Accounting for Cured Patients in Cost-Effectiveness Analysis

Megan Othus, PhD1,∗, Aasthaa Bansal, PhD1, Lisel Koepl, MS1, Samuel Wagner, PhD2, Scott Ramsey, PhD1
1Fred Hutchinson Cancer Research Center, Seattle, WA, USA; 2Oncology Division, Bristol-Myers Squibb, New York, NY, USA

ABSTRACT

Background: Economic evaluations often measure an intervention effect with mean overall survival (OS). Emerging types of cancer treatments offer the possibility of being “cured” in that patients can become long-term survivors whose risk of death is the same as that of a disease-free person. Describing cured and noncured patients with one shared mean value may provide a biased assessment of a therapy with a cured proportion. Objective: The purpose of this article is to explain how to incorporate the heterogeneity from cured patients into health economic evaluation. Methods: We analyzed clinical trial data from patients with advanced melanoma treated with ipilimumab (Ipi; n = 137) versus glycoprotein 100 (gp100; n = 136) with statistical methodology for mixture cure models. Both cured and noncured patients were subject to background mortality not related to cancer. Results: When ignoring cured proportions, we found that patients treated with Ipi had an estimated mean OS that was 8 months longer than that of patients treated with gp100. Cure model analysis showed that the cured proportion drove this difference, with 21% cured on Ipi versus 6% cured on gp100. The mean OS among the noncured cohort patients was 10 and 9 months with Ipi and gp100, respectively. When ignoring cured proportions, we found that the incremental cost-effectiveness ratio (ICER) when comparing Ipi with gp100 was $324,000/quality-adjusted life-year (QALY) (95% confidence interval $254,000–$600,000). With a mixture cure model, the ICER when comparing Ipi with gp100 was $113,000/QALY (95% confidence interval $101,000–$154,000). Conclusions: This analysis supports using cure modeling in health economic evaluation in advanced melanoma. When a proportion of patients may be long-term survivors, using cure models may reduce bias in OS estimates and provide more accurate estimates of health economic measures, including QALYs and ICERs. Keywords: cure models, oncology, overall survival, survival analysis.

Copyright © 2017, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

Progress in the treatment of cancer has led to some patients being cured of their disease in the sense that they become long-term survivors whose risk of death is the same as that of a person who did not have cancer. Cured patients can induce heterogeneity in the overall survival (OS) of a patient population that may not be adequately described with traditional statistical analyses. For example, when some patients are cured, the mean OS of the full patient population is equal to the weighted average of the OS among cured and the OS among uncured patients, weighted by the relative proportions. The mean OS of cured patients is often much longer than the mean OS of uncured patients, and may in fact exceed the observation period of clinical studies. Grouping all patients together and reporting one shared mean value may provide an incomplete assessment of a therapy that cures a proportion of patients. In addition, statistical methods that do not account for cured patients may provide biased assessments of OS.

Mean OS is a measure frequently used in health economic evaluation, for example, in the evaluation of the mean effect of a treatment. This article aims to explain how to incorporate the heterogeneity from cured patients into health economic evaluation. In particular, we will explain how to modify the calculation of quality-adjusted life-years (QALYs) and incremental cost-effectiveness ratios (ICERs) to account for cured patients. The ICER summarizes the additional value of a treatment and is defined as the ratio of the difference between mean treatment costs and the difference in mean treatment effects:

\[
\frac{(\text{Mean cost treatment}_1 - \text{Mean cost treatment}_2)}{(\text{Mean effect treatment}_1 - \text{Mean effect treatment}_2)}
\]

Mean effects are often measured directly with mean OS or with survival weighted by quality of life, or QALYs. As an illustration, we will use clinical trial data from patients with advanced-stage melanoma treated with ipilimumab (Ipi) to show how health economic evaluations that explicitly account for cured patients can differ from standard analyses that do not model cured patients.

Cure Models

Statistical methodology for cure models has been an active area of research for more than 50 years. The most popular framework for cure models is to assume that the study population is a mixture of patients who are cured and patients who are not cured...
and to explicitly model this mixture [1–4]. In this framework, regression models can be used to estimate the probability that a patient is cured and to predict the survival of patients who are not cured.

