

Evaluating the Trade-off between Quantity and Quality of Life through Quality-Adjusted Survival Analysis

Lawrence Rasouliyan

ICON Clinical Research, Medical Affairs Statistical Analysis, Barcelona, Spain

ABSTRACT

Clinical trials often employ many different endpoints to evaluate the benefits and risks associated with treatment. These endpoints, however, may not always align with one another. In oncology studies, for instance, a more efficacious treatment, which prolongs survival, may also be associated with greater toxicity, which decreases quality of life. Quality-adjusted survival analysis (QASA) provides a useful tool to calculate summary measures that incorporate both survival and overall quality of life so that the trade-off can be compared across treatments. Topics to be covered include construction of quality-adjusted Kaplan-Meier curves, calculation of quality-adjusted life years (QALYs), and examination of an exploratory method to evaluate quality-adjusted median survival.

INTRODUCTION

Taking a therapy commonly entails a trade-off between benefits and risks. For instance, does the benefit of a sleeping medication outweigh the possible side effect of dizziness? Does the benefit of delaying cancer progression for a finite amount of time outweigh the toxicity of chemotherapy? The question of “is it worth it?” typically comes into play when sizing up the trade-off between desired outcomes and potential side effects.

In disease indications with particularly poor survival prognoses, this evaluation of trade-off evolves into a question of quantity of life versus quality of life. The evaluation of new therapies in these disease areas commonly focuses on assessing endpoints related to benefit and risk, where the benefit manifests itself as prolonged survival, and the risk manifests itself as treatment toxicity, which in turn translates to poorer quality of life.

In clinical trials, these endpoints are reported separately from one another so that the advantages and disadvantages of the therapy can be evaluated within the context of the particular patient's situation. If the endpoints related to survival and quality of life are both in favor of one particular treatment, then the treatment choice is clear. In situations, however, where a particular treatment has favorable survival endpoints but unfavorable quality of life endpoints, then further evaluation may be desirable. In the latter scenario, it may also be useful to calculate a summary measure that incorporates both survival and quality of life in order to gain further insight into the quantity versus quality trade-off and to compare this trade-off across treatments.

BACKGROUND

In clinical studies, the benefit endpoint of prolonged survival can be evaluated through methods of survival analysis, and the risk endpoint of toxicity can be evaluated through assessment of health-related quality of life (HR-QoL) over time. Quality-adjusted survival analysis (QASA) is a useful tool that integrates these two techniques into a summary measure that assesses the quantity versus quality of life trade-off. Brief background will be provided on survival analysis and HR-QoL before delving into the techniques of QASA.

SURVIVAL ANALYSIS¹

Survival analysis is an analytic method where the outcome of interest is the time to the occurrence of an event of interest. In many cases, and in the context of this study, the event of interest is death. Of particular importance is understanding the behavior of the survival function (denoted as S), which is the probability that the time of death (denoted as T) is later than a given time (denoted as t). That is,

$$S(t) = \Pr(T > t) \quad (\text{Equation 1})$$

The survival function must be non-increasing, and it is typically assumed to be equal to 1 at time 0. Furthermore, the survival function is assumed to approach 0 as time approaches infinity. Thus, at the beginning of observation, the instantaneous probability of survival is 100%, and as time approaches infinity, the instantaneous probability of survival approaches 0%.

An important attribute of the survival function is that it takes into account a form of missing data known as censoring. Censoring is a very common phenomenon in survival data, and it describes the notion that not all patients are observed throughout the entire course of follow-up. Censoring can take various forms; though, it primarily occurs as right censoring (unobserved data are to the “right” or later than of the patient’s last point of observation). Censoring can also be described as informative (where the censoring is related to some set of patient characteristics) or non-informative (where the censoring occurs more or less in a randomly regardless of patient characteristics). In the context of this study, the case of non-informative right censoring will be assumed for all patients who experience censoring.

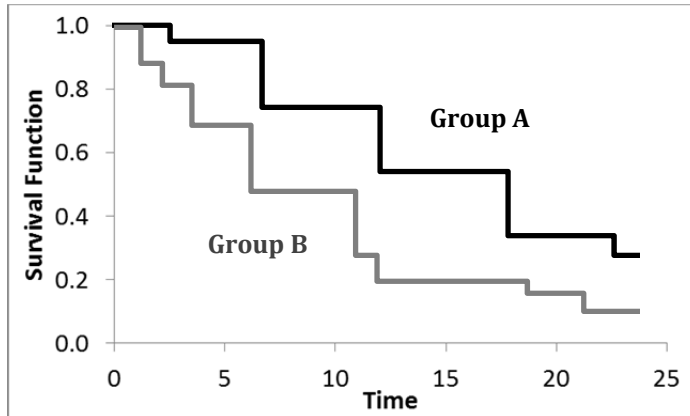
A common method of estimating the survival function is through the use of a Kaplan-Meier estimator (also known as product-limit estimator), which is the non-parametric maximum likelihood estimate of the survival function. Using this method, the survival function estimate is assumed to be constant over time (illustrated by a horizontal line segment) until the point in time where a death occurs. At this point, the survival function estimate steps downward to its new value (illustrated by another horizontal line segment) until another death occurs. Hence, at each observed death (denoted as the i th death), the survival function estimate can be calculated as follows:

$$\hat{S}(t) = \prod_{t_i < t} \frac{(n_i - d_i)}{n_i} \quad (\text{Equation 2})$$

where n_i denotes the number of patients at risk just prior to time i (i.e., the number of patients still alive minus the number of patients who have been censored), and d_i denotes the number of deaths at time i . Although, many methods (both parametric and non-parametric) exist for estimating the survival function, in this example, the Kaplan-Meier method will be used.

