Journal of Statistical Research *ISSN 0256 - 422 X*
2019, Vol. 53, No. 2, pp. 165-189

INFERENCE ON MEAN QUALITY-ADJUSTED LIFETIME USING JOINT MODELS FOR CONTINUOUS QUALITY OF LIFE PROCESS AND TIME TO EVENT

XIAOTIAN GAO*

Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA, USA

Email: xig31@pitt.edu

XINXIN DONG

Takeda Development Center Americas, Inc., Deerfield, IL,USA

CHAERYON KANG

Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA, USA

ABDUS S. WAHED

Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA, USA

SUMMARY

Quality-adjusted lifetime (QAL) has been considered as an objective measurement that summarizes the quantitative and qualitative health aspects in a unitary and meaningful way. The idea is to account for each individual's health experience in adjusting the overall survival, with death at one extreme and perfect health at the other extreme. In existing literature, the health states are defined to be discrete and the number of states is taken to be finite. Therefore, QAL can be calculated as a sum of time spent at each health state multiplied by the corresponding weight. In this paper we propose an estimator of the mean QAL when the quality of life process is assumed to be continuous and observed with error over time at fixed time points via joint modeling of quality of life and time to event outcomes. We prove the asymptotic properties of the proposed estimator and study its finite sample performance through simulation. We illustrate its application with data from a hepatitis C clinical trial study.

Keywords and phrases: Quality-adjusted Survival; Accelerated Failure Time; Joint modeling; Survival analysis

* Corresponding author

1 Introduction

Quality-adjusted lifetime (QAL), introduced by Goldhirsch et al. (1989), has been considered as an objective measurement that summarizes the quantitative and qualitative health aspects in a unitary and meaningful way. Since its introduction QAL has occupied a significant area of clinical and biomedical research in treatment comparisons, especially for chronic diseases such as cancer, cardiovascular diseases, and acquired immunodeficiency syndrome (AIDS). The idea of QAL is to divide each individual's health experience into several intermediate states, with death at one extreme and perfect health at the other extreme. The time spent in each state is weighted by a utility coefficient ranging from zero to one. The value of the utility coefficient is decided according to clinically specific evaluation of the quality of life (QOL) that state renders. The perfectly healthy state has value of one while the absorbing state equivalent to death has value of zero. This leads to a utility function over time which takes the value of the utility coefficient of the state occupied at that time. QAL is defined as an integration of this utility function over the survival duration.

In existing literature, the health states are defined to be discrete and the number of states is taken to be finite. For finite discrete states, QAL can be calculated as a sum of time spent at each health state multiplied by the corresponding weight and considered as equivalent to time with perfect health (Gelber et al., 1991). Inference about the distribution of QAL in various forms has been investigated by many authors (Korn, 1993; Zhao and Tsiatis, 1997; Laan and Hubbard, 1999; Pradhan and Dewanji, 2009).

In addition to estimating the distribution of QAL, the mean of QAL has also been considered as an important parameter of interest in treatment comparisons. Cole et al. (1993) considered a Cox type parametric regression model to estimate mean QAL using bootstrap variance estimation. Hwang et al. (1996) proposed an estimator of the expected quality-adjusted survival with quality of life information collected from a cross-sectional survey. Huang and Louis (1999) presented a nonparametric approach to estimate the expected quality-adjusted survival time for a general multistate process with incomplete follow-up data. Zhao and Tsiatis (2000) derived a class of consistent and asymptotically normal estimators of mean QAL with right-censored data in a more general setting. Wang and Zhao (2006) developed a regression method to investigate the covariate effects on the mean QAL with right-censored data. Andrei and Murray (2007) proposed regression models for the mean of the QOL adjusted restricted survival time using pseudo-observations. Zhao and Wang (2008) considered the empirical likelihood method for the regression model of mean QAL with right-censored data.

In all previous studies the QOL function was assumed to be a step function, taking constant values at various states. However, in most applications, QOL is measured using continuous scales with scores ranging over some specified intervals. For example, in the study that motivated our research, Study of Viral Resistance to Antiviral Therapy of Chronic Hepatitis C (Virahep-C), quality of life was assessed in the form of symptom assessment via a visual analog scale (VAS) ranging from 0.00 to 10.00 (Bonkovsky et al., 2007). In most cases, a reasonable approach to draw inference on QAL is to utilize the QOL measurements, which are usually collected repeatedly with error during the course of treatment or the study period. This approach will not only make use of more information from continuous QOL scores, but also relax the assumption that individuals' health status remain the same between clinical visits.

