Original Research Article

MDM Medical Decision Making

Medical Decision Making
2023, Vol. 43(7-8) 949–960

DOI: 10.1177/0272989X231201621
journals.sagepub.com/home/mdm
Sage

Cost-Effectiveness Analysis for Therapy Sequence in Advanced Cancer: A Microsimulation Approach with Application to Metastatic Prostate Cancer

Elizabeth A. Handorf, J. Robert Beck, Andres Correa, Chethan Ramamurthy, and Daniel M. Geynisman

Purpose. Patients with advanced cancer may undergo multiple lines of treatment, switching therapies as their disease progresses. We developed a general microsimulation framework to study therapy sequence and applied it to metastatic prostate cancer. **Methods.** We constructed a discrete-time state transition model to study 2 lines of therapy. Using digitized published survival curves (progression-free survival, time to progression, and overall survival [OS]), we inferred event types (progression or death) and estimated transition probabilities using cumulative incidence functions with competing risks. We incorporated within-patient dependence over time; first-line therapy response informed subsequent event probabilities. Parameters governing within-patient dependence calibrated the model-based results to a target clinical trial. We applied these methods to 2 therapy sequences for metastatic prostate cancer, wherein both docetaxel (DCT) and abiraterone acetate (AA) are appropriate for either first- or second-line treatment. We assessed costs and quality-adjusted life-years (5-y QALYs) for 2 treatment strategies: DCT → AA versus AA → DCT. **Results.** Models assuming within-patient independence overestimated OS time, which corrected with the calibration approach. With generic pricing, AA → DCT dominated DCT → AA, (higher 5-y QALYs and lower costs), consistent for all values of calibration parameters (including no correction). Model calibration increased the difference in 5-y QALYs between treatment strategies (0.07 uncorrected v. 0.15 with base-case correction). Applying the correction decreased the estimated difference in cost (−$5,360 uncorrected v. −$3,066 corrected). Results were strongly affected by the cost of AA. Under a lifetime horizon, AA → DCT was no longer dominant but still cost-effective (incremental cost-effectiveness ratio: $19,463). **Conclusions.** We demonstrate a microsimulation approach to study the cost-effectiveness of therapy sequences for advanced prostate cancer, taking care to account for within-patient dependence.

Corresponding Author
Elizabeth A. Handorf, Rutgers University School of Public Health, Cancer Institute of New Jersey, 120 Albany St., New Brunswick, NJ 08901, USA; (ehandorf@cinj.rutgers.edu).

Highlights

- We developed a discrete-time state transition model for studying therapy sequence in advanced cancers.
- Results are sensitive to dependence within patients.
- A calibration approach can introduce dependence across lines of therapy and closely match simulation outcomes to target trial outcomes.

Keywords

microsimulation model, calibration, therapy sequence

Date received: December 20, 2022; accepted: August 31, 2023

Introduction

In modern oncology practice, patients with advanced cancers often undergo multiple lines of therapies, switching treatments when their disease progresses. The clinically optimal order of therapies may be unclear and is rarely studied by prospective randomized trials, with new treatment paradigms continuing to evolve. In many cancers, specific agents are medically appropriate in several lines of treatment, which motivates the question: Which therapies should be given, and in what order? New therapies are often expensive and may provide only marginal improvements in survival and quality of life, making the cost-effectiveness of different therapy sequences highly relevant for patient and provider decision making.

Although generally understudied, the cost-effectiveness of therapy sequences has been examined in several advanced cancers, including BRAF wild-type melanoma,[1] EGFR mutated non–small-cell lung cancer,[2] HER2+ breast cancer,[3,4] and KRAS wild-type colorectal cancer.[5] These analyses often use state-based Markov models.[1,3–5] A major limitation of this approach is the "memoryless" property of the Markov process; that is, the probability of moving from one health state to another depends only on the current state. Semi-Markov models relax this assumption, but when prior health states inform future events, such models require many additional states to encode prior information and can become unwieldy. This has forced previous studies of therapy sequence to make unrealistic simplifying assumptions.

In the therapy-switching problem, several analytic challenges make modeling complex. First, staggered start times of later lines of therapy can lead to outcomes and transition probabilities dependent on both state time and total model time. Second, the probability of death is highly dependent on prior progression events, as patients are much more likely to die after disease progression. Third, patients' outcomes are dependent over time. Patients who do not respond well to first-line therapy and experience early progression are more likely to do poorly in subsequent lines of treatment due to the inherent aggressive biology of their cancer or other factors such as comorbid conditions. Fourth, not all first-line trial patients would be eligible for second-line trials, as

Rutgers University School of Public Health, Cancer Institute of New Jersey, USA (EAH); Fox Chase Cancer Center, Cancer Prevention and Control Program, Philadelphia, PA, USA (JRB); Department of Surgical Oncology, Fox Chase Cancer Center, Philadelphia, PA, USA (AC); Division Hematology/Oncology, Mays Cancer Center UT Health San Antonio, San Antonio, TX, USA (CR); Department of Hematology/Oncology, Fox Chase Cancer Center, Philadelphia, PA, USA (DMG).
The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: EAH received research funding from Eli Lilly and Pfizer; honoraria from Pfizer Geynisman; consultancy or advisory roles at Pfizer, Exelixis, AstraZeneca, Seattle Genetics/Astellas, Merck, Myovant Sciences, and Bristol-Myers Squibb; research funding from Genentech, Merck, Calithera Biosciences, Astellas Pharma, and Harpoon therapeutics. JRB has stock ownership in GlaxoSmithKline. CR received honoraria from Gilead Sciences; consulting or advisory role at Exelixis, Seagen; research funding from Dispersol, Gilead Sciences, Mirati Therapeutics, Novartis, Nuvation Bio, Seagen. This work was supported, in part, by the Fox Chase Cancer Center core grant P30 CA0069 and NIH U54 CA2217. All work was conducted at Fox Chase Cancer Center, within the Biostatistics and Bioinformatics Facility, Department of Surgical Oncology, and Department of Hematology/Oncology. This work has been presented at the Society for Medical Decision Making North American meetings (2018 and 2022). A preprint of this work is available: https://arxiv.org/abs/2210.05086.

some patients' overall health would deteriorate. The exclusion of such patients from second-line trials would lead to an overestimate of survival for the full first-line cohort (i.e., survivorship bias).

