:*
Markov modeling
State-transition modeling
Discrete event simulation
Time-to-event
Cost-effectiveness
Individual patient data

ABSTRACT

*Background:* Individual patient data, e.g. from clinical trials, often need to be extrapolated or combined with additional evidence when assessing long-term impact in cost-effectiveness modeling studies. Different modeling methods can be used to represent the complex dynamics of clinical practice; the choice of which may impact cost-effectiveness outcomes. We compare the use of a previously designed cohort discrete-time state-transition model (DT-STM) with a discrete event simulation (DES) model.

*Methods:* The original DT-STM was replicated and a DES model developed using AnyLogic software. Models were populated using individual patient data of a phase III study in metastatic colorectal cancer patients, and compared based on their evidence structure, internal validity, and cost-effectiveness outcomes. The DT-STM used time-dependent transition probabilities, whereas the DES model was populated using parametric distributions.

*Results:* The estimated time-dependent transition probabilities for the DT-STM were irregular and more sensitive to single events due to the required small cycle length and limited number of event observations, whereas parametric distributions resulted in smooth time-to-event curves for the DES model. Although the DT-STM and DES model both yielded similar time-to-event curves, the DES model represented the trial data more accurately in terms of mean health-state durations. The incremental cost-effectiveness ratio (ICER) was €172,443 and €168,383 per Quality Adjusted Life Year gained for the DT-STM and DES model, respectively.

*Conclusion:* DES represents time-to-event data from clinical trials more naturally and accurately than DT-STM when few events are observed per time cycle. As a consequence, DES is expected to yield a more accurate ICER.

## 1. Introduction

Healthcare expenditures have increased importantly over the last decades, especially in oncology due to expensive novel targeted agents and personalized treatments based on molecular markers in order to provide patients with the best possible care [1,2]. Cost-effectiveness analysis of such novel medical technologies is becoming increasingly relevant, as it may inform treatment, resource allocation, and research prioritization decisions. This is illustrated by the standardized approaches to value cancer treatment options in terms of efficacy and costs for clinicians [3,4] and guidance for performing cost-effectiveness analysis alongside clinical trials [5].

High quality individual patient data (IPD) on health outcomes, resource use, and care procedures, e.g. obtained from randomized controlled trials (RCTs), are the preferred source of evidence for cost-effectiveness analysis. However, single individual patient datasets do not always provide all (or the only) evidence required for estimating the (long-term) cost-effectiveness of medical technologies [6,7], indicating

* Corresponding author at: Health Technology and Services Research Department, Technical Medical Centre, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
*E-mail address:* h.koffijberg@utwente.nl (H. Koffijberg).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.canep.2018.09.008
Received 8 June 2018; Received in revised form 27 September 2018; Accepted 29 September 2018

the need for cost-effectiveness models to synthesize evidence from additional sources or to extrapolate beyond the time horizon of e.g. RCTs [5,8]. Such cost-effectiveness models should adequately represent clinical practice and, therefore, reflect the true nature of the evidence used to define them, including evidence obtained from RCTs and other sources of IPD. In other words, *the model should match the evidence*.

The primary outcome of many clinical oncology studies is the time until an event of interest occurs, e.g. the patients' overall survival or progression-free survival from the moment of randomization, which are typically recorded continuously over time. However, the most frequently applied cost-effectiveness modeling method, i.e. discrete-time state-transition modeling (DT-STM) [9], uses transition probabilities over discrete time cycles with a fixed length to represent the progression of time. For example, in an DT-STM with time cycles of three weeks patients can only progress to another health state after this predefined and rigid time length, even though in daily practice patients may progress at any time instead of only at a multiple of three weeks. The length of these time cycles needs to be chosen so that the complex dynamics of clinical practice are appropriately represented [9]. For DT-STM to represent clinical practice better, shorter cycle lengths would be preferable [10]. Although half-cycle corrections may be applied to avoid bias and to better approximate clinical practice [11], this still insufficiently allows complex clinical dynamics if the cycle length is too long [12].

Using shorter cycles lengths can be disadvantageous, mainly because of increase in number of cycles that needs to be simulated. Besides increasing the computational burden of the simulation [9,12], the larger number of cycles makes it more challenging to represent the uncertainty in the transition probabilities, as the uncertainty in the numerous cycle-specific probabilities needs to be reflected while also maintaining the correlation between them. Furthermore, because the expected number of observations within a cycle decreases with decreasing cycle length, the likelihood of substantial irregularities in transition probabilities between successive cycles is expected to increase. These irregularities are likely to impact the simulation outcomes and do not correspond to clinical practice, as the probability of an event is commonly expected to be similar between successive moments, i.e. the transition-curves follow a smooth pattern over time.