In this paper we present an approach for consistent estimation of the mean QAL when the QOL process is observed in continuous scale with error over time at fixed time points. Specifically, we use a utility function to map the QOL scores to an interval $[0, 1]$ in the definition of QAL. The QOL process itself is modeled using a linear mixed model. To improve efficiency of the estimation, we jointly model the longitudinal quality of life

scores and the survival time. We then borrowed the idea from Zhao and Tsiatis (2000) and Robins et al. (1994) to derive a consistent and asymptotically normal estimator for mean QAL and its asymptotic variance, where the QOL score trajectories are estimated from the data using the joint modeling.

The remainder of this article is organized as follows. In Section 2, we provide details of the proposed methods. Asymptotic properties of this estimator are outlined in Section 3 and their finite sample performance is investigated via simulation in Section 4. Finally, we demonstrate an application of the proposed method to a hepatitis C study in Section 5.

2 Estimation of Mean QAL via Joint Modeling

2.1 Notation and Definition of QAL

We introduce some notations used to model the survival and the QOL processes. For ith individual, let T_i^* be the potential event time, such as time to death, time to disease relapse or time to infection, and C_i be the potential right censoring time, such as time to lost to follow-up, or time to the end of study. Let L denote some predetermined restricted length of time after which a reasonable proportion of individuals are still being followed. The restricted survival time of interest is $T_i = \min(T_i^*, L)$. For all i, we denote the observed time and event indicator respectively by $Y_i = \min(T_i, C_i)$ and $\Delta_i = I\{T_i \leqslant C_i\}$.

Suppose $W_i(t)$ represents the dynamic QOL score of ith individual at time t. Let Q be a twice differentiable and monotonic function mapping QOL to interval $[0, 1]$, with 0 representing the worst QOL such as death and 1 representing the best QOL such as the normal health status. Then QAL of ith individual with QOL $W_i(t)$ and the restricted survival time T_i can be defined as

$$U_i = \int_0^{T_i} Q\{W_i(t)\}dt. \tag{2.1}$$

The exact functional form of Q is chosen by experts to reflect their belief on how continuous QOL scores converted to weights of life time. A natural choice for Q is the inverse logit function:

$$Q\{W_i(t)\} = logit^{-1}\{qW_i(t)\}I(t \leqslant T_i), \tag{2.2}$$

where a predefined scale parameter q changes the scale of QOL scores in the QAL function, and $logit^{-1}(x) = 1/(1+e^{-x})$. For the analysis of Virahep-C data with QOL ranging from 0 to 10, we use $Q(x) = \{1 + \exp(x - 5)\}^{-1}$, which is applying (2.2) to centered QOL scores with $q = 1$.

Since the result of QAL analysis depends on the choice of utility, sensitivity analysis has been recommended in the absence of any clear clinical justification of the use of a particular utility function. For example, Gelber et al. (1995) suggested implementing a "threshold utility analysis", that illustrates the treatment comparison results for all combinations of utility coefficient values (weights) for different discrete QOL status. This allows the results of the analysis to be interpreted for individual patients based on clinician-patient consultation on the trade-off between the intervention effectiveness and the tolerance to side effects. Following the same, we propose a sensitivity analysis based on a family of utility functions indexed by q. We consider the following family of utility functions:

$$Q(x, q) = \begin{cases} \{1 + \exp[(qx - 5)/(2q - 1)]\}^{-1} & \text{for } 0.5 < q < 1, \\ \{1 + \exp(qx - 5)\}^{-1} & \text{for } q \geq 1. \end{cases} \tag{2.3}$$

The utility function $Q(\cdot, q)$ in (2.3) maps from [0,10] to an approximate [0,1] monotonically. The utilities with a larger value of q differentiate lower QOL scores more, while those with a smaller value of q differentiate higher QOL scores more (Figure 1). For example, to compare two extreme cases, the utility with $q = 5$ differentiates QOL scores only in the region between 0 and 2, as all scores greater than 2 are mapped to values below 0.01. On the other hand, the utility with $q = 0.6$ differentiates QOL scores only in the region between 7 and 10, as all scores less than 7 are mapped to values above 0.98. Note that these utility functions only work on the QOL scores ranging from 0 to 10, and higher scores suggest worse quality of life. For different support of QOL scores, one can re-scale as appropriate.