A plot of the Kaplan-Meier estimate by time (known as the Kaplan-Meier plot) is often used to visualize the behavior of the survival function as a series of downward steps. To illustrate common concepts in survival analysis, an example of a Kaplan-Meier plot for two groups is provided in Figure 1:

Figure 1. Illustrative example of Kaplan-Meier plot



In this illustrative example, the survival function for patients in Group B falls more rapidly than that for patients in Group A, indicating that patients in Group B generally overall have shorter survival times. The median survival time, which is an important measure in clinical trials, is the point in time where the survival function reaches the value of 0.5. Median survival is defined as long as at least half of the observed patients die before the end of follow-up. If this point in time never occurs in a clinical study, then median survival time cannot be calculated.

The mean survival time is defined as the area under the survival curve. It is important to state that, as is common with clinical studies, not every patient experiences the event during the time of observation. In order to calculate a true mean survival time, we would need to observe death in every single patient, whenever that may occur. That is, we would need to calculate the area under the curve after the survival function has reached 0. For this reason, mean survival time is typically not an outcome of interest in clinical trials. In outcomes research and health economics studies, however, mean survival up until a certain time point may be of interest. In this case, the measure is known as the restricted mean survival time because the calculation is restricted to observations until that certain point in time. The restricted mean is calculated as the area under the survival function from time 0 until the given time point of interest. A common time point of interest is the overall median follow-up time across all patients.

Another useful tool that can be used when analyzing time to event data, assuming that certain assumptions are met, is the Cox Proportional Hazards model. Details of this model have been explained elsewhere and beyond the scope of this paper. However, this model yields a measure known as the hazard ratio (HR), which characterizes the relative likelihood of experiencing the event. If the HR is equal to 1, the likelihood of experiencing the event is equal between treatment groups. If the HR is greater than 1, an increase risk exists in the study group relative to the referent group. If the HR is less than 1, a decreased risk exists in the study group relative to the referent group.

HEALTH-RELATED QUALITY OF LIFE²

Health-related quality of life (HR-QoL) is a perceived attribute of general well-being as it relates to a person's health status. A multitude of instruments aimed at measuring HR-QoL have been developed and range from disease-specific assessment to overall general assessment.

A commonly used instrument for measuring general HR-QoL is the EuroQol (EQ-5D). Details of the EQ-5D have been previously described.² In brief, the EQ-5D consists of 5 questions, each related to a particular domain of HR-QoL (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). After applying weight-based algorithms, an overall summary score of HRQoL, known as the EQ-5D utility index, can be calculated. This utility index score has a maximum value of 1, which represents full health. A value of 0 represents death. Negative values are possible (the minimum value is -0.59), which describe health states that are perceived as worse than death.

For this study, the EQ-5D utility index score will be used to assess HR-QoL.

QUALITY-ADJUSTED SURVIVAL ANALYSIS^{3,4,5}

Quality-adjusted survival analysis (QASA) is an analytic tool that utilizes survival analysis techniques while adjusting for HR-QoL, so that the quantity versus quality of life trade-off can be quantified into summary measures. Because HR-QoL data may or may not be collected or available during the course of follow-up in many clinical studies, several analytic approaches exist that implement different assumptions of health state trajectories and disease progression (e.g., constant quality of life assumption, quality-adjusted time without symptoms and toxicity).^{3,4}

For this study, we will be analyzing the case where repeated HR-QoL measurements at intervals are available over the entire course of the follow-up (requiring the least amount of assumptions regarding health state and disease progression).

An important concept and routinely used summary measure in health economic evaluation is the quality-adjusted life year (QALY).⁵ The QALY is the primary outcome of interest in QASA and is analogous to a quality-adjusted mean survival time. QALYs can be thought of as years of life weighted by quality of life, and they are calculated by multiplying the years of life spent in a health state by the HR-QoL weight of that particular health state. Therefore, 1 year spent in full health, 2 years spent at 50% health, and 4 years spent at 25% health are each equivalent to 1 QALY. Utility scores (such as the EQ-5D utility score) provide particularly good HR-QoL weights because they are anchored at 1 and 0 to represent full health and death, respectively.

This approach of weighting survival time by HR-QoL is used to construct the quality-adjusted survival plots, which, in turn can be used to calculate QALYs. To construct the quality-adjusted survival plots, we simply multiply the Kaplan-Meier survival estimates obtained in the survival analysis by the group-specific mean EQ-5D utility scores applicable to the given points in time. The QALY is then calculated as the area under the quality-adjusted survival curve:

$$QALY_L = \int_0^L S(t)Q(t)dt \quad (\text{Equation 3})$$

where $S(t)$ is the overall survival function and $Q(t)$ is quality of life over time. Because the QALY is effectively the mean survival time adjusted for quality of life, its calculation is commonly restricted up to a certain point in time, just as in the calculation of mean survival time. This restricted point in time is denoted by L in Equation 3. Rather than using years as the unit of time, it is also possible to use days or months in which case the QALY becomes quality-adjusted mean survival time in the unit of choice.

In situations where HR-QoL is assumed to remain relatively constant across the observation period, the $Q(t)$ in Equation 3 can be pulled outside of the integral, and repeated-measures models can be implemented to calculate an overall global mean HR-QoL for each treatment group. In these situations, the survival functions for each treatment group are ultimately multiplied by a constant representing mean HR-QoL. For this study, however, we will not be assuming constant HR-QoL over time.