We propose using microsimulation models,[6] which are discrete-time state-transition models that aggregate simulated individual-level trajectories.[7–10] These models are particularly useful when individual factors inform future transition probabilities. These models require simulating a large number of patients to ensure stability of estimates, making them computationally intensive.[6] Nevertheless, modern computing solutions have made these models more practical to run.[8] We use microsimulation models to address each of the problems listed above.

This study is motivated by 2 therapy options for metastatic prostate cancer, abiraterone acetate (AA) and docetaxel (DCT). The cost-effectiveness of these 2 treatments has been of substantial interest, for both first-line and second-line treatment.[11–13] Although AA has a better side-effect profile, both DCT for first-line therapy followed by AA after progression, and AA followed by DCT, have been consistent with the standard of care. Until very recently, AA had a list price upwards of $10,000/month, while generic DCT has been available years. Although generic AA is now available, the question of therapy order is not unique to this situation. As new, expensive agents come on the market, patients, physicians, and payers will need to make value judgments when considering not just whether to give a specific therapy but also when the therapy should be given.

Methods

Microsimulation Models for Multiple Lines of Therapy

Model framework. Because of the complex dependence over time, we propose a discrete time state-based microsimulation model. Figure 1 shows the general structure of our model. Patients start in Line 1 therapy, staying in that state until disease progression. At progression, they would switch therapies and move to the Line 2 state. After progression on Line 2 therapy, the patient would move into the Extensive Disease state, a condition wherein patients have poor quality of life and require expensive care. In this state, patients may receive anticancer therapy or supportive therapy, depending on the cancer. Patients could also move directly from Line 1 to Extensive Disease; the proportion of patients who do not go on to receive Line 2 therapy would depend on disease characteristics. In prostate cancer, in which patients are generally healthy after first-line treatment, few patients would go directly to Extensive Disease, but in a more aggressive disease like pancreatic cancer, this probability would be much higher. Patients can move from any state to Death, an absorbing endpoint.

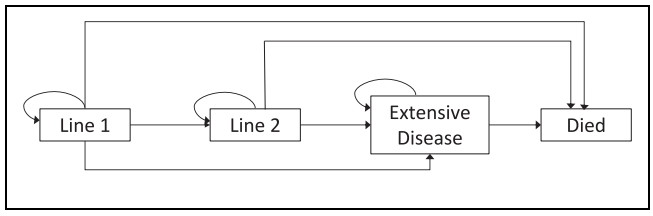

Figure 1 General model structure for 2 lines of therapy.

Estimation of transition probabilities. Many cost-effectiveness studies use published survival curves from clinical trials to inform model transition probabilities. To obtain the correct transition probabilities for this model, one must take into account the dependence of progression and death. Overall survival (OS) curves from clinical trials of first-line therapy cannot be used directly to estimate probabilities of death, as doing so would result in too many individuals dying prior to cancer progression. We propose to use a competing risks framework to estimate transition probabilities.[14] That is, we will estimate the cumulative incidence function of progression accounting for the competing risk of death.

We abstract data from published studies using commercial software to digitize survival curves, which produces a data set defining the survival step function (probabilities and times). Then, the method proposed by Guyot et al[15] allows us to combine the digitized curves with published numbers at risk to infer the underlying event and censoring times. Briefly, this algorithm uses an iterative process to determine the number of censoring events that occur in each interval defined by a risk table corresponding the Kaplan-Meier curves. It first assumes censoring occurs uniformly across each interval and adjusts the number of censoring events until the inferred individual data and published data match. This method has been shown to successfully re-create survival statistics corresponding to published Kaplan-Meier curves. After we apply this algorithm to the progression-free survival (PFS) and OS curves, we have vectors of times (T) and censoring indicators (δ): T_{OS} and δ_{OS} for OS and T_P and δ_P for PFS. For the censoring indicators, 1 indicates that an event occurred, and 0 indicates that the observation was censored.

Death is considered an event when estimating PFS. We propose an algorithm to determine which events should be classified as deaths versus progressions. One

can look for close matches between death event times (abstracted from the OS curve) and progression/death event times (abstracted from the PFS curve). PFS events that occur at the same time as deaths (within some margin of error) are assumed to be deaths. We can therefore obtain a vector of event type indicators (δ_{CR}) for use in a competing risks model. We code δ_{CR} such that 0 denotes a censored observation, 1 denotes progression, and 2 denotes death.

Formally, we initially set $\delta_{CR} = \delta_P$ and $T_{CR} = T_P$.

$$\forall \{T_{OS_i} | \delta_{OS_i} = 1\},$$
$$\text{Let } A_i = \left\{ j | T_{CR_j} \in [T_{OS_i} - \varepsilon, T_{OS_i} + \varepsilon\ ] \ \&\ \delta_{CRj} = 1 \right\}$$
$$\text{set } \delta_{CRa_1} = 2 \text{ if } |A_i| > 0.$$

where ε is the allowable margin of error. More intuitively, this algorithm initially sets the vector of event type indicators as the vector of PFS censoring indicators. For each death event, it looks for progression events within some acceptable margin. If 1 or more progression event meets this criteria (if events are close in time or tied), it sets the first of them to be a death event. Note that not all death events will have a corresponding time in the PFS curve, as most patients progress before death. After one has obtained an inferred data set with event times and event types, cumulative incidence curves can be estimated using standard methods for competing risks.[16] Some studies report progression differently, considering patients who die to be censored instead of failed. The resulting survival functions can be termed time to progression (TTP). We henceforth refer to PFS when including death as an event and TTP when deaths are censored. A similar process to that described above can be used to infer δ_{CR} if the trial reports TTP, by looking for close matches between death events and censoring times in the TTP data set.

The performance operating characteristics of this approach are difficult to quantify in the general case, as they depend on many factors including the distribution of survival/progression times, the resolution of the published figures, the skill of the abstractor at manually selecting points on the survival curve using the DigitizeIt software, and the number of samples. Generally, the fewer events present on the curves, and the higher the resolution of the image, the easier it should be to match steps on the PFS/TTP and OS curves. Therefore, in lieu of a formal assessment, we performed a proof-of-concept analysis that uses a public data set containing death and progression times of cancer patients. See Appendix B.