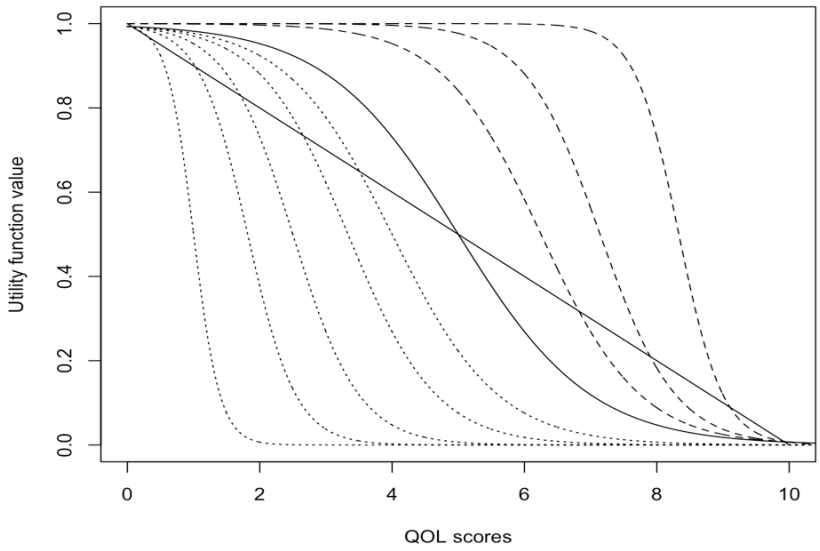

Figure 1: Samples of utility functions indexed by q. Dots curves represent utilities with $q > 1$; solid curve represents $q = 1$; dashed curves represent $q < 1$. From the left to the right, $q = \{5, 2.75, 2, 1.5, 1.25, 1, 0.8, 0.7, 0.6\}$.

2.2 Estimation of the Mean QAL

Our objective is to draw inferences on the mean QAL, defined as

$$\mu = E(U_i) = E\left\{\int_0^{T_i} Q\left(W_i(t)\right) dt\right\}.$$

Estimation of μ requires an exact path of $W_i(t)$ up to time T_i along with the distribution of T_i. If $W_i(t)$, for all $t \leqslant Y_i$ were known for all i, one could estimate μ by using the inverse probability weighted estimator defined

in Zhao and Tsiatis (2000):

$$\hat{\mu}_{ZT} := \frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i U_i}{\hat{K}(T_i)}, \tag{2.4}$$

where U_i is defined in (2.1) and $\hat{K}(\cdot)$ is the Kaplan-Meier estimator of $K(t) = pr(C_i > t)$, the marginal survival function of C_i.

We note that this estimator requires $W_i(t)$ to be known for all uncensored individuals prior to occurrence of the event. In practice, however, $W_i(t)$ could only at best be measured at discrete time points. In other words, even if all individuals attend all scheduled visits, we would still be missing information in between visits. Furthermore, in reality, some individuals could have missed some visits in between, or dropped out of the study. Therefore, it is important to model the QOL process as a continuous manner for inference on the mean QAL from unbalanced and censored data.

2.3 Estimation of QOL Process via Joint Models

Joint modeling of longitudinal and survival data, originally proposed by Wulfsohn and Tsiatis (1997), has been frequently used in biomedical research. Joint modeling of QOL and survival process is intuitively appealing due to the fact that QOL of an individual is inherently related to the length of the individual's life span, and thus it will potentially lead to more reasonable and efficient inference. We first specify a linear mixed model for the QOL process based on the observed QOL data and covariates. The QOL process $W_i(t)$ is only observable with errors at finite time points, say at $\{t_{ij}\}_{j=1}^{n_i}$, for $0 \leqslant t_{ij} < t_{i(j+1)} \leqslant L$. Let us denote these observations by $\{W_{ij}\}_{j=1}^{n_i}$. We use $\{X_i(t), Z_i(t)\}$ to denote the time-dependent covariates that account for the between and within subject variations of QOL. As argued in Tsiatis and Davidian (2004) and Zeng et al. (2005), the key assumption for joint modeling inference is that $\{X_i(t), Z_i(t)\}$ are observable processes, or at least observable at the time when they determine $W_i(t)$ and T_i through a finite dimensional parametric model. In most applications in existing literature, $\{X_i(t), Z_i(t)\}$ include baseline covariates, some deterministic function of time, or the combination of both. The QOL process is modeled using the linear mixed model as follows:

$$W_{ij} = f_M\{X_i(t_{ij}), \beta\} + g\{Z_i(t_{ij}), b_i\} + e_{ij}, \tag{2.5}$$

where $b_i \sim N(0, \Sigma)$, $e_{ij} \sim N(0, \sigma^2)$, $E[g\{Z_i(t_{ij}), b_i\}] = 0$, and $\{X_i(t), Z_i(t)\}$, b_i, e_{ij}, are pairwise independent for all t. Here $f_M\{X_i(t), \beta\}$ is the mean QOL trajectory of individuals with same covariate processes $X_i(t)$ (which are assumed to generate the true QOL process via (2.1)), while $g\{Z_i(t), b_i\}$ accounts for the longitudinal correlation within the same individual. Measurement errors e_{ij} are independent white noise. This model not only allows the assessment of change in the QOL process over time using fixed effects, but also accounts for within subject correlation through subject-specific random effects.