It is important to note that, unlike the overall survival function, the quality-adjusted survival plot may not necessarily be non-increasing. In situations where HR-QoL may increase over time, the quality-adjusted survival plot may have sudden increases at points in time where the HR-QoL increases.

EXPLORATORY METHOD FOR QUALITY-CORRECTED MEDIAN SURVIVAL

Because overall mean survival times are generally not calculated in clinical trials and because QALYs are analogous to quality-adjusted restricted mean survival times, the results of QASA are not directly comparable to the primary results of clinical trials (which compare median survival times between groups). Currently, no standard method exists for comparing median survival times after adjusting for quality of life.

In an effort to make QASA results directly more comparable to primary outcomes in clinical trial results, an exploratory method for calculating quality-adjusted median survival will be proposed.

To illustrate this exploratory method, we will assume two treatment groups labeled as Group A and Group B. We will also assume that Group B has a shorter median survival time than Group A (as in the illustrative example in Figure 1). We propose first to calculate a survival-weighted mean HR-QoL score for Group A and for Group B. This value would be the ratio of the quality-adjusted mean survival time to the overall (unadjusted) mean survival time for the particular group. Regardless of group, the calculation would be restricted to the median survival time for Group B (denoted as t_{MedB}). The ratio of the survival-weighted mean HR-QoL scores between treatment groups (Group B to Group A) represents a quality "correction" factor (denoted as QCF). Hence, the QCF is calculated as follows:

$$QCF = \frac{\int_0^{t_{MedB}} S_B(t) Q_B(t) dt}{\int_0^{t_{MedB}} S_B(t) dt} \bigg/ \frac{\int_0^{t_{MedB}} S_A(t) Q_A(t) dt}{\int_0^{t_{MedB}} S_A(t) dt} = \frac{\left[\int_0^{t_{MedB}} S_B(t) Q_B(t) dt \right] \left[\int_0^{t_{MedB}} S_A(t) dt \right]}{\left[\int_0^{t_{MedB}} S_A(t) Q_A(t) dt \right] \left[\int_0^{t_{MedB}} S_B(t) dt \right]} \quad (\text{Equation 4})$$

By then multiplying the median survival time for Group B by the QCF, the result can ultimately serve as a quality-corrected median survival time for Group B. This quality-corrected median survival time can serve a reference point to which we can compare the Group A median survival time. Then, we can ask whether the difference between the Group A median survival time and the Group B quality-corrected median survival time is significantly different than 0. That is,

$$t_{MedA} - [(QCF)(t_{MedB})] = 0 \quad (\text{Equation 5})$$

where t_{MedA} is the Group A median survival time, t_{MedB} is the Group B median survival time, and QCF is the quality of life correction factor defined in Equation 4. The term "quality correction" is introduced here to indicate a correction for the relative difference in HR-QoL between the two groups. This term differs from "quality adjustment," which always involved a multiplication factor less than or equal to one.

The difference expressed in Equation 5, along with confidence intervals can be calculated through bootstrap methods. This exploratory method can provide insight as to whether median survival times between treatment groups are different after adjusting for HR-QoL.

For instance, suppose that the median survival times for Groups A and B were 10 months and 8 months, respectively, and that their survival-weighted mean HR-QoL scores at 8 months (which is the median survival time for Group B) were 0.8 and 0.9, respectively. The QCF would be equal to $0.9 / 0.8 = 1.125$. By multiplying the median survival time for Group B by the QCF, we obtain a quality-corrected median survival of 9 months (8×1.125). Hence, the 2 month observed difference in median survival between groups becomes a 1 month "corrected" difference after accounting for HR-QoL.

METHODS

In order to illustrate the use of QASA, an oncology clinical trial example will be used. To avoid any suggestion that the author of this paper is making independent clinical inferences about the particular disease indication or about the performance of commercially available treatments, the generic term "cancer" will be used to denote the disease, and the generic terms "treatment" and "placebo" will be used to denote the two possible therapeutic regimens. Furthermore, rather than using actual clinical trial data, a simulated dataset will be generated to mimic the general overall attributes of actual clinical trial results.

DATA SOURCE AND SIMULATION

Data were simulated for 1000 cancer patients, 500 to receive the treatment and 500 to receive the placebo. Patients were followed for 60 months or until censoring. Study visits where quality of life was assessed were specified at study entry (baseline) and at the following time points after baseline: Months 1, 6, 12, 18, 24, 30, 36, 42, 48, 54, and 60.

Survival times and censoring times in months were simulated through random number generation following the Weibull distribution by using the RAND function in SAS®. For patients in the treatment group, the shape parameter was specified as 1.71, and the scale parameters were specified as 24 for patients who died before the end of follow up (n=251) and 28 for patients who were censored before the end of follow up (n=149). For patients in the placebo group, the shape parameter was specified as 1.03, and the scale parameters were specified as 16 for patients who died before the end of follow up (n=262) and 23 for patients who were censored before the end of follow up (n=230). The remaining patients in each group were observed for the entire study duration and censored at 60 months.

Target mean HR-QoL scores, as assessed by the EQ-5D utility, were specified. Variation around these target HR-QoL scores for each patient and visit were simulated through random number generation following the uniform distribution by using the RANUNI function in SAS®. Assumptions in the simulation included a variation of ± 0.15 points around the target means and that patients who experienced death at some point during follow-up would have on average lower HR-QoL than patients who were censored by approximately 0.10 points.

The SAS® code use to generate the simulated dataset is provided in Appendix A.