Model calibration. For second-line treatment and beyond, trial results may be too optimistic for the full cohort of patients. We assume that our study sample is trial eligible at the time of first-line treatment; however, some patients will have deteriorated health at the time of progression and may not meet all inclusion criteria for the second-line trials. The second-line trial results therefore may not be fully generalizable to our cohort, as the studies would tend to exclude sicker patients. Note that we assume that most patients will still be treatable with second-line therapy; this assumption is a variation on a well-known phenomenon that clinical trial populations differ systematically from populations treated in clinical practice.[17,18] Patients with more aggressive disease will tend to have worse response and shorter TTP in all lines of therapy.[19] This can result in a substantial overestimate of OS for the cohort. We propose a calibration approach that simultaneously addresses both of these issues.

We can increase the event probabilities in Line 2 and Extensive Disease states when time in first line (T_1) is short by applying a hazard ratio (HR) to the survival curves, where the HR is some nonincreasing function of T_1. The function defining the HR is denoted as $G(T_1)$.

This will result in a correlation between the time spent in Line 1 and Line 2 states. Further, if $G(t) > 1$ for all or most values of T_1, then it can correct the overestimate of survival outcomes for later lines of therapy. One simple and easy-to-interpret functional form is a linear decrease in G(t) from some maximum HR (θ) to a HR of 1 at some maximum time (ω).

$$G(T_{1i}) = \begin{cases} \theta - \frac{(\theta-1)}{\omega} T_{1i} & T_{1i} < \omega \\ 1 & T_{1i} \geq \omega \end{cases}$$

where T_{1i} = time spent by subject i in Line 1 before first disease progression, θ = largest HR applied to any subject, and ω = smallest time spent in Line 1 for which no penalty will be applied for later lines of therapy.

We propose using a bounded Nelder-Mead algorithm to choose parameters (e.g., θ and ω) that best calibrate the OS curve to the trial data. This is a simplex method that uses gradient descent to find parameters that minimize the objective function.[20,21] Nelder-Mead is particularly useful here because it does not require a fully defined parametric function and can optimize multiple parameters in a constrained parameter space. We propose minimizing the sum of squared error (SSE) between the microsimulation model-based OS curve, $S_{mod}(t)$, and the target OS curve from the relevant trial, $S_{tgt}(t)$. As we use a discrete time model, a natural choice is to calculate the SSE based on the difference between the 2 survival curves at the start of each cycle. We use the following cost function:

$$F(\theta, \omega | \boldsymbol{T}_1) = \sum_{c=1}^{C} \left(S_{mod}(t = c | \theta, \omega, \boldsymbol{T}_1) - S_{tgt}(t = c) \right)^2$$

where C is the total number of cycles and $\mathbf{T_1}$ is the vector of times on Line 1. We then find

$$\arg\min_{\theta, \omega} F(\theta, \omega | \boldsymbol{T}_1) \text{ subject to } \theta \in [1, \infty),\ \omega \in [1, \infty)$$

We can also use bounds to restrict the range of the final parameter estimates to clinically plausible values. The optimization should be repeated with different starting values to ensure local optima are not selected and with many randomization seeds as $S_{mod}(t = c | \theta, \omega, \boldsymbol{T}_1)$ will vary stochastically. After identifying the best values, the resulting model-based survival curves can be tested for lack of fit against the target survival curves using the Komolgorov-Smirnov test. In this analysis, we applied the same adjustment (identical parameters) to both the Line 2 and Extensive Disease states. A more flexible adjustment, if needed, could allow for different adjustments for the 2 different states.

Prostate Cancer Application

Background. These modeling methods were motivated by a study of metastatic hormone-sensitive prostate cancer (mHSPC). This refers to de novo metastatic prostate cancer (patients initially diagnosed with M1 disease) or disease refractory to local treatment not previously treated with hormonal therapy. For decades, the accepted first-line treatment for mHSPC was androgen deprivation therapy (ADT) alone. After first progression, patients are said to have metastatic castrate-resistant prostate cancer (mCRPC), and other agents would be added to ADT. Docetaxel (DCT) is a cytotoxic chemotherapy which has long been a standard therapy for mCRPC. Aberaterone acetate (AA) is a next-generation androgen receptor signaling inhibitor that has been approved in mCRPC since 2011 and has a better safety profile than cytotoxic chemotherapies.[22] Starting in 2015, 2 seminal phase III trials changed the treatment paradigm for mHSPC. The CHAARTED trial demonstrated a TTP and OS benefit for DCT + ADT versus ADT alone.[23] Next, in 2017, the LATITUDE trial demonstrated PFS and OS benefits for AA + ADT versus ADT alone.[22] These trials were selected due to the similarity of their inclusion criteria; nevertheless, some differences were present in the patient populations (see Appendix Table A1).

First-line docetaxel and abiraterone in mHSPC patients have never been compared head-to-head. Meta-analyses demonstrate a potential PFS benefit for AA, but the OS between the 2 treatments is largely similar.[24] According the National Comprehensive Cancer Network guidelines (version 3.2022),[25] either AA + ADT or DCT + ADT are considered category 1 treatments (i.e., those with the strongest level of evidence) for mHSPC. In addition, in patients who progress after either treatment, switching to the other is clinically appropriate. Therefore, treatment sequences of either AA followed by DCT (AA→DCT) or DCT followed by AA (DC→TAA) are appropriate according to consensus guidelines (for simplicity, we drop "+ADT" going forward, but note that all patients will continue to receive ADT in addition to other therapies). In this case where OS outcomes are expected to be similar, the cost-effectiveness of either treatment strategy is of interest for payers, patients, and policy makers. We use this clinical example to illustrate our microsimulation methods.

Model structure and inputs. Figure 2 shows the prostate cancer model, adapted from the model in Figure 1. Adverse events may occur after either Line 1 or Line 2 therapy. Patients may stop active treatment without yet experiencing disease progression; this is part of the planned course of treatment for DCT (6 cycles on Line 1, 10 cycles on Line 2). Although AA should ideally be taken until disease progression, some patients may stop early (as was seen in the trial). Patients off therapy but without disease progression are in the Post-Line1 or Post-Line 2 state.

Transition probabilities between Line 1, Line 2, and Extensive Disease were defined using the methods described above. Survival times were first extracted from digitized survival curves from relevant clinical trials (CHAARTED, LATITUDE, MANSAIL, and COU-AA-302).[22,23,26,27] Calibration parameters were then determined, and the adjusted survival curves defined the transition probabilities. Probability of death from Extensive Disease was based on the PROSILECA trial (which enrolled heavily pretreated patients),[28] and these probabilities were also adjusted using the calibration correction. DigitizeIt software (DigitizeIt version 2.0, Braunschweig Germany) was used to numerically define the survival curves. We assumed a low, fixed (10%) probability of transitioning directly from Line 1 to Extensive Disease based on clinical expertise, due to the relatively healthy population at baseline and the typical course of this disease.