The survival process is modeled by the accelerated failure time (AFT) model due to its appealing interpretation and robustness to non-proportionality of hazards. Specifically, for the ith individual with baseline covariates V_i related to survival outcome, conditional hazard and survival functions for T_i are modeled respectively as follows:

$$\lambda\left(t|b_i, \tilde{X}_i(t), \tilde{Z}_i(t), V_i\right) = \lambda_o(t)\exp[\alpha \cdot \{f_M(X_i(t), \beta) + g(Z_i(t), b_i)\} + \eta V_i],$$

$$S\left(T_i|b_i, \tilde{X}_i(T_i), \tilde{Z}_i(T_i), V_i\right) = \exp\Big[-\int_0^{T_i}\lambda_o(u)\exp\{f_M(X_i(u), \beta) + g(Z_i(u), b_i) + \eta^T V_i\}du\Big],$$

where $\tilde{X}_i(t)$ and $\tilde{Z}_i(t)$ denote the histories of covariate processes up to time t, and $\lambda_o(t)$ denotes the baseline hazard function with parameter Γ. We denote the whole parameter by $\theta^T = (\beta^T, vec^T(\Sigma), \sigma, \Gamma^T, \eta^T, \alpha)$. Following Wulfsohn and Tsiatis (1997), the complete likelihood function is given by

$$\sum_i \log \left[\left\{ S\left(T_i|b_i, \tilde{X}_i(T_i), \tilde{Z}_i(T_i), V_i\right) \right\} \left\{ \lambda\left(T_i|b_i, \tilde{X}_i(T_i), \tilde{Z}_i(T_i), V_i\right) \right\}^{\Delta_i} \right.$$
$$\left. \times\, f_b(b_i)(2\pi\sigma^2)^{-n_i/2} \prod_j \exp\{-|W_i(t_{ij}) - f_M(X_i(t_{ij}), \beta) - g(Z_i(t_{ij}), b_i)|^2/2\sigma^2\} \right], \tag{2.6}$$

where $f_b(b_i) = (2\pi)^{-d_b/2}\det(\Sigma)^{-1/2} \exp\left(-b_i^T\Sigma^{-1}b_i/2\right)$, and d_b denotes the dimension of b_i. As b_i is not observed, (2.6) is maximized using the Expectation-Maximization (EM) algorithm. Given the form of the functions λ_o, f_M, and g, it is straightforward to derive or approximate (via Gaussian Quadrature) the score equation and Hessian of (2.6) with respect to θ. Further details could be found in Wulfsohn and Tsiatis (1997) and Rizopoulos (2012).

Once the parameters of the QOL model in (2.6) are estimated, the QAL of ith non-censored individual can be estimated by

$$U_i(\hat{\beta}) = \int_0^{T_i} Q\left(\hat{W}_i(t)\right) dt = \int_0^{T_i} logit^{-1}\{q f_M(X_i(t)); \hat{\beta}\} dt,$$

where $\hat{\beta}$ is the maximum likelihood estimator of β from model (2.6). Finally the mean QAL is estimated by

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i U_i(\hat{\beta})}{\hat{K}(T_i)}. \tag{2.7}$$

3 Asymptotic Properties of the Estimator

Throughout this paper, Q takes the form of (2.2) with a prespecified value of q, unless otherwise specified. The same approach can be used for other Q functions. For simplicity, we now denote $Q\{q f_M(X_i(t); \beta)\}$ as $f(X_i(t), \beta)$. We note that $f(X_i(t), \beta) \in [0, 1], \forall\, t$ and β. We use $\dot{f}(X_i(t), \beta)$ and $\ddot{f}(X_i(t), \beta)$ to denote the score vector and Hessian matrix of $f(X_i(t), \beta)$ with respect to the vector β evaluated at $(X_i(t), \beta)$, respectively. Euclidean norm is denoted as $\|\cdot\|$. We only state the results in this section, and all technical details are deferred to Appendix, Section A, B, and C.