Discrete event simulation (DES) is an alternative modeling technique to which the challenges associated with discrete time cycles do not apply. Events can occur at any time in a DES model, because the time to these events are typically modeled using smooth time-to-event distributions, e.g. Gamma or Weibull distributions [13]. In DES, the behavior of a system is translated into an ordered sequence of well-defined events, which comprise specific changes in the system's state at a specific point in time [13]. DES is well suitable for modeling clinical processes, as it is able to incorporate patient-level characteristics and clinical histories, competing resources, and interactions between different actors, e.g. physicians and patients [14]. Although originating from the operations research field, DES is increasingly being used for cost-effectiveness modeling [15].

Several studies have compared the use of DT-STM and DES for cost-effectiveness analyses of medical technologies. Using the same model structure and evidence, quantitative outcomes such as the incremental cost-effectiveness ratio (ICER), are unlikely to be substantially different between these modeling methods [16,17]. However, substantial differences in outcomes may occur, if the use of DES results in a more appropriate representation of clinical practice compared to DT- STM, for example by including patient characteristics or considering resource constraints [18]. Especially in the scenario in which insufficient observations are available for the chosen cycle length, and irregularities in the cycle-specific transition probabilities are substantial when using DT-STM, the use of DES might be preferable.

The objective of this study is to compare the evidence structure and outcomes of a recently published cost-effectiveness DT-STM [19] with those of a newly developed DES model. The comparison will be performed based on the dataset of the randomized clinical phase III CAIRO3 study, in which maintenance treatment with capecitabine and bevacizumab (CAP-B) or observation in metastatic colorectal cancer patients after six induction cycles of capecitabine, oxaliplatin, and bevacizumab (CAPOX-B) was evaluated [20]. The results of this study should facilitate a better understanding of the potential impact of selecting a modeling method for cost-effectiveness modeling studies informed by IPD.

## 2. Methods

### 2.1. Maintenance treatment in metastatic colorectal cancer

The CAIRO3 study (NCT00442637) is a randomized clinical phase III study, which was carried out by the Dutch Colorectal Cancer Group (DCCG) in 64 hospitals in the Netherlands. A total of 558 metastatic colorectal patients with stable disease or better after six cycles of CAPOX-B induction therapy were randomized to either receive CAP-B maintenance treatment or observation until progression, which is referred to as the post-induction stage. CAPOX-B treatment was to be reintroduced upon progression on either maintenance or observation, and continued until second progression (primary end-point), which is referred to as the reintroduction stage. Although second progression was the primary end-point of the CAIRO3 study, the cost-effectiveness analysis of the CAIRO3 study also considered additional lines of treatment after second progression [19], which is referred to as the salvage therapy stage. Study results have been previously published [20].

### 2.2. State-transition model

A cohort DT-STM, i.e. Markov model, was originally developed for the cost-effectiveness analysis of the CAIRO3 study and included four health states: post-induction, reintroduction, salvage therapy, and death (Fig. 1a). The model was defined using cohort level cycle-specific transition probabilities, which were estimated from the CAIRO3 trial using Life Tables in IBM SPSS Statistics software, version 23, IBM Corp. (Armonk, NY, USA). This indicates that the probability of moving from one state to another depended only on the time passed since the start of the simulation, e.g. time from randomization until first progression. Half-cycle correction was applied and 100 cycles of three weeks were simulated in total. The DT-STM was built using TreeAge Pro Healthcare v.2014, TreeAge Software (Williamstown, MA, USA), and is described in detail elsewhere [19].

To facilitate an adequate comparison between the two modeling methods, the DT-STM was first replicated in AnyLogic multi-method simulation software, v.7.3, The AnyLogic Company (Chicago, IL, USA), the environment also used for developing the DES model. This replicated DT-STM was then compared to the original DT-STM to assess potential variation in outcomes due to the use of different software environments. In total, 100 events were generated at intervals of three weeks, corresponding to the setup in the original DT-STM. Following each event, the occupation of the health states was recorded and used to calculate health and economic outcomes at the corresponding point in time. The model was validated by structured "walk-throughs", comparing (intermediate) results with calculations by hand, extreme value analysis, trace analysis, and cross validation with the original DT-STM during model development, and sensitivity analysis using the final model [21,22].

### 2.3. Discrete event simulation model

The DES model was defined on patient-level using AnyLogic software and according to the ISPOR-SMDM Modeling Good Research Practice Task Force guidelines [14]. The model was defined to have the same health states as the DT-STM (Fig. 1b). Although DES allows for constrained resources to be accounted for, resource use was not