Other model inputs were similar to prior studies of first-line treatment of mHSPC.[11] Selected model inputs are shown in Table 1, with all model specifications listed

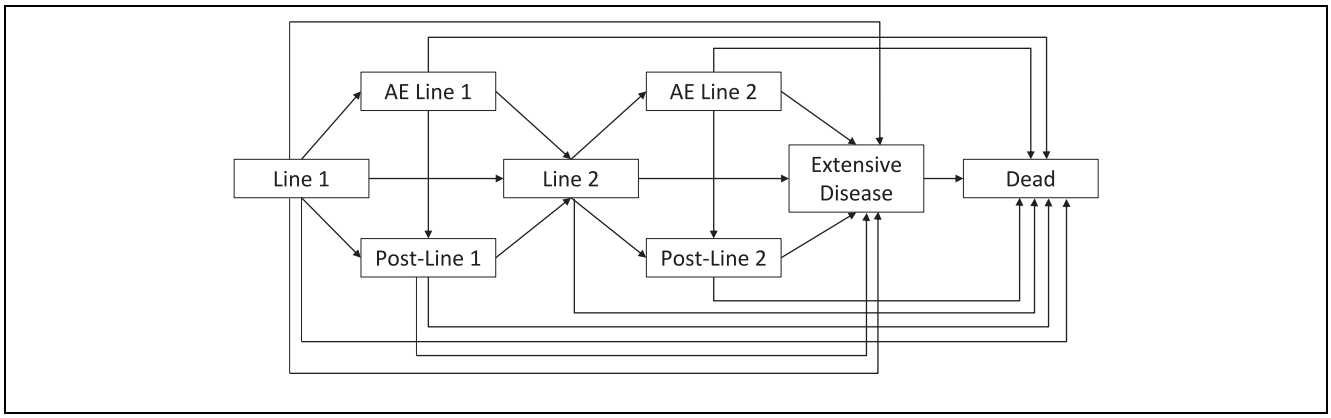

Figure 2 States and allowable transitions in prostate cancer model.

Table 1 Selected Model Inputs Other Than Calibrated State Transitions

Costs (per 3 Week Cycle)	Value	Reference
AA: 2021 branded	$2,396	Amazon[29]
AA: 2021 generic	$296	Amazon[29]
AA: on patent	$6,560	AccessPharmacy[30]
DCT	$2,388	AccessPharmacy[30]
Treatment of febrile neutropenia	$19,675	Lyman et al.[31]
Utility Values	**Value**	**Reference**
L1 SD on AA	0.83	Chi et al.[32]
L1 SD on DCT	0.78	Morgans et al.[33]
L2 SD on AA	0.725	Zhong et al.,[34] Chi et al.[32]
L2 SD on DCT	0.675	Zhong et al.,[34] Chi et al.[32]
Extensive disease	0.62	Zhong et al.[34]
Adverse Event Probabilities	**Value**	**Reference**
L1 Fatigue (AA)	0.02	Fizazi et al.[26]
L1 Neutropenia (DCT)	0.12	Sweeney et al.[23]
L1 Febrile Neutropenia (DCT)	0.061	Sweeney et al.[23]

AA, abiraterone acetate; DCT, docetaxel; SD, stable disease.

in the Supplementary Material. We included adverse events that were relatively common and likely to substantially affect cost and/or quality of life. These included grade 3+ fatigue and neutropenia with and without fever. The probabilities of these events were taken from the respective clinical trials.[22,23,26,27] We assumed that afebrile neutropenia was treated with filgrastim, whereas febrile neutropenia required hospitalization.

Our model included costs of medications, medication administration, physician office visits, and hospitalizations. Medication costs were obtained from retail pharmacies[29] or by average wholesale prices, adjusted to reflect anticipated discounting.[30,34] Costs of medication administration and office visits were taken from the Centers for Medicare and Medicaid Services physician fee schedule.[35] The expense of treating patients in the Extensive Disease state was estimated using results of a prior cost-effectiveness study of third-line treatments for metastatic prostate cancer (cabazitaxel after treatment with docetaxel and androgen inhibitor).[36] Costs of hospitalization for febrile neutropenia were taken from prior literature.[31] All costs were calculated in 2021 dollars.[37]

Quality-of-life adjustments were made via utilities. These were taken from the relevant literature.[33,34,38,39]

Where possible, we used quality-of-life data from the clinical trials of interest.[32] Utilities were lower when on treatment with DCT than when on treatment with AA. We assumed that half of the patients on line 2 therapy would have symptomatic disease (e.g., bone metastasis or other symptoms that reduce quality of life). We assumed that all patients in the Extensive Disease state would have cancer symptoms and a poorer quality of life. We used a 5-y time horizon, as this was the longest follow-up supported by trial data, with results reported as 5-y restricted quality-adjusted life-years (5-y for consistency QALYs). In a lifetime scenario, we use a 25-y horizon. Transition probabilities for the lifetime model after 5 y were based on parametric extrapolations of the trial survival curves and the probability of death from the Social Security Administration 2019 actuarial life tables for men.[40] Cost and effect outcomes were discounted at 3% per year. This model used a payer perspective in the United States of America. Cycle length was 3 wk (based on DCT treatment schedule).

Software. All analyses were performed using R software version 3.6. The code framework for the microsimulation model was adapted from Krijkamp et al.[8] Software to infer the survival times was published by Guyot et al.[15] Cumulative incidence functions were fit using the R package cmprsk.[41] Nelder-Mead optimization was performed using the package neldermead.[42] Our code to run this model is available in the Appendix and in our github repository https://github.com/BethHandorf/CEA_therapy_switching.

Human subjects and funding. This study included only summary data from published sources; no patient-level data were incorporated the model. This work was funded in part by grants from the National Cancer Institute. The funding agency had no role in this study.