Theorem 1. *Let $\hat{\mu}_n$ be $\hat{\mu}$ defined as in* (2.7) *and $\hat{\beta}_n$ be $\hat{\beta}$ defined in the previous section. If all the following conditions are satisfied, then $\hat{\mu}_n$ converges to μ in probability.*

- *A1. $\hat{\beta}_n$ converges to β_o in probability.*
- *A2. For t$\in [0, L]$ and some constant C, $pr(\sup_i \|X_i(t)\| < C) = 1$.*
- *A3. For arbitrarily fixed α with $\|\alpha\| < C$, in some open neighborhood B_α around β_o, $\dot{f}(\alpha, \beta)$ exists and is continuous on α ; $\ddot{f}(\alpha, \beta)$ exists and satisfies the Lipchitz condition, that is, for a constant M and arbitrary x and y in B_α, $\|\ddot{f}(\alpha, \beta)x - \ddot{f}(\alpha, \beta)y\| \leqslant M\|x - y\|$.*

The consistency of the maximum likelihood estimator under standard regularity conditions in joint models has been well-established (Tsiatis and Davidian, 2004; Zeng et al., 2005) and A1 of Theorem 1 follows directly

from their results. A2 is reasonable as, $X_i(t)$ denotes all covariate processes, for example, age, height and weight, and is usually bounded uniformly for the population of research interest. A3 follows with Q being a smooth link function (in our case the inverse logit function) and polynomials of predictors.

Theorem 2. *Let λ^c and S be the marginal hazard function of C_i and marginal survival function of T_i, respectively. Under the assumptions A1-A3 in Theorem 1, $n^{1/2}(\hat{\mu}_n - \mu)$ converges in law to a normal distribution with mean 0 and variance σ^2, where $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2$, $\sigma_1^2 = E\{U_i^2(\beta_o)\} - \mu^2$, $\sigma_2^2 = \int_0^L \lambda^c(u) E\big[I(T_i \geqslant u)\{U_i(\beta_o) - G(u)\}^2\big]/K(u)du$, $\sigma_3^2 = \Phi(\beta_o)^T \mathbf{I}^{-1}(\theta_o)\Phi(\beta_o)$, and $\Phi(\beta_o) = E\big\{\int_0^{T_i} \dot{f}(X_i(t), \beta_o)\, dt\big\}$. Here $\mathbf{I}^{-1}(\theta_o)$ is the β component of the inverse Fisher information matrix at θ_o and $G(u) = E\{U_i(\beta_o) I(T_i \geqslant u)\}/S(u^-)$.*

For $N^c(t) = \sum_{i=1}^n I(Y_i \leqslant t\,, \Delta_i = 0)$, the variance σ^2 could easily be estimated by plugging empirical estimators and the Nelson estimator into individual terms of σ. More specifically, we denote

$$\hat{G}(u) = \frac{1}{n\hat{S}(u^-)} \sum_{i=1}^n \frac{\Delta_i U_i(\hat{\beta}_n) I(T_i \geqslant u)}{\hat{K}(T_i)}, \qquad \hat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i U_i^2(\hat{\beta}_n)}{\hat{K}(T_i)} - \hat{\mu}_n^2;$$

$$\hat{\Phi} = \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i}{\hat{K}(T_i)} \int_0^{T_i} \dot{f}(X_i(t), \hat{\beta}_n)\, dt, \qquad \hat{\sigma}_3^2 = \hat{\Phi}^T \cdot \mathbf{I}^{-1}(\hat{\theta}) \cdot \hat{\Phi};$$

$$\hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^n \frac{\Delta_i}{\hat{K}(T_i)} \int_0^L \left[\frac{\{U_i(\hat{\beta}_n) - \hat{G}(u)\}^2}{\hat{K}(u)}\right] \left[\frac{I(Y_i \geqslant u)}{\sum_{i=1}^n I(Y_i \geqslant u)}\right] dN^c(u).$$

Then

$$\hat{\sigma}^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2 + \hat{\sigma}_3^2. \tag{3.1}$$

4 Simulation Study

4.1 Simulation Set-up

We evaluated the performance of the proposed mean QAL estimator (2.7) and its variance estimator (3.1) under various scenarios. The quality of life score, W_{ij}, for individual i at time t_{ij} was generated as follows:

$$\begin{aligned} W_{ij} &= f_M\{X_i(t_{ij}), \beta\} + g\{Z_i(t_{ij}), b_i\} + e_{ij} \\ &= (\beta_0 + \beta_2 X_i) + (\beta_1 + \beta_3 X_i) t_{ij} + b_{0i} + b_{1i} t_{ij} + e_{ij}\,, \end{aligned} \tag{4.1}$$

where $X_i \sim$ Bernoulli(0.5), $b_i \sim N(0, \Sigma)$, $e_{ij} \sim N(0, \sigma^2)$, and b_i, e_{ij} and X_{ij} were pairwise independent. We set $\sigma = 1$ and $\Sigma = \text{diag}(1, 0.5)$. Variables b_i and e_{ij} were truncated to 2 standard deviations. Each individual had up to seven measurements with individual time $t_i = \{t_{ij} : t_{ij} = j, j = 0, 1, \cdots, 6\}$.