SURVIVAL ANALYSIS AND HR-QoL

Overall survival analysis was performed using two procedures in SAS®. PROC LIFETEST was used to generate Kaplan-Meier estimates and plots, along with associated summary statistics such as survival time estimates (25th percentile, median, and 75th percentile) and the log-rank test. PROC PHREG was used to construct a Cox Proportional Hazards model, which generated a hazard ratio and 95% confidence interval:

```
proc lifetest data=surv01 method=km;
  strata tx;
  time t_mo*death(0);
run;

proc phreg data=surv01;
  class tx;
  model t_mo*death(0) = tx / risklimits;
run;
```

In the above code, tx denotes treatment group (1=Treatment, 2=Placebo), t_mo denotes time in months, and death denotes an indicator variable for death (1=died, 0=did not die). In PROC PHREG, the risklimits option in the MODEL statement indicates the calculation of the hazard ratio and 95% confidence interval.

Mean HR-QoL scores, as measured by the EQ-5D utility, were plotted over time by treatment group. The following SAS® code was used:

```
proc gplot data=qol02;
  plot qol*month=tx / vaxis=axis1 haxis=axis2 legend=legend1;
  legend1 position=(inside bottom center) mode=protect across=1 frame
    offset=(+29,+29) label=(position=top j=c "Therapy") value=(j=1);
  axis1 order=0 to 1 by .1 minor=(number=1) label=(angle=90 "Mean EQ-5D Utility Score");
  axis2 order=0 to 60 by 6 minor=none offset=(0) label=("Months of Follow-Up");
  symbol1 line=1 i=j w=9 value=dot h=4 color='black';
  symbol2 line=1 i=j w=9 value=dot h=4 color='light gray';
  format month 6.0 qol 6.2 tx ftx.;
run;
```

QUALITY-ADJUSTED SURVIVAL ANALYSIS

Survival data were merged with HR-QoL data, and through a series of data steps, a dataset was created at the level of treatment group (values 1 for treatment patients and 2 for placebo patients) and day of follow-up (ranging from 0 to 1826, which corresponds to 60 months of follow-up). Hence, the analytic dataset contained 3654 observations.

Each observation contained the Kaplan-Meier survival function estimate and mean HR-QoL at that point in time. The product of these two variables yielded the quality-adjusted survival function estimate for each observation. A macro was created to calculate the area under the curve from time 0 to the specified time point of interest using the trapezoidal method. This macro was applied to both the survival function and quality-adjusted survival function.

Quality-adjusted Kaplan-Meier survival plots by treatment group were generated. On the same figure, the original (unadjusted) overall survival plots by treatment group were overlaid so that visual comparisons could be made.

Overall mean survival times and quality-adjusted mean survival times were calculated for the following time points of interest: overall median follow-up time across all patients (20.7 months or 631 days) and end of study (60 months or 1826 days). Bootstrap methods (1000 samples with replacement) were used to calculate 95% confidence intervals around the estimates and p-values assessing the difference between quality-adjusted mean survivals at the specified time points of interest.

As an exploratory method of assessing quality-adjusted median survival, the QCF was calculated (Equation 4) and the difference between the Group A median survival time and the quality “scaled” median survival time was calculated (Equation 5). Bootstrap methods (1000 samples with replacement) were used to calculate the 95% confidence interval and p-value.

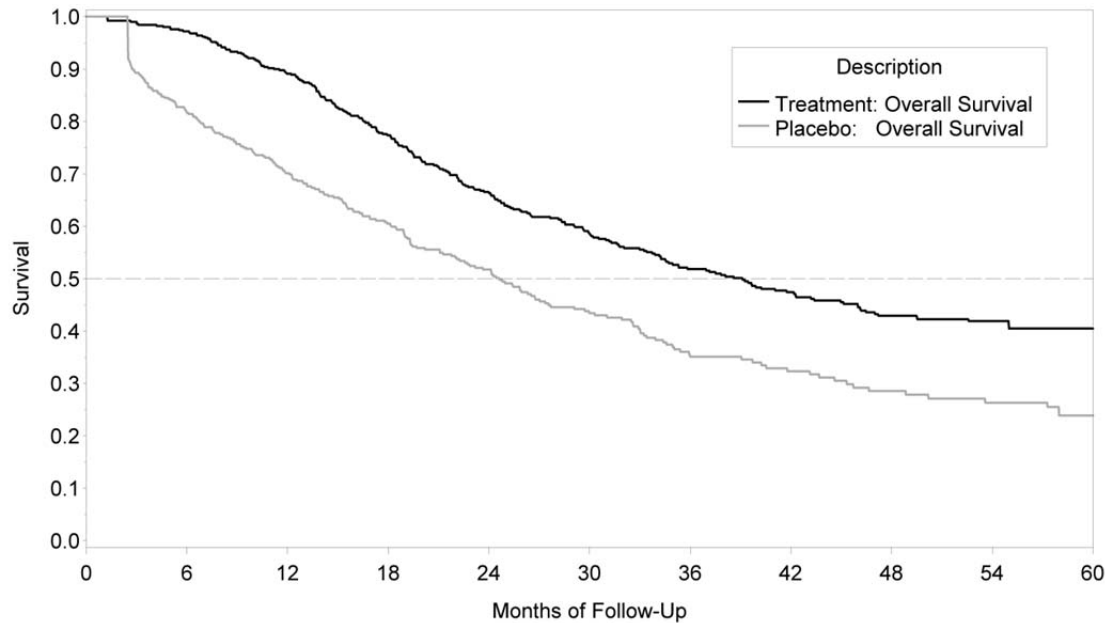
SAS® code for generating the analytic dataset and conducting the QASA can be obtained from the author.