Results

We present the results of our microsimulation approach in the context of our prostate cancer application. First, we used the method described in section 2.1.2 to separately identify progression and death events. For Line 1 DCT, there were a total of 397 observations. Of these 181 were progression events. For this study, death was censored in the calculation of the TTP curves. Our method recategorized 38 of the 216 censored observations as deaths while on Line 1 therapy (9.6% identified as death events). For Line 1 AA, there were a total of 597 observations, and 83 of the 247 progression events were recategorized as deaths (for this trial, death was considered an event). Overall, 13.9% of AA progression events were recategorized. When analyzing line 2 PFS curves, 45/526 (8.6%) DCT observations were categorized as deaths, and 60/564 (10.6%) AA observations were categorized as deaths. We note that in our proof-of-concept analysis using a data set with known outcomes, the algorithm underestimated the proportion of death events (see Appendix B), which may have occurred here as well. Nevertheless, the proportion of death events here seems reasonable, as most prostate cancer patients die after their cancer has progressed substantially (i.e., Extensive Disease state).

Our optimization approach identified $(\theta,\omega) = (2.21,87)$ as the best calibration parameters for Line 1 DCT and (5.07,36) for Line 1 AA. This indicates that the DCT survival curve best matches the trial data with a modest to small HR applied to Line 2 and Extensive Disease states and at least some small survival penalty applied to all patients in the simulation (87 total cycles). Contrary to this, for AA, a very large HR is applied to subjects with short times on Line 1 therapy, but no additional hazards are applied to patients who are progression-free on Line 1 therapy for greater than 36 cycles (2.1 y). The results of applying these calibration parameters are shown in Figure 3 and Table 2. With no calibration, the model substantially overestimates survival, but the calibration approach greatly improves fit. Interestingly, both sets of optimal parameters resulted in a nonsignificant test for lack of fit for both DCT and AA OS curves. Because of this, and because of the longer follow-up from the DCT trial, in our clinical example, we used the DCT optimal parameters as our base case. Although this may introduce some bias from residual lack of fit, we chose to prioritize reducing the risk of overfitting the model.

Next, we evaluated the cost-effectiveness of the 2 treatment strategies. We assessed how estimates changed when we varied the calibration parameters. We calculated outcomes using 1) the optimal DCT parameters, 2) the optimal AA parameters, and 3) the optimal parameters applied to each respective treatment. We compared these results to those of no correction. The results of the models are shown in Table 3A. We see that the uncalibrated model gives the highest estimates for both costs and 5-y QALYs for both treatment strategies. The difference in 5-y QALYs was the smallest for the uncalibrated model and was larger after corrections, particularly when applying the same correction to both strategies. Although costs were lower in the calibrated models, the difference in costs varied. When applying the optimal AA correction, the 2 treatment strategies had

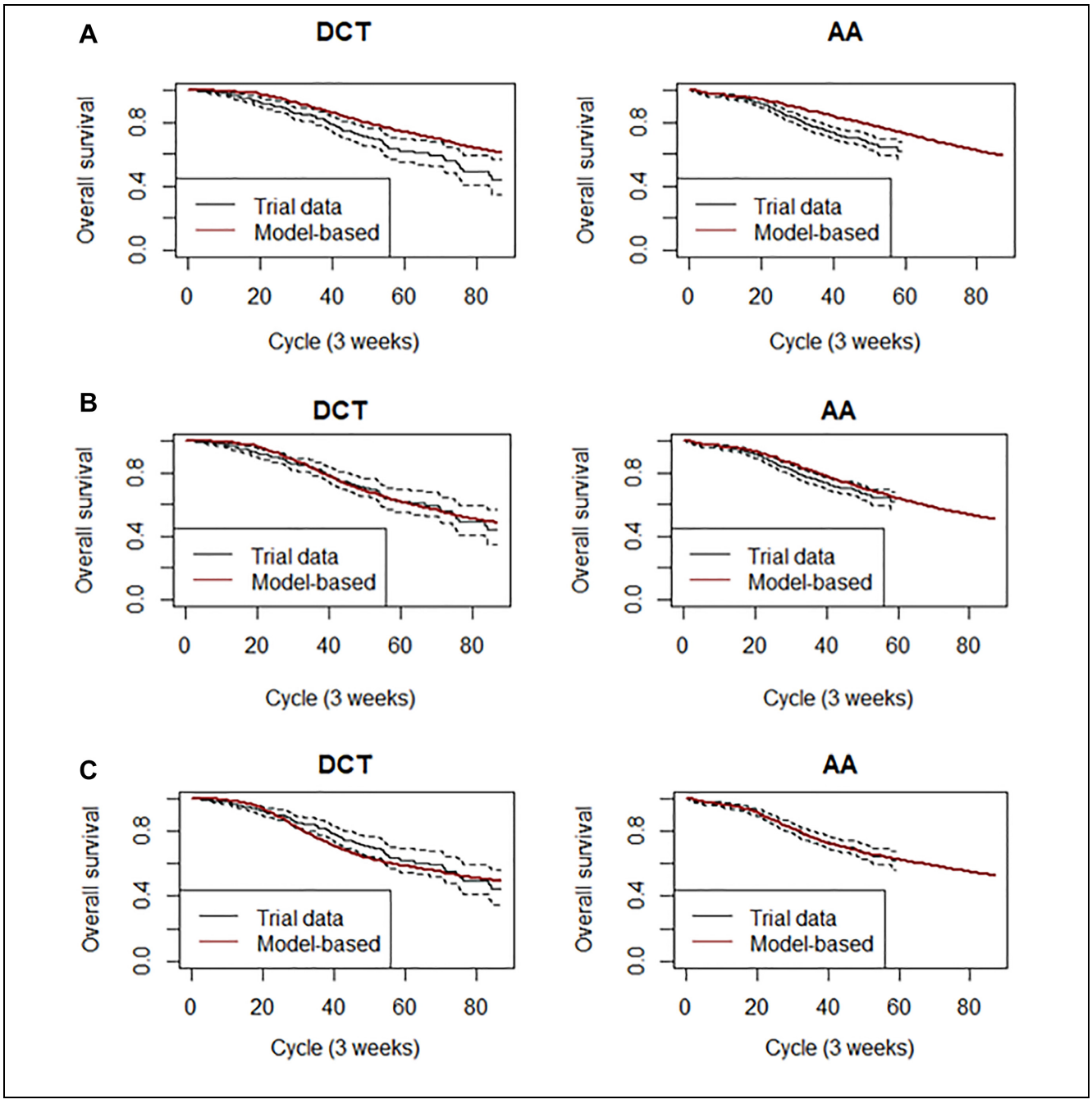

Figure 3 Calibration of model-based curves to trial-based overall survival. (A) No calibration. (B) Optimal calibration parameters for docetaxel. (C) Optimal calibration parameters for abiraterone acetate.

very similar costs, while the largest difference was found when applying different optimal corrections to the respective treatment strategies.