At present, there are no diagnostic tests that can assess whether an individual patient is cured of his or her cancer. So, long-term follow-up is the ultimate way to determine whether a cured subpopulation exists. For studies with limited follow-up, cure models may be able to provide only limited estimates, and the confidence intervals (CIs) of estimates should reflect the ambiguity. Cure models are useful when survival curves indicate a possibly heterogeneous patient population, with some patients failing quickly and others having long survival. These models do not identify individual patients as cured but rather supply a probability that a patient is cured.

Logistic regression is a common choice to model the probability that a patient is cured [5]. Both patients who are cured and patients who are not cured are subject to “background” mortality not related to cancer. Patients who are not cured are subject to additional mortality from their cancer, and parametric survival models are often used to estimate this excess mortality. Mathematically, the survival for a population with a cure fraction can be written as follows:

\[ S(t;x) = S_0(t;x) \left[ p(x) + (1 - p(x)) S_1(t;x) \right], \tag{1} \]

where \( S(t) \) denotes the survival at time \( t \), \( S_0(t; x) \) denotes the background mortality at time \( t \) conditional on covariates \( x \), \( p(x) \) denotes the probability of being cured conditional on covariates \( x \), and \( S_1(t; x) \) denotes the mortality due to disease (e.g., cancer) at time \( t \) conditional on covariates \( x \) [6,7]. We note that \( S_0, p, \) and \( S_1 \) can be written more generally to depend on different covariates, but we focus on the scenario with shared covariates without loss of generality. \( S_0 \) can be calculated from external data; for our application we used age- and sex-matched mortality data from the US Social Security life tables. We modeled \( p(x) \) with logistic regression and considered Weibull and lognormal parametric formulations of \( S_0 \).

## Cure Models and Health Economic Evaluation

Economic evaluations of competing interventions often estimate mean OS or QALYs for each intervention from clinical trials. However, for clinical trials with heterogeneity due to cured patients, the survival curves plateau and not drop to 0 during the finite follow-up of the trial. If the observed survival is not 0 at the end of the observation period, the mean value cannot be estimated without constructing a model. Parametric models such as the Weibull and lognormal can be used to calculate the mean. If a population contains a mixture of cured and uncured patients, the mean survival of the population should be calculated as the weighted average of the mean survival times of each of the cured and uncured subpopulations, weighted by the relative proportions. In the model in Equation 1, the mean OS of the cured proportion is the mean of the background mortality \( S_0 \), whereas the mean OS for the uncured patients is a function of both the background mortality \( S_0 \) and the disease-related mortality \( S_1 \). In many cancer applications, the mortality from \( S_1 \) is much higher than the background mortality \( S_0 \). The mean of a random variable with survival function \( S(t) \) is equal to \( \int_0^\infty S(t) dt \), and so the mean OS for cured patients is equal to \( \int_0^\infty S_0(t) dt \) and the mean OS for uncured patients is equal to \( \int_0^\infty S(t) S_1(t) dt \).

Cured and uncured patients will also have different costs because the cured patients will have long-term follow-up costs that are associated with long-term surveillance of their cancer and related medical costs. Similar to the calculation of mean OS, the mean costs associated with a therapy should be calculated as the weighted average of the mean costs for cured patients and the mean costs for uncured patients, weighted by the relative proportions. One issue in the estimation of mean costs is that survival times are censored on some study subjects and we do not observe their total costs. A naive sample average of the total observed costs can give biased results. To address this issue, we will use the nonparametric Kaplan-Meier sample average (KMSA) estimator to calculate mean costs [8]. The KMSA technique partitions the time period of interest into small intervals and uses cost histories to determine the mean cost \( M \) as follows:

\[ M = \sum_{i=1}^{\infty} \frac{\tilde{S}_i}{C_i}, \tag{2} \]

where \( C_i \) is the average cost over the \( i \)th interval conditional on surviving until the beginning of the interval and \( \tilde{S}_i \) is the probability of being alive at the beginning of the \( i \)th interval, estimated using the KMSA estimator. Lin et al. [9] demonstrated that the Kaplan-Meier estimator is unbiased and consistent as long as 1) censoring is independent in time and 2) the time intervals for the cost analysis are sufficiently narrow.