RESULTS

SURVIVAL ANALYSIS RESULTS

Kaplan-Meier plots for patients in the treatment and placebo groups are depicted in Figure 1. From visual inspection, the overall survival function generally appeared greater throughout the course of follow-up for treatment patients than for placebo patients. Moreover, for placebo patients a relatively steep decline in the survival function was observed early in follow-up (at around the 3 to 4 month mark) followed by a progressively slower rate of decline throughout the course of the study. For treatment patients, in contrast, the rate of decline in the survival function appeared slower at the beginning of follow-up, then became slightly more rapid for a period of time, then progressively slowed until the end of observation. Another visual detail of interest was that the overall survival function met the 0.5 mark (denoting median survival time) approximately 14 to 15 months earlier in time for placebo patients compared to treatment patients.

Figure 2. Kaplan-Meier plots of overall survival for study population



Associated survival analysis statistics are summarized in Table 1. The numbers and percentages of patients who experienced death, who were censored during the course of follow-up, and who were followed until the end of the study were tabulated, as specified in the dataset simulation. The median survival times for placebo and treatment patients were 24.7 months (95% CI: 21.1, 27.6 months) and 39.1 months (95% CI: 33.7, 45.2 months), respectively. The widths of the 95% confidence intervals around medial survival time estimates suggested a significant difference between study groups favoring treatment patients.

The log-rank test yielded a p-value of <0.001, which provided further evidence the survival functions were different between treatment groups. Another observation of interest is that the 25th percentile survival times (roughly the point in time where 25% of the patients died) were 9.4 months and 19.1 months for placebo and treatment patients, respectively. The 75th percentile survival time for placebo patients was 58.0 months, while this value was undefined for treatment patients because 75% of treatment patients had not yet died at the time of study end.

Cox Proportional Hazards model results yielded a hazard ratio of 0.581 (95% CI: 0.488, 0.692). These results indicate that the instantaneous likelihood of death for treatment patients is 0.581 times that of placebo patients. Hence, a significant protective effect from death is observed among treatment patients relative to placebo patients.

Table 1. Overall survival analysis statistics of study population

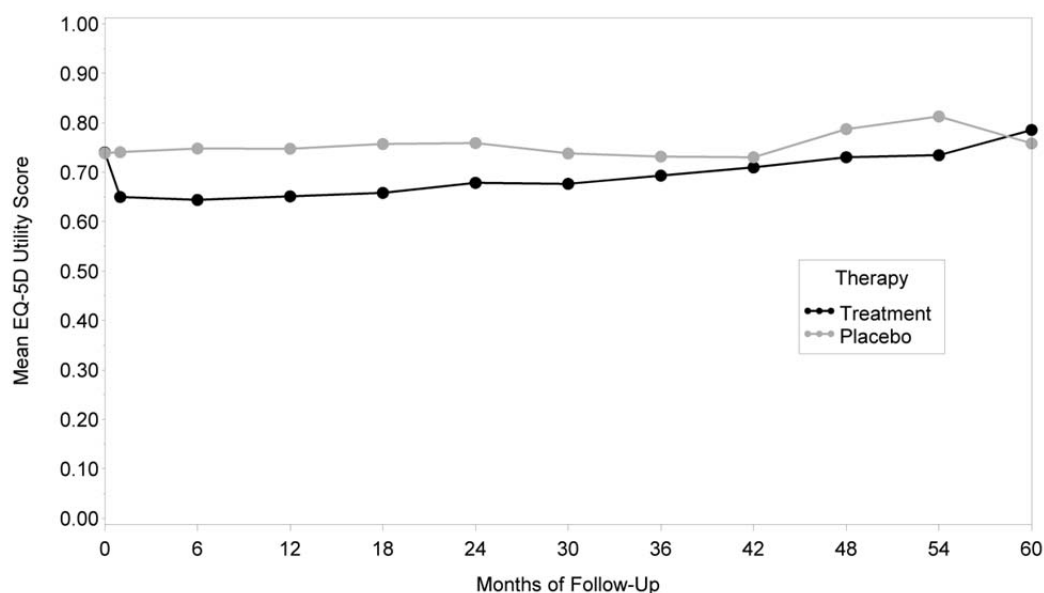
	Placebo (N = 500)	Treatment (N = 500)
Number of subjects, n (%)		
Died (event)	262 (52%)	251 (50%)
Censored before end of study	230 (46%)	149 (30%)
Followed until end of study	8 (2%)	100 (20%)
Estimates for overall survival (months)		
25th Percentile (95% CI)	9.4 (7.1, 11.7)	19.1 (17.0, 21.1)
Median (95% CI)	24.7 (21.1, 27.6)	39.1 (33.7, 45.2)
75th Percentile (95% CI)	58.0 (45.4, NR)	NR (NR, NR)
Log-rank test		
P value	<0.001	
Hazard ratio		
Estimate (95% CI)	0.581 (0.488, 0.692)	

NR = Not reported because value was undefined.

HEALTH-RELATED QUALITY OF LIFE RESULTS

Patients in both treatment groups initiated the study with very similar mean HR-QoL scores: 0.740 for treatment patients and 0.739 for placebo patients. At Month 1 follow-up, mean HR-QoL decreased sharply for treatment patients (score = 0.650), while it remained relatively constant for placebo patients (score = 0.741). Over the subsequent months of follow-up, mean HR-QoL scores for treatment patients decreased a bit further, hit a minimum of 0.644 at Month 6, slowly climbed with fluctuation to 0.736 at Month 54, then ended with a score of 0.786 at Month 60. For placebo patients, mean HR-QoL scores remained relatively constant over the entire course of follow-up with fluctuation, hitting a minimum of 0.735 at Month 30 and a maximum of 0.823 at Month 54.