In additional scenarios, we varied the cost of AA, as it has recently become available as a generic, and prior analyses showed that the cost-effectiveness of first-line AA was highly dependent on AA cost.[11] Table 3B shows the results of these models, using the optimal calibration parameters for DCT. (As discussed above, we chose to use the optimal DCT parameters for analyses when

Table 2 Fit of Model-Based Results with Various Calibration Parameters

	Strategy Treating with DCT First			Strategy Treating with AA First		
	3-y OS	**5-y OS**	**p-val[a]**	**3-y OS**	**5-y OS**	**p-val[a]**
Trial data	0.697	0.443		0.658	Not available	
No calibration	0.789	0.613	0.00051	0.776	0.595	0.0023
DCT optimal parameter	0.681	0.487	0.86	0.696	0.509	0.80
AA optimal parameter	0.630	0.495	0.29	0.664	0.531	0.99

AA, abiraterone acetate; DCT, docetaxel; OS, overall survival.
[a]p-val testing the model-based results to the true trial results using the Komolgorov-Smirnov test for lack of fit.

Table 3 Cost-Effectiveness Model Results

A. Effect of Varying Calibration Parameters on Model Results (AA-Generic Cost)

Calibration Parameter	Strategy	Cost	5-y QALY	Δ Cost	Δ 5-y QALY	ICER
No correction	DCT first	$132,144	2.96			
$(\theta,\omega) = (1,1)$	AA first	$126,784	3.03	−$5,360	0.07	Dominant
DCT optimal	DCT first	$119,458	2.72			
$(\theta,\omega) = (2.21,87)$	AA first	$116,392	2.87	−$3,066	0.15	Dominant
AA optimal	DCT first	$112,393	2.63			
$(\theta,\omega) = (5.07,36)$	AA first	$111,577	2.82	−$816	0.19	Dominant
AA: $(\theta,\omega) = (2.21,87)$	DCT first	$119,458	2.72			
DCT: $(\theta,\omega) = (5.07,36)$	AA first	$111,577	2.82	−$7,881	0.10	Dominant

B. Effect of Varying AA Cost on Model Results (DCT Optimal Calibration $(\theta,\omega) = (2.21,87)$)

AA Cost	Strategy	Cost	5-y QALY	Δ Cost	Δ 5-y QALY	ICER
Generic (2021)	DCT first	$119,458	2.72			
	AA first	$116,392	2.87	−$3,066	0.15	Dominant
Branded (2021)	DCT first	$136,849	2.72	NA	NA	NA
	AA first	$211,615	2.87	$74,766	0.15	$512,794
On-patent	DCT first	$171,330	2.72	NA	NA	NA
	AA first	$400,412	2.87	$229,082	0.15	$1,571,189

C. Effect of Extending Model to Lifetime Horizon (DCT Optimal Calibration $(\theta,\omega) = (2.21,87)$)

AA Cost	Strategy	Cost	QALY	Δ Cost	Δ QALY	ICER
Generic (2021)	DCT first	$190,698	4.57			
	AA first	$199,980	5.05	$9,282	0.48	$19,463

AA, abiraterone acetate; DCT, docetaxel; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.

varying AA cost.) We see that, consistent with other studies, the on-patent pricing results in the AA→DCT strategy being not cost-effective, with an ICER > $1.5 million/5-y QALY. Using current brand name pricing, AA→DCT is still not cost-effective; however, with generic pricing, this strategy becomes cost-saving and dominates the DCT→AA strategy. Finally, we extended the model to a lifetime horizon, using the DCT optimal calibration parameters and the generic AA cost. In the main model, approximately half of the cohort was alive at the end of follow-up, so consideration of a lifetime horizon is warranted. Here, AA→DCT was no longer the dominant strategy, as costs were slightly increased; however, it was still highly cost-effective, with an ICER of $19,463/QALY (see Table 3C).

Discussion

We have developed a microsimulation strategy to model treatment sequence across multiple lines of therapy. Our

method allows for event probabilities to be correlated within patients across study states and can calibrate model-based results to those of target trials. We believe that our model could be adapted to many studies of treatment sequence, although subsequent validation is necessary. Multiple lines of therapy are very common in advanced cancers, but our work could potentially be extended to apply to other progressive diseases, such as chronic infections, rheumatic disease, or cardiovascular diseases.

When we first looked at our clinical example of metastatic prostate cancer, we tried to avoid incorporating dependence within patients by using the OS curves directly to define probabilities of death. However, this led to 32% of patients dying while on Line 1 therapy, which is clinically unrealistic in this relatively healthy population.

Using the competing risks framework solved this problem but led to calibration issues. One alternative approach we tried was to increase the probability that patients would move directly from Line 1 to Extensive Disease; however, we could not calibrate results unless an unrealistically high proportion of patients forego second-line treatment. Although this is a metastatic cancer population, prostate cancer is still relatively indolent, and patients may survive for years with their disease. We therefore kept the proportion of patients who forgo Line 2 treatment as low and fixed (10%). Instead, our calibration approach recognizes that patients with poor outcomes on Line 1 therapy are more likely to have worse responses to other therapies, due to the underlying nature of their disease. Our approach, using competing risks estimates with recalibrated transition probabilities, enabled us to create a model with higher fidelity to real patient experiences.

Our proposed methods have several limitations. One drawback of the microsimulation approach in general is its use of discrete time to approximate an underlying continuous process. This can be problematic for several reasons, as discussed by Graves et al.[43] They recommend using discrete event simulation models with continuous time as an alternative. Adapting our framework to a DES model is an important area for future research. However, we note that the ability of the microsimulation model to closely match OS results from relevant trials demonstrates its utility in our clinical scenario.

Our calibration method assumes that the second-line trial data may not be fully generalizable to our microsimulation cohort and that patients' outcomes on different lines of therapy are correlated. Data justifying these assumptions are limited, which is a weakness of our approach. Our calibration approach used a simple functional form for the correction. The linear correction is unlikely to be exactly correct; however, it is simple to estimate and interpret. Other functional forms could be considered, and any monotonic decreasing function of the HR would result in a positive correlation between time spent on Line 1 and Line 2 therapies. One could consider a flexible spline-based approach; however, caution would be needed to avoid model overfitting. In addition, each iteration is time-consuming to fit, so models with more than 2 parameters may become computationally burdensome. In our clinical example, the linear correction was sufficient to provide good calibration, although there was some residual lack of fit (specifically, OS was still overestimated in the first 3 y for DCT). It is possible that allowing different probabilities for the transition directly from Line 1 to Extensive Disease could have improved fit; however, we chose to keep this probability consistent across arms due to a lack of direct information on these parameters.