### Statistical Methods

Survival was estimated using the Kaplan–Meier method. Parametric survival models without a cure fraction were estimated. Parameters for \( p(x) \) and \( S_0(t; x) \) from Equation 1 were estimated using the score equations from the log-likelihood in the study by Lambert [7] (a reference for implementing a version of Equation 1 in the statistical program Stata [StataCorp, College Station, TX]). Mean survival for uncured patients [equal to \( \int_0^\infty S_0(t) dt \) defined as the average cost over the \( i \)th interval conditional on surviving until the beginning of the interval and \( \tilde{S}_i \) is the probability of being alive at the beginning of the \( i \)th interval, estimated using the KMSA estimator. Lin et al. [9] demonstrated that the Kaplan-Meier estimator is unbiased and consistent as long as 1) censoring is independent in time and 2) the time intervals for the cost analysis are sufficiently narrow.

#### Ipi Case Study

Ipi is a monoclonal antibody that targets cytotoxic T-lymphocyte–associated antigen 4, a protein receptor that downregulates the immune system to allow cytotoxic T-lymphocytes to continue to target cancer cells. Ipi has been approved by the United States for treatment of unresectable stage III and metastatic melanoma, by Canada for treatment of unresectable stage III and metastatic melanoma in patients who have failed or failed to tolerate other therapies, and by the European Union for first-line and second-line treatment of metastatic melanoma.

We considered patient-level data from a randomized trial in patients with unresectable stage III and IV melanoma [10]. The trial randomized patients to three arms: an Ipi arm, an active control arm with a cancer vaccine derived from the melanosomal glycopeptide 100 (gp100), and an Ipi + gp100 arm. In the following text, we focus attention on the gp100 and Ipi arms for clearer exposition. In these two arms, the median age was 58 years with a range of 19 to 91 years, with no significant difference in age between arms (Wilcoxon P value = 0.998). OS was measured from the date of randomization to the date of death from any cause, with patients last known to be alive censored at the date of last contact. The solid lines in Figure 1 display Kaplan–Meier estimates of OS. The median follow-up of censored patients is 1.8 years. The plateau at the tail of the curve indicates that more than 15% of the patients on the Ipi arm could be long-term survivors.

Because the Kaplan–Meier estimates do not drop to 0 at the end of follow-up, the empirical curve cannot be used to estimate the mean survival in this patient population. Previous work considered approaches to estimate the mean OS of this
population by assuming parametric models for either the whole curve or the tail of the curve [11]. We considered an alternative approach and modeled this population as a mixture population.

For the background mortality $S_0$, we used age- and sex-matched life table data for the 2010 US Social Security area population. Although background rates from population-based sources also contain mortality associated with melanoma, in practice this has little effect on the parameter estimates [12]. For the excess mortality due to melanoma among noncured patients, we considered Weibull and lognormal distributions. These two distributions are flexible enough to model various potential hazard functions; the Weibull distribution can model monotonic hazard functions and the lognormal distribution can model a hazard function that first rises and then falls. On the basis of the graphical fit of the data, we present results for the Weibull models here. Results for the mixture cure model and a standard analysis that does not model a cure fraction are presented in Table 1.

When not explicitly modeling a cured proportion, the Weibull analysis indicates that the Ipi arm has a mean OS that is 8 months (0.7 years) longer than that of the gp100 arm. In the mixture cure analysis, the mean OS for noncured patients is similar in the two arms, 9 months on the gp100 arm and 10 months on the Ipi arm. The mean OS for cured patients is also similar in both arms and expected to be 26 years. Under this model, the survival differences between the arms are driven by the proportion of cured patients, with 15% more patients being cured on the Ipi arm than on the gp100 arm. Figure 1 provides the Kaplan-Meier estimates, the standard Weibull model estimates, and the cure model estimates of survival, with the model estimates projected out to 10 years past randomization. The cure model fits the observed data well and the projected survival shows a slow decline in expected mortality. The standard Weibull model that does not include a cure proportion does not fit the tail of the observed data well and essentially assumes that all patients would have died within 10 years.