Graphical representation of mean HR-QoL scores over time by treatment group are presented in Figure 3. Actual mean scores by treatment group and study visit are available in the data simulation code in Appendix A.

Figure 3. Health-related quality of life over time by treatment group

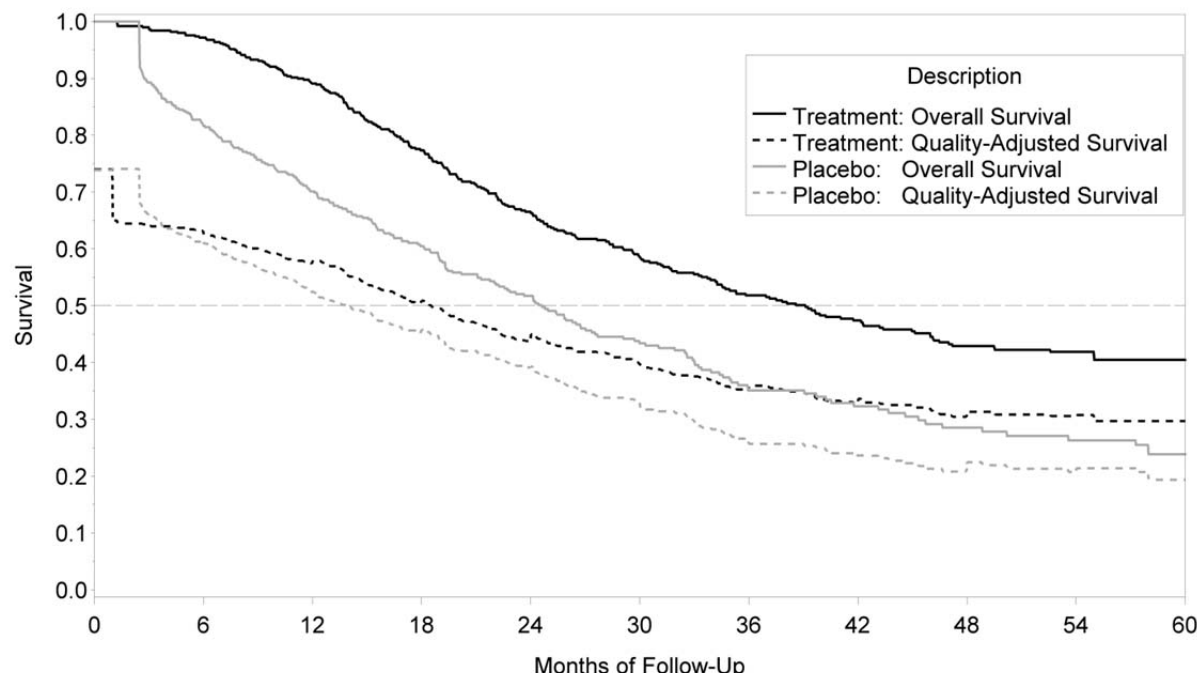
QUALITY-ADJUSTED SURVIVAL ANALYSIS RESULTS

Quality-adjusted Kaplan-Meier plots by treatment group are depicted in Figure 4 (also overlaid with the unadjusted Kaplan-Meier plots). The unadjusted survival functions are depicted with solid lines while the quality-adjusted survival functions are

depicted with dashed lines. The quality-adjusted functions begin at approximately 0.740 at time 0, which corresponds to the respective HR-QoL values at study entry. From visual inspection, the quality-adjusted survival function appears to be generally higher for placebo patients early in follow-up until approximately Month 5 at which point the curves appear to cross. After this point, the quality-adjusted survival function appears generally greater for treatment patients; however the distance between the two curves appears much smaller in magnitude than that between the unadjusted survival curves.

Table 2 summarizes the results of mean survival time and quality adjusted mean survival time at two restricted time points of interest: overall median follow-up (20.7 months) and at end of study (60 months). At median follow-up (20.7 months), treatment patients had an overall mean survival time of 18.5 months, while placebo patients had a value of 15.5 months, which yielded a difference of 3.0 months between groups (95% CI: 2.3, 3.8). After adjusting for HR-QoL, the adjusted mean survival times were 12.1 months and 11.6 months for treatment and placebo patients, respectively, yielding a difference of 0.5 quality-adjusted months between groups (95% CI: -0.0, 1.2).

Figure 4. Quality-adjusted (and unadjusted) Kaplan-Meier plots of overall survival for study population



At the end of study (60 months), treatment patients had an overall mean survival time of 38.5 months, while placebo patients had a value of 29.4 months, which yielded a difference of 9.0 months between groups (95% CI: 6.3, 12.0). After adjusting for HR-QoL, the adjusted mean survival times were 26.0 months and 22.2 months for treatment and placebo patients, respectively, yielding a difference of 3.8 quality-adjusted months between groups (95% CI: 1.6, 6.2).