A limitation of combining data from multiple clinical trials is that the patient populations may differ. Here, the mHSPC DCT patients were younger, had higher Gleason scores, and were less heavily pretreated than patients in the mHSPC AA trial. This could bias survival outcomes in either direction; however, the differences were clinically minor and are not expected to substantially alter long-term outcomes. Using secondary data, in which we do not have actual patient progression and death times, is also a weakness of this analysis. Although our algorithm has face validity, we did find in a proof-of-concept analysis that it cannot perfectly reconstruct individual patient data. This should motivate increased sharing of de-identified individual patient data and use of electronic health record–derived data sets to increase the fidelity of survival estimates. Nevertheless, if no such data are available, our method is a reasonable alternative for modelers to consider.

As the focus of this manuscript is methods development, we did not perform a full sensitivity analysis to vary each parameter used in the illustrative cost-effectiveness example. Instead, we limited the sensitivity analysis to the most pertinent parameters: those used in the calibration. To address second-order uncertainty for the calibration parameters, one could consider uncertainty analyses using a bootstrap or a Bayesian calibration approach. This is an important area for future studies. As the price of AA has recently changed substantially, we varied its cost in this article. A full cost-effectiveness analysis of modern treatments for mHSPC is of interest for future studies, as the treatment paradigms for this disease continue to evolve.[44,45]

In our clinical example, we show how a lack of calibration can alter model estimates, in this case, by

underestimating the benefit of AA→DCT over DC→TAA (0.07 5-y QALYs without calibration v. 0.10 to 0.19 5-y QALYS with calibration). We also showed that under on-patent pricing, the AA→DCT strategy is generally not cost-effective. These results are consistent with prior studies of line 1 therapy (where AA is found to not be cost-effective),[11,12] although we note that here we are testing the full planned treatment strategy, not just first-line treatment. These results changed dramatically with generic AA, where the AA→DCT strategy dominates DCT→AA, being both more effective and less costly. Here we used AA prices from 1 retail pharmacy; future research should obtain more comprehensive estimates. The average wholesale price of generic AA ranges from $510 to $11,665 per month,[30] indicating substantial variability in costs. In our scenario analysis with a lifetime horizon, AA→DCT was no longer dominant but still highly cost effective, with an ICER of $19,463/QALY. This difference was attributable to durable responses to first-line AA, with 35% of patients in our model continuing on treatment for more than 5 y.

In the future, we plan to use this model to assess novel frontline options in prostate cancer such as triplet therapy (AA + DCT + ADT v. darolutamide + DCT + ADT), which has emerged in recent months. We will also extend this model by incorporating patient-level information derived from electronic health records. Models informed by actual patient outcomes on sequences of therapies will better capture dependence structures across lines of therapy.

ORCID iDs

Elizabeth A. Handorf https://orcid.org/0000-0003-0445-8978
J. Robert Beck https://orcid.org/0000-0001-6673-3205

Availability of data and software

Data and software used in this manuscript are publicly available in the repository https://github.com/BethHandorf/CEA_therapy_switching

Supplemental Material

Supplementary material for this article is available online at https://doi.org/10.1177/0272989X231201621.

References

1. Kohn CG, Zeichner SB, Chen Q, Montero AJ, Goldstein DA, Flowers CR. Cost-effectiveness of immune checkpoint inhibition in BRAF wild-type advanced melanoma. *J Clin Oncol*. 2017;35(11):1194.
2. Hirsh V, Singh J. Optimal sequencing strategies in the treatment of EGFR mutation–positive non–small cell lung cancer: clinical benefits and cost-effectiveness. *Am J Health Syst Pharm*. 2020;77(18):1466–76.
3. Diaby V, Adunlin G, Ali AA, et al. Cost-effectiveness analysis of 1st through 3rd line sequential targeted therapy in HER2-positive metastatic breast cancer in the United States. *Breast Cancer Res Treat*. 2016;160(1):187–96.
4. Diaby V, Alqhtani H, van Boemmel-Wegmann S, et al. A cost-effectiveness analysis of trastuzumab-containing treatment sequences for HER-2 positive metastatic breast cancer patients in Taiwan. *Breast*. 2020;49:141–8.
5. Riesco-Martínez MC, Berry SR, Ko Y-J, et al. Cost-effectiveness analysis of different sequences of the use of epidermal growth factor receptor inhibitors for wild-type KRAS unresectable metastatic colorectal cancer. *J Oncol Pract*. 2016;12(6):e710–23.
6. 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.
7. Gibson EJ, Begum N, Koblbauer I, et al. Cohort versus patient level simulation for the economic evaluation of single versus combination immuno-oncology therapies in metastatic melanoma. *J Med Econ*. 2019;22(6):531–44.
8. Krijkamp EM, Alarid-Escudero F, Enns EA, Jalal HJ, Hunink MM, Pechlivanoglou P. Microsimulation modeling for health decision sciences using R: a tutorial. *Med Decis Making*. 2018;38(3):400–22.
9. Schofield D, Zeppel M, Tan O, Lymer S, Cunich MM, Shrestha RN. A brief, global history of microsimulation models in health: past applications, lessons learned and future directions. *Int J Microsimulation*. 2018;11(1):97–142.
10. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ*. 2006;15(12):1295–310.
11. Ramamurthy C, Handorf EA, Correa AF, Beck JR, Geynisman DM. Cost-effectiveness of abiraterone versus docetaxel in the treatment of metastatic hormone naïve prostate cancer. *Urol Oncol*. 2019;37(10):688–95.
12. Hird AE, Magee DE, Cheung DC, Matta R, Kulkarni GS, Nam RK. Abiraterone vs. docetaxel for metastatic hormone-sensitive prostate cancer: a microsimulation model. *Can Urol Assoc J*. 2020;14(9):E418–27.
13. Pollard ME, Moskowitz AJ, Diefenbach MA, Hall SJ. Cost-effectiveness analysis of treatments for metastatic castration resistant prostate cancer. *Asian J Urol*. 2017;4(1):37–43.
14. Aalen O. Nonparametric estimation of partial transition probabilities in multiple decrement models. *Ann Stat*. 1978;6(3):534–45, 12.
15. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012;12(1):9.
16. Collett D. *Modelling Survival Data in Medical Research*. New York: CRC Press; 2015.