The mixture cure models can include covariates to further explore and describe the mixture population. We fit a Weibull mixture cure model including the following pretreatment covariates (covariates that were prespecified for subgroup analysis in Hodi et al. [10] as potential prognostic factors): age (quantitative), sex (male vs. female), performance status (0 vs. 1), stage (M0, M1a, and M1b vs. M1c), lactate dehydrogenase (a known prognostic factor for advanced melanoma; elevated vs. normal), previous interleukin-2 (IL-2) therapy, and previous chemotherapy. Table 2 presents the results for this analysis.

Similar to what was observed in the univariate analysis in Table 1, patients treated with Ipi had an improved probability of cure compared with gp100, but there was no significant difference in survival for those who were not cured ($S_0$). Patients who had received IL-2 therapy previously were more likely to be cured than patients who had not received IL-2 therapy previously, and patients with stage M1c disease were less likely to be cured than patients with stage M0, M1a, or M1b disease. Increased performance status, lactate dehydrogenase level, and previous chemotherapy were associated with poorer survival for patients who were not cured.

To calculate QALYs for each arm, we used utility weights from Tromme et al. [13] for patients with stage IV melanoma. Within a disease state, utility weights were taken to be equal in the two arms. Tromme et al. [13] provided utility weights from the start of treatment (utility = 0.583) and from the start of remission (utility = 0.796). For the standard Weibull analysis that did not model a cure fraction, we assumed that the response rate was equal to the cure proportion from the cure model. The cohort that did not respond had a constant utility weight from the start of treatment (0.583). For the cohort that did respond, the utility weight for the first 4 months was equal to the start of treatment weight (0.583) and the utility weight for the remaining time was equal to the remission utility weight (0.796). Scheduled tumor assessments on the trial were done at baseline and at week 12. Patients with stable disease or better had confirmatory scans at week 16. The 4-month cutoff for the remission time was chosen on the basis of this confirmatory scan schedule. For the mixture cure model analysis, the utility weight for the cohort that was not cured was constant and equal to the start of treatment weight (0.583). For the cohort that was cured, the utility weight for the first 4 months was equal to the start of treatment weight (0.583) and the utility weight for the remaining time was equal to the remission utility weight (0.796). Table 3 presents the results of these calculations.

For cost calculations, we used melanoma cost data from Seidler et al. [14], which allocated charges to three phases: the first 4 months (initial phase), the last 6 months (terminal phase), and the time in between (interim phase). We assumed the full initial-phase costs and the full terminal-phase costs for all

---

**Table 1 – Standard analysis and mixture cure model analysis of Ipi data.**

<table>
<thead>
<tr>
<th>Summary Statistic</th>
<th>gp100 (95% CI)</th>
<th>Ipi (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weibull analysis</strong> (without cure modeling)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean OS (y)</td>
<td>0.90 (0.75–1.11)</td>
<td>1.60 (1.23–2.10)</td>
</tr>
<tr>
<td><strong>Mixture cure model analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cure proportion (%)</td>
<td>6 (5–15)</td>
<td>21 (13–30)</td>
</tr>
<tr>
<td>Mean OS of cured patients (y)</td>
<td>26 (24–28)</td>
<td>26 (24–28)</td>
</tr>
<tr>
<td>Mean OS of noncured patients (y)</td>
<td>0.75 (0.59–0.91)</td>
<td>0.83 (0.69–0.99)</td>
</tr>
</tbody>
</table>

CI, confidence interval; gp100, glycoprotein 100; Ipi, ipilimumab; OS, overall survival.
patients on both arms in the analyses, even though some patients lived for less than 6 months. In analyses that accounted for cured patients, terminal-phase costs for cured patients were taken to be equal to the 2012 national average costs of the last year of life of patients without cancer from the Dartmouth Atlas of Healthcare [14]. All cost estimates from the literature were adjusted to 2014 costs with medical cost inflation [15]. The KMSA method was used to calculate interim-phase costs. The cost of Ipi was assumed to be $120,000 for all patients on the Ipi arm [15]. Because gp100 is not approved by the Food and Drug Administration, no additional costs for gp100 were assumed. The results are presented in Table 3.