Table 2. Quality-adjusted (and unadjusted) mean survival at restricted time points

Restricted time point	Group	Mean survival in months (95% CI)	Quality-adjusted mean survival in months (95% CI)
Median follow-up (20.7 months)	Treatment (TX)	18.5 (18.1, 18.9)	12.1 (11.8, 12.4)
	Placebo (PL)	15.5 (14.9, 16.1)	11.6 (11.1, 12.1)
	Difference (TX - PL)	3.0 (2.3, 3.8)	0.5 (-0.0, 1.2)
	P value	<0.001	0.052
End of study (60 months)	Treatment (TX)	38.5 (36.6, 40.4)	26.0 (24.6, 27.5)
	Placebo (PL)	29.4 (27.3, 31.8)	22.2 (20.4, 23.9)
	Difference (TX - PL)	9.0 (6.3, 12.0)	3.8 (1.6, 6.2)
	P value	<0.001	<0.001

It is important to note that the choice of restricted time point may be arbitrary or based on a particular point of interest. When this time point is later in the course of follow-up, the area under the curve calculations may become less stable. Due to death and loss to follow-up, the effective sample sizes steadily decrease over time. Therefore, the behavior of the Kaplan-Meier survival functions may be erratic at the tails. Parametric methods (which were not used in this study) are commonly used to estimate survival functions in an effort to address the potentially unstable behavior later in follow-up.

EXPLORATORY METHOD FOR QUALITY-CORRECTED MEDIAN SURVIVAL RESULTS

Results of the exploratory method of assessing quality-adjusted median survival are presented in Table 3. The overall (unadjusted) mean survival times for placebo and treatment patients were 17.6 and 21.2 months, respectively, while the quality-adjusted mean survival times were 13.2 and 13.9, respectively. These values corresponded to survival-weighted HR-QoL scores of 0.75 in placebo patients and 0.66 in treatment patients, yielding a QCF of 1.14. After multiplying the median survival time for placebo patients by the QCF, the quality-corrected median survival for placebo patients was 28.2 months (24.7×1.14). This quality correction yielded a difference in median survival between groups of 10.9 months (95% CI: 10.6, 11.4); whereas, the difference in observed median survival was 14.4 months. Hence, the exploratory method suggested that the “gap” between median survival times would be shortened by approximately 3.5 months after accounting for quality of life.

Table 3. Exploratory method of quality-corrected median survival (all times in months)

	Placebo (PL)	Treatment (TX)
Overall (unadjusted) mean survival ^a (95% CI)	17.6 (16.8, 18.4)	21.2 (20.7, 21.7)
Quality-adjusted mean survival ^a (95% CI)	13.2 (12.5, 13.8)	13.9 (13.5, 14.3)
Survival-weighted HR-QoL	0.75	0.66
Quality correction factor [QCF] (95% CI)	1.14 (1.13, 1.16)	
Observed median survival	24.7	39.1
Difference in observed median survival [TX - PL]	14.4	
Quality-corrected median survival [change in PL only]	28.2	39.1
Difference in corrected median survival [TX - PL] (95% CI)	10.9 (10.6, 11.4)	

^a. Restricted at median survival time for placebo patients (24.7 months)

95% confidence intervals obtained from 1000 bootstrap samples.

CONCLUSION

Quality-adjusted survival analysis (QASA) is a useful method for determining an overall summary measure that combines survival and quality of life. Through analysis of a simulated oncology study (in which the results mimicked those of an actual clinical trial), topics explored included survival analysis, assessment of HR-QoL over time, construction of quality-adjusted Kaplan-Meier curves, calculation of quality-adjusted life years (QALYs), and the examination of an exploratory method to evaluate quality-adjusted median survival.

The results indicated that while patients on treatment had better survival outcomes, they also had a poorer quality of life profile. After adjusting for HR-QoL through QASA techniques, it was determined that the trade-off between quantity and quality of life still favored treatment patients relative to placebo patients. Currently, no standard method is available for adjusting median survival times for quality of life. Through the proposed exploratory method, further insight can be gained to determine whether median survival times between treatments differ after correcting for HR-QoL. Perhaps this exploratory method can be of benefit when it is desirable to compare quality-adjusted results to the primary results of clinical trials.

REFERENCES

1. Kleinbaum DG. Survival analysis: A self-learning text. Springer Verlag, New York, 1996.
2. Kind P. The EuroQol instrument: an index of health-related quality of life. In: Spilker B (ed). Quality of Life and Pharmacoeconomics in Clinical Trials, 2nd edn. Philadelphia: Lippincott-Raven, 1996.
3. Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. Stat Med 1990 Nov;9(11):1259-76.
4. Glasziou PP, Cole BF, Gelber RD, Hilden J, Simes RJ. Quality adjusted survival analysis with repeated quality of life measures. Stat Med 1998 Jun 15;17(11):1215-29.

5. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010;96:5-21. Epub 2010 Oct 29.
6. SAS Institute, Inc., 2008. *SAS/STAT® 9.2 User's Guide*. Cary, NC: SAS Institute Inc.

ACKNOWLEDGMENTS

I would like to acknowledge and thank Dave P. Miller and David J. Pasta for their review of the paper.

CONTACT INFORMATION

The author welcomes questions and comments. Please direct inquiries to

Lawrence Rasouliyan, MPH
Senior Research Manager, Medical Affairs Statistical Analysis
ICON Clinical Research
Lawrence.Rasouliyan@iconplc.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.