We used the following AFT model to simulate the survival outcome:

$$\lambda(t|b_i, X_i, V_i) = \lambda_o(t) \exp\left[\alpha\{f_M(t; \beta, X_i) + g(t; \beta, b_i)\} + \eta V_i\right], \tag{4.2}$$

where $\{b_i, X_i\}$ was the same as in (4.1) to associate the survival outcomes with quality of life scores. The parameter α was taken to be -0.01. In (4.2), λ_o was the baseline hazard function of a Weibull distribution with the shape and scale parameters as $a = 1.5$ and $b = 5$, respectively. The vector of covariates V_i consisted

of constant 1 and a variable generated from a standard normal distribution. The truncation time L was the third quartile of the survival time to ensure more than a small proportion of individuals was followed after L. The censoring time C_i was generated from a Weibull distribution, where the parameters were chosen to ensure target censoring proportions.

The true value of the mean QAL was calculated by generating a large population data set (N $= 10^7$) with uncensored survival outcomes as described above. The true value of mean QAL for a restricted time period $[0, L]$ was calculated as the average of individual QAL defined as $U_i = \int_o^{T_i} logit^{-1}\{qf_M(t; \beta, X_i)\}dt$ with a predefined scale parameter of $q = 0.1$, although sensitivity analysis used various q values.

We considered multiple scenarios that comprised a combination of the following factors: (i) censoring proportion, (ii) regression coefficients (β), (iii) sample size, and (iv) Q functions. We considered two levels of censoring: low(20%) and moderate (40%) censoring. The regression coefficients $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ were chosen to compare the performance of the proposed estimator in the presence of two subgroups (categorized by X) with different trend of QOL: (i) two subgroups had QOL in the opposite direction ($\beta = (2, -3, 1, 4)$), (ii) both subgroups had increasing QOL in the same direction while the slopes were different ($\beta = (7, 3.5, 2.5, 1)$). We considered sample sizes of 200, 400, or 800. Functions Q were chosen to be inverse logit as in (2.2) or a simple linear function. Parameters in the joint model were estimated using the R package JM (Rizopoulos, 2010).

We compared our estimator to the inverse probability weighted (IPW) estimator proposed by Zhao and Tsiatis (2000), denoted by $\hat{\mu}_{ZT}$. We note that this IPW estimator was developed with a discrete and known QOL process, and thus it is not applicable to the continuous QOL setting. For comparison purpose, therefore, we computed the IPW estimator at the true value of β (β_o), by $\hat{\mu}_{ZT}^{(o)} = n^{-1}\sum_{i=1}^{n}[\{\Delta_i U_i(\beta_o)\}/\hat{K}(T_i)]$ with a known QOL. Therefore, the estimator $\hat{\mu}_{ZT}^{(o)}$ could be treated as the oracle estimator of the mean QAL among all estimators with the same inverse probability weights. Note that in practice, this estimator would be unavailable because true QOL will be unknown. The R code of the simulation study is available upon request to the authors.

4.2 Results of the Simulation Study

In summary, the finite sample performance of the proposed method was similar to that of the optimal IPW estimator. Table 1 shows the results of simulations for two subgroups with QOL trends in the opposite direction with function Q chosen to be (2.2). Bias, Monte Carlo standard error (MCSE), Monte Carlo mean of standard error estimate (ESE) and coverage probability of the 95% confidence interval were summarized over 1,000 Monte Carlo samples. Both the optimal and the proposed estimators had negligible absolute biases ($\leqslant 0.003$) and comparable variabilities and empirical coverage probabilities. The coverage probabilities of the 95% confidence intervals for the proposed estimator were between 93.5% and 96.5%, compared to 94.5% to 96% of those from the optimal estimator. The standard error estimates for the proposed estimator closely matched the MC standard errors. The accuracy and the precision of the mean estimates improved with larger sample sizes and smaller censoring proportions, as evidenced by the decrease bias and difference between MCSE and ESE.

We observed similar results when Q was taken to be the simple linear function that maps from the range of QOL to $[0, 1]$ (Table 2) except that the estimates were more variable, and the coverage probabilities were less stable (92.6% – 97.4%) compared to those observed in Table 1, where the two subgroups had opposite QOL trends. This might be due to the fact that linear function Q is more sensitive to extreme QOL scores compared to the inverse logit function.

Table 1: Simulation results for two subgroups with QOL in different directions using the inverse logit Q (true $\mu = 1.853$). We compare two estimators: the proposed mean QAL estimator ($\hat{\mu}$) and the estimator assuming the true QOL trajectories are known ($\hat{\mu}_{ZT}^{(o)}$). Absolute bias (Bias), the standard error of estimates in Monte Carlo samples multiplied by $\sqrt{n}$ (MCSE), Monte Carlo mean of the variance estimates multiplied by $\sqrt{n}$ (SES), and coverage probability of the 95% confidence interval (95% CP) are reported.