Incorporating potential cure fractions increases the QALY values for both arms. We note that the time horizon is a lifetime based on the life table data used for the background mortality. The QALY gain is greater in the Ipi arm, reflecting the higher proportion of cured patients with that therapy. We conservatively assumed a utility less than 1 for all patients at all time points, but in some scenarios it may be reasonable to expect that utilities for cured patients may reach a value closer to 1 over time, which would further increase the QALY. Similar to the QALY calculations, costs also increase in the cure model compared with the standard model, because of the increased costs associated with cured patients expected to live a mean of 26 years. We conservatively assumed constant interim-phase costs for all cured patients, but in some scenarios it may be reasonable to expect costs for cured patients to decrease over time, which would decrease the overall mean costs. With a standard Weibull model that does not incorporate potential cure factions, we found the ICER when comparing Ipi with gp100 was $324,000/QALY (95% CI $254,000–$600,000). With the mixture cure model, the ICER when comparing Ipi with gp100 was $113,000/QALY (95% CI $101,000–$154,000).

### Table 2 – Multivariable Weibull mixture cure model results.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Cure model</th>
<th>S model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>Ipi (ref. = gp100)</td>
<td>2.01 (1.00–4.06)</td>
<td>0.05</td>
</tr>
<tr>
<td>Age (y)</td>
<td>1 (0.98–1.03)</td>
<td>0.93</td>
</tr>
<tr>
<td>Male (ref. = female)</td>
<td>0.81 (0.42–1.57)</td>
<td>0.53</td>
</tr>
<tr>
<td>Performance status 1 (ref. = 0)</td>
<td>0.77 (0.40–1.48)</td>
<td>0.43</td>
</tr>
<tr>
<td>Elevated LDH (ref. = normal)</td>
<td>0.52 (0.24–1.36)</td>
<td>0.11</td>
</tr>
<tr>
<td>Previous IL-2 (ref. = no previous IL-2)</td>
<td>2.38 (1.14–4.99)</td>
<td>0.022</td>
</tr>
<tr>
<td>Previous chemotherapy (ref. = no previous chemotherapy)</td>
<td>1.09 (0.33–3.58)</td>
<td>0.88</td>
</tr>
<tr>
<td>Stage M1c (ref. = M0, M1a, M1b)</td>
<td>0.43 (0.20–0.91)</td>
<td>0.027</td>
</tr>
</tbody>
</table>

CI, confidence interval; HR, hazard ratio; IL-2, interleukin-2; Ipi, ipilimumab; LDH, lactate dehydrogenase; OR, odds ratio; ref. reference.

### Table 3 – QALY and cost analysis summary.

<table>
<thead>
<tr>
<th>Summary Statistic</th>
<th>gp100 (95% CI)</th>
<th>Ipi (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>QALYs</td>
<td>0.53 (0.43–0.66)</td>
<td>0.99 (0.74–1.36)</td>
</tr>
<tr>
<td>Mean costs</td>
<td>0.89 (0.66–1.23)</td>
<td>2.38 (2.15–2.72)</td>
</tr>
<tr>
<td>(&gt;$100,000)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Weibull analysis (without cure modeling)

Mixture cure model analysis

| QALYs             | 1.67 (1.31–2.38) | 4.72 (3.10–6.37) |
| Noncured patients | 0.43 (0.33–0.52) | 0.48 (0.40–0.57) |
| Cured patients    | 20.36 (18.88–21.88) | 20.70 (19.15–22.28) |
| Mean costs        | 1.75 (1.47–3.05) | 5.20 (4.00–6.41) |
| (>$100,000)       |                |             |
| Noncured patients | 0.83 (0.74–0.92) | 2.08 (2.00–2.18) |
| Cured patients    | 15.46 (14.31–16.61) | 16.92 (15.74–18.11) |

CI, confidence interval; gp100, glycoprotein 100; Ipi, ipilimumab; QALYs, quality-adjusted life-years.