17. Elting LS, Cooksley C, Bekele BN, et al. Generalizability of cancer clinical trial results: prognostic differences between participants and nonparticipants. *Cancer*. 2006;106(11):2452–8.
18. Wang W, Ma Y, Huang Y, Chen H. Generalizability analysis for clinical trials: a simulation study. *Stat Med*. 2017;36(10):1523–31.
19. Shore N, Heidenreich A, Saad F. Predicting response and recognizing resistance: improving outcomes in patients with castration-resistant prostate cancer. *Urology*. 2017;109: 6–18.
20. Nelder JA, Mead R. A simplex method for function minimization. *Comput J*. 1965;7(4):308–13.
21. Box M. A new method of constrained optimization and a comparison with other methods. *Comput J*. 1965;8(1): 42–52.
22. Ryan CJ, Smith MR, de Bono JS, et al. Abiraterone in metastatic prostate cancer without previous chemotherapy. *N Engl J Med*. 2012;368(2):138–48.
23. Sweeney CJ, Chen YH, Carducci M, et al. Chemohormonal therapy in metastatic hormone-sensitive prostate cancer. *N Engl J Med*. 2015;373(8):737–46.
24. Kassem L, Shohdy KS, Abdel-Rahman O. Abiraterone acetate/androgen deprivation therapy combination versus docetaxel/androgen deprivation therapy combination in advanced hormone-sensitive prostate cancer: a network meta-analysis on safety and efficacy. *Curr Med Res Opin*. 2018;34(5):903–10.
25. Freedman-Cass D, Berardi R, Shead D. NCCN guidelines version 3.2022 prostate cancer. 2022. https://isotopia-global.com/wp-content/uploads/2022/04/NCCN-guidlines-prostate-cancer-2022.pdf
26. Fizazi K, Tran N, Fein L, et al. Abiraterone plus prednisone in metastatic, castration-sensitive prostate cancer. *N Engl J Med*. 2017;377(4):352–60.
27. Petrylak DP, Vogelzang NJ, Budnik N, et al. Docetaxel and prednisone with or without lenalidomide in chemotherapy-naive patients with metastatic castration-resistant prostate cancer (MAINSAIL): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Oncol*. 2015;16(4):417–25.
28. Eisenberger M, Hardy-Bessard A-C, Kim CS, et al. Phase III study comparing a reduced dose of cabazitaxel (20 mg/m2) and the currently approved dose (25 mg/m2) in postdocetaxel patients with metastatic castration-resistant prostate cancer—PROSELICA. *J Clin Oncol*. 2017;35(28):3198–206.
29. Amazon. Amazon pharmacy. Available from: www.amazon.com.
30. AccessPharmacy. Drug monographs. Available from: https://accesspharmacy.mhmedical.com/drugs.aspx [Accessed 23 November, 2021].
31. Lyman GH, Lalla A, Barron RL, Dubois RW. Cost-effectiveness of pegfilgrastim versus filgrastim primary prophylaxis in women with early-stage breast cancer receiving chemotherapy in the United States. *Clin Ther*. 2009;31(5): 1092–104.
32. Chi KN, Protheroe A, Rodríguez-Antolín A, et al. Patient-reported outcomes following abiraterone acetate plus prednisone added to androgen deprivation therapy in patients with newly diagnosed metastatic castration-naive prostate cancer (LATITUDE): an international, randomised phase 3 trial. *Lancet Oncol*. 2018;19(2):194–206.
33. Morgans AK, Chen Y-H, Sweeney CJ, et al. Quality of life during treatment with chemohormonal therapy: analysis of E3805 chemohormonal androgen ablation randomized trial in prostate cancer. *J Clin Oncol*. 2018;36(11):1088.
34. Zhong L, Pon V, Srinivas S, et al. Therapeutic options in docetaxel-refractory metastatic castration-resistant prostate cancer: a cost-effectiveness analysis. *PLoS One*. 2013;8(5): e64275.
35. Centers for Medicare & Medicaid Services. Medicare physician fee schedule look-up tool. 2021. Available from: https://www.cms.gov/medicare/physician-fee-schedule/search [Accessed 23 November, 2021].
36. Zhang P-F, Xie D, Li Q. Cost-effectiveness analysis of cabazitaxel for metastatic castration resistant prostate cancer after docetaxel and androgen-signaling-targeted inhibitor resistance. *BMC Cancer*. 2021;21(1):35.
37. Statistics USBoL. Consumer price index for all urban consumers: medical care in U.S. city average [CPIMEDSL] FRED, Federal Reserve Bank of St. Louis, 2022. Available from: https://fred.stlouisfed.org/series/CPIMEDSL
38. Bayoumi AM, Brown AD, Garber AM. Cost-effectiveness of androgen suppression therapies in advanced prostate cancer. *J Natl Cancer Inst*. 2000;92(21):1731–9.
39. Younis T, Rayson D, Skedgel C. The cost–utility of adjuvant chemotherapy using docetaxel and cyclophosphamide compared with doxorubicin and cyclophosphamide in breast cancer. *Curr Oncol*. 2011;18(6):288–96.
40. Social Security Administration. Actuarial life table. Available from: https://www.ssa.gov/oact/STATS/table4c6.html
41. Gray B. cmprsk: subdistribution analysis of competing risks. 2.2-10 ed2020. https://cran.r-project.org/web/packages/cmprsk/index.html
42. Bihorel SB, Michael. neldermead: R port of the 'scilab' neldermead module. 1.0-11 ed2010. https://cran.r-project.org/web/packages/neldermead/index.html
43. Graves J, Garbett S, Zhou Z, Schildcrout JS, Peterson J. Comparison of decision modeling approaches for health technology and policy evaluation. *Med Decis Making*. 2021;41(4):453–64.
44. Fizazi K, Foulon S, Carles J, et al. Abiraterone plus prednisone added to androgen deprivation therapy and docetaxel in de novo metastatic castration-sensitive prostate cancer (PEACE-1): a multicentre, open-label, randomised, phase 3 study with a 2 × 2 factorial design. *Lancet*. 2022;399(10336):1695–707.
45. Davis ID, Martin AJ, Stockler MR, et al. Enzalutamide with standard first-line therapy in metastatic prostate cancer. *N Engl J Med*. 2019;381(2):121–31.