Censoring	Sample size	Estimator	Bias	MCSE	ESE	95% CP
20%	200	$\hat{\mu}$	0.002	1.187	1.157	93.5%
		$\hat{\mu}_{ZT}^{(o)}$	0.002	1.132	1.157	94.5%
	400	$\hat{\mu}$	0.002	1.166	1.160	94.5%
		$\hat{\mu}_{ZT}^{(o)}$	0.002	1.160	1.159	95.2%
	800	$\hat{\mu}$	<0.001	1.143	1.162	94.6%
		$\hat{\mu}_{ZT}^{(o)}$	<0.001	1.165	1.161	94.7%
40%	200	$\hat{\mu}$	0.001	1.290	1.361	96.1%
		$\hat{\mu}_{ZT}^{(o)}$	<0.001	1.235	1.257	95.5%
	400	$\hat{\mu}$	0.002	1.264	1.358	96.0%
		$\hat{\mu}_{ZT}^{(o)}$	0.003	1.258	1.257	95.5%
	800	$\hat{\mu}$	<0.001	1.250	1.362	96.5%
		$\hat{\mu}_{ZT}^{(o)}$	<0.001	1.214	1.261	96.0%

Table 2: Simulation results for two subgroups with QOL in different directions using the linear Q (true $\mu = 2.482$). We compare two estimators: the proposed mean QAL estimator ($\hat{\mu}$) and the estimator assuming the true QOL trajectories are known ($\hat{\mu}^{(o)}_{ZT}$). Absolute bias (Bias), the standard error of estimates in Monte Carlo samples multiplied by $\sqrt{n}$ (MCSE), Monte Carlo mean of the variance estimates multiplied by $\sqrt{n}$ (SES), and coverage probability of the 95% confidence interval (95% CP) are reported.

Censoring	Sample size	Estimator	Bias	MCSE	ESE	95% CP
20%	200	$\hat{\mu}$	0.005	1.901	1.847	93.4%
		$\hat{\mu}^{(o)}_{ZT}$	0.004	1.889	1.768	92.0%
	400	$\hat{\mu}$	<0.001	1.810	1.847	95.1%
		$\hat{\mu}^{(o)}_{ZT}$	<0.001	1.791	1.772	94.7%
	800	$\hat{\mu}$	0.002	1.785	1.848	95.5%
		$\hat{\mu}^{(o)}_{ZT}$	0.002	1.756	1.773	95.3%
40%	200	$\hat{\mu}$	0.006	2.146	2.286	95.0%
		$\hat{\mu}^{(o)}_{ZT}$	0.003	2.073	1.949	92.6%
	400	$\hat{\mu}$	0.001	1.991	2.280	96.8%
		$\hat{\mu}^{(o)}_{ZT}$	<0.001	1.985	1.994	94.8%
	800	$\hat{\mu}$	0.001	2.003	2.280	97.4%
		$\hat{\mu}^{(o)}_{ZT}$	0.002	1.974	1.996	95.9%

The simulation results indicated that the finite sample performances of the proposed estimator were reasonably equivalent to those of the optimal IPW estimator, which assumed the true QOL scores to be known, despite estimating the QOL process in the former. When two subgroups had increasing QOL in the same direction with different slopes, results were similar to those presented in Table 1 and Table 2, and therefore deferred to the Appendix, Section D.

5 Application to Hepatitis C Study

5.1 Hepatitis C Study

The Study of Viral Resistance to Antiviral Therapy of Chronic Hepatitis C (Virahep-C) was designed to evaluate and compare the effectiveness of peginterferon and ribavirin combination therapy for hepatitis C between African Americans (AA) and Caucasian Americans (CA) (Conjeevaram et al., 2006). All patients were treated for 24 weeks with peginterferon and ribavirin and evaluated for a response at the end of week 24. We focused on 187 (74 AA; 113 CA) patients whose viral load at the end of initial treatment was non-detectable (<600 IU/mL), and had at least 2 measurements of symptom available after initial treatment. These responding patients received peginterferon combination therapy for an additional 24 weeks (treatment period) and were

followed up for 48 more weeks (follow-up period).