### Discussion

Standard survival modeling techniques may be insufficient to calculate cost-effectiveness summaries in situations in which therapies essentially cure a proportion of patients treated. To address this issue, we developed a mixture cure model approach that can be applied to estimate outcomes and costs. In the case of Ipi for patients with advanced melanoma, we find that the cure model approach produces substantially different estimates of incremental cost-effectiveness than do traditional approaches. Compared with standard modeling, the cure modeling approach will substantially increase costs and QALYs for the fraction of cured patients, whereas it will slightly reduce costs and QALYs for the noncured population. The degree of change between standard modeling and cure modeling is dependent on the cure proportions, costs, and utilities. Thus, in situations in which treatments provide a fraction of patients with durable remissions from their illnesses, we believe that the mixture cure model approach may be superior to standard approaches in estimating incremental cost-effectiveness.

There are a number of clinical situations in which treatments can cure a proportion of patients. Probably the most common situations are those involving blood cancers in children, including Hodgkin’s disease in younger adults, in which the proportion of patients cured can exceed 80% [16]. More recently, treatments for patients with advanced melanoma have been found to produce long-term survival results in approximately 10% of patients treated. Of note, in both these situations, patients who are not cured after initial therapy face severely reduced overall expected survival and high short-term mortality.

We note situations in which cost-effectiveness analysts might consider mixture cure model techniques versus standard survival modeling. The first, shown here, is when a small proportion of patients display substantially better survival than does the remainder of the treated population. The components of interest here are the absolute difference in mean survival between cured and noncured groups and the proportion of patients who achieve the better outcome. Although there is no threshold difference in survival that makes mixture cure modeling superior to standard survival modeling, cure models may be warranted whenever there is sufficient follow-up that supports evidence of a heterogeneous population with some long-term survivors and there exists a hypothesis or explanation for such a heterogeneous population. With enough follow-up, a fraction of long-term
survivors can be identified in survival plots that exhibit a plateau at the tail of the curve, or that show late long-term events as the study population ages. Follow-up times for many clinical trials are insufficient to detect populations with markedly improved outcomes such as have been reported for melanoma. Applying cure models requires a certain “maturity” of data such that these differences can be identified.

We note that this is not a comprehensive cost-effectiveness analysis of this data set. We have used these melanoma data to feature differences in two alternative statistical methodologies for calculating survival and studied the impact the different statistical methodologies had on two health economic measures, the QALY and the ICER. More comprehensive cost-effective analyses have been performed for several health authorities, including the UK National Institute for Health Care and Excellence [17].

Our analysis had limitations. We used utility and cost data available from the literature in our analysis, but in melanoma these data are not necessarily well-suited for a cure modeling framework. Additional follow-up is warranted to study whether utilities and costs change over time among patients who are long-term survivors. It may be reasonable to expect that utilities would increase and costs would decrease, which would impact our ICER calculation. The follow-up from the clinical trial is less than 5 years and it will be of interest to see whether the projected trend is consistent with what is observed with the actual follow-up of the trial.

Conclusions

We found that using a cure modeling approach to account for long-term survivors in a clinical trial for advanced melanoma significantly changed several health economic measures. In particular, the ICER comparing the two arms decreased by 65% because of the significantly higher QALYs for cured patients and one arm having a greater proportion of cured patients. The cure modeling framework is likely to be useful for health economic analysis for other cancers and diseases.

Acknowledgments

We thank Srividya Kotapati, Michael Lees, and Tai-Tsang Chen of Bristol-Myers Squibb for their critical review of the manuscript.

Source of financial support: Financial support for this study was provided entirely by a contract with Bristol-Myers Squibb. The funding agreement ensured the authors’ independence in designing the study, interpreting the data, writing, and publishing the report. Dr. S. Wagner is employed with the sponsor.

Supplemental Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2016.04.011 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

REFERENCES

[17] Ipilimumab for previously treated advanced (unresectable or metastatic) melanoma | Guidance and guidelines | NICE. http://www.nice.org.uk/guidance/ta268/chapter/1-guidance.