APPENDIX A – SAS® CODE USED TO GENERATE THE SIMULATED DATASET

```
* Note that the RAND function is started with a single seed. However, the state of the
process cannot be captured by a single seed. Hence, an exact duplicate of the dataset used
in the analysis with randomly-generated survival times cannot be re-generated w/ this code.

* Simulate Survival Data. Times in months. Will also calculate time in days in data step;
* Overall length of study = 60 months; %let studylength = 60;
* Conversion from months to days; %let conv = 30.4368499;

* Treatment group (tx=1);          * Placebo group (tx=2);
%let numdeath1 = 251;             %let numdeath2 = 262;
%let numltfu1 = 149;             %let numltfu2 = 230;
%let meanstop11 = 24;            %let meanstop21 = 16;
%let meanstop10 = 38;           %let meanstop20 = 23;
%let minstopdie1 = 1.3;         %let minstopdie2 = 2.5;
%let minstopcen1 = 0.9;         %let minstopcen2 = 0.9;
%let maxstopdie1 = 55;          %let maxstopdie2 = 58;
%let maxstopcen1 = 60;          %let maxstopcen2 = 60;
%let shape1 = 1.71;             %let shape2 = 1.03;

* Survival data simulation. 1000 patients (1/2 in each group);
data surv01;
do patid = 1 to 1000;
  * Treatment group;
  *Died;
  if (patid le &numdeath1.) then do;
    tx=1; death=1;
    t_mo = max(&minstopdie1., (min(&maxstopdie1., (rand('weibull', &shape1., &meanstop11.)))));
    t_dy = int(t_mo*&conv.); output;
  end;
  *Lost to Follow up;
  else if (&numdeath1. lt patid le (&numdeath1.+&numltfu1.)) then do;
    tx=1; death=0;
    t_mo = max(&minstopdie1., (min(&maxstopdie1., (rand('weibull', &shape1., &meanstop10.)))));
    t_dy = int(t_mo*&conv.); output;
  end;
  *Completed study;
  else if ((&numdeath1.+&numltfu1.) lt patid le 500) then do;
    tx=1; death=0;
    t_mo = &studylength.;
    t_dy = int(t_mo*&conv.); output;
  end;
  * Placebo group;
  *Died;
  else if (500 lt patid le (500+&numdeath2.)) then do;
    tx=2; death=1;
    t_mo = max(&minstopdie2., (min(&maxstopdie2., (rand('weibull', &shape2., &meanstop21.)))));
    t_dy = int(t_mo*&conv.); output;
  end;
  *Lost to Follow up;
  else if ((500+&numdeath2.) lt patid le (500+&numdeath2.+&numltfu2.)) then do;
    tx=2; death=0;
    t_mo = max(&minstopdie2., (min(&maxstopdie2., (rand('weibull', &shape2., &meanstop20.)))));
    t_dy = int(t_mo*&conv.); output;
  end;
  *Completed study;
  else if ((500+&numdeath2.+&numltfu2.) lt patid le 1000) then do;
    tx=2; death=0;
    t_mo = &studylength.;
    t_dy = int(t_mo*&conv.); output;
  end;
end;
label patid = "Patient ID"
      tx = "Treatment Group: 1=Treatment, 2=Placebo"
      death = "Event: 1=Death, 0=Not Death"
      t_mo = "Time to event in Months"
```

```

t_dy = "Time to event in Days";
run;

* Simulate QoL (EQ-5D utility) data;
data qol01;
  input visit$ month timecat qtx1 qtx2; format qtx1 qtx2 6.3;
  label visit = "Study Visit"
        month = "Month"
        timecat = "Sequential Time Category"
        qtx1 = "Mean EQ-5D Utility Score for patients in Treatment Group (Tx1)"
        qtx2 = "Mean EQ-5D Utility Score for patients in Placebo Group (Tx2)";
  cards;
BL      0      1      0.741 0.739
M1      1      2      0.651 0.741
M6      6      3      0.644 0.744
M12     12     4      0.648 0.739
M18     18     5      0.651 0.745
M24     24     6      0.661 0.749
M30     30     7      0.652 0.731
M36     36     8      0.655 0.718
M42     42     9      0.664 0.710
M48     48     10     0.681 0.735
M54     54     11     0.684 0.740
M60     60     12     0.735 0.737
;
run;

data qol02 (keep=timecat visit month tx qol);
  set qol01 (in=a) qol01 (in=b);
  if a then do;
    tx=1; qol=qtx1;
  end;
  if b then do;
    tx=2; qol=qtx2;
  end;
  label tx = "Group: 1=Treatment, 2=Placebo"
        qol = "Mean EQ-5D Utility Score";
run;

%let seed1 = 56; %let seed2 = 415;
* Dataset qol03 will contain HR-QoL scores for each observed patient at each study visit;
data qol03;
  set _null_;
run;

%macro eq5ddat (mth=);
data eq5d&mth.;
  merge saslib.surv01 (where=(t_mo ge &mth.))
        qol02 (where=(month eq &mth.));
  by tx;
  plusmin = ranuni(&seed2.);
  if (plusmin le 0.5) then mult=1;
  else if (plusmin gt 0.5) then mult=-1;
  * Assume that on average patients who ultimately die during follow-up will have poorer QoL;
  if (death eq 1) then qolpt = (qol - .05) + (mult*.15*ranuni(&seed1.));
  else if (death eq 0) then qolpt = (qol + .05) + (mult*.15*ranuni(&seed1.));
run;

data qol03 (drop=mult plusmin);
  set qol03 eq5d&mth.;
  label qolpt = "HR-QoL score at study visit";
run;
%mend eq5ddat;

%eq5ddat (mth=0); %eq5ddat (mth=1); %eq5ddat (mth=6); %eq5ddat (mth=12); %eq5ddat (mth=18);
%eq5ddat (mth=24); %eq5ddat (mth=30); %eq5ddat (mth=36); %eq5ddat (mth=42); %eq5ddat
(mth=48); %eq5ddat (mth=60);

```