Clinically, two events are of main interest: "break-through" is defined as an increase in viral level to above or equal 600 IU/mL during the 24 week treatment following the response, and "recurrence" is defined as the re-appearance of viral level above the detectable unit at any time after the response. Although the risk of hepatitis C virus (HCV) recurrence is the major concern, poor quality of life of HCV patients has been reported in many HCV studies (Chong et al., 2003) which might lead to stopping treatment and eventually increase the risk of HCV recurrence (Bonkovsky et al., 2007). Therefore, it is of clinical interest to consider time to breakthrough and time to recurrence adjusting for some quality of life measure.

One quality of life function measured in Virahep-C study was symptom assessment, measured via a VAS (Bonkovsky et al., 2007). Symptom assessment questionnaires were given to each patient at every four weeks during the treatment period, and also at weeks 28, 36, 48 in the follow-up phase after the treatment. On a continuous line starting at 0 and ending at 10, patients marked their overall well-being, with 0 indicating "very good" and 10 indicating "awful" status. We used this symptom assessment scores to calculate mean symptom-adjusted time to break-through as well as mean symptom-adjusted time to recurrence.

The survival time to HCV recurrence was restricted to 50 weeks so that there were still enough patients at risk in the study. Restricted survival time to recurrence in Caucasian Americans was significantly shorter compared to African Americans (Figure 2a, p-value of the log-rank test is 0.01). By week 50, 70.3% of African Americans and 54.7% of Caucasian Americans had experienced a recurrence of the disease. In total, 45 patients were censored, of whom 16 were African Americans. The mean restricted time to recurrence for Caucasian Americans (mean of 46.1 weeks with 95% CI $[44.4, 47.9]$) was significantly longer than that for African Americans (mean of 40.3 weeks with 95% CI $[37.1, 43.6]$).

During the treatment period, Caucasian Americans had relatively higher symptom scores (worse well-being) at all but the first time point compared to African Americans. While the median scores of Caucasian Americans seemed to remain at a high level steadily during the treatment period, those of African Americans seemed to have a decreasing trend, suggesting an improvement of overall symptoms. However, Caucasian Americans had an apparent improvement right after the treatment had ended, and the median scores of Caucasian Americans remained to be lower in all the follow-up visits (weeks 28, 36 and 48). Interestingly, Caucasian Americans appeared to experience more symptoms during the treatment period than African Americans, but became similar or even experienced less symptom than African Americans after the treatment was completed. This observation suggests that the racial difference in time to breakthrough might differ from the racial difference in time to recurrence when adjusted for the symptom score.

Therefore, we applied our method to estimate the symptom adjusted mean time to breakthrough and time to recurrence by restricting the study time to the end of treatment (0 - 24 week) and the end of follow up (0 - 50 week), respectively. For symptom-adjusted time to recurrence, we fitted a linear mixed model as the longitudinal component with B-spline with boundary knots at 0 and 50 and interior knot at 24 as fixed effects, along with race and interaction between race and spline terms. A random intercept was also included to account for the within subject correlation. For symptom-adjusted-time to breakthrough, a similar random intercept model was used replacing the spline terms by B-spline with boundary knots at 0 and 24. The AFT model for the survival outcome included log-transformed baseline viral load as an additional time-independent covariate. The formulas of marginal models are listed below. Note that for breakthrough and recurrence, splines

terms $b_k(t)$ are different.

$$W_{ij} = \sum_{k=1}^{K_b} \beta_k b_k(t_{ij}) + \phi \mathrm{I}(\mathrm{race} = \mathrm{AA}) + \sum_{k=1}^{K_b} \psi_k b_k(t_{ij}) \mathrm{I}(\mathrm{race} = \mathrm{AA}) + b_i + e_{ij} \,,$$

$$\lambda_i(t|\mathrm{Data}) = \lambda_o(t) \exp\Big[\alpha\big\{\sum_{k=1}^{K_b} \beta_k b_k(t) + \phi \mathrm{I}(\mathrm{race} = \mathrm{AA}) + \sum_{k=1}^{K_b} \psi_k b_k(t) \mathrm{I}(\mathrm{race} = \mathrm{AA}) + b_i\big\} + \eta \log(\mathrm{vload}_i)\Big],$$

where b_i is the normal random effect and e_{ij} are normal errors. The utility function Q was chosen to be inverse logit and applied to centered symptom scores. We provide the analysis results with $q = 1$ and sensitivity analysis results with different q in Section E in the Appendix.

5.2 Analysis Results

Predicted symptom scores for both African American and Caucasian patients from the joint modeling are presented in Figure 2c. Note that the same separate mixed model that ignored survival outcomes slightly underestimated symptoms scores trajectories, especially after the treatment period when recurrent hepatitis became more frequent.

Figure 3 summarizes the analysis results for the mean time to breakthrough and the mean time to recurrence. While Caucasian Americans were observed to have an estimated one week higher mean time to breakthrough compared to African Americans [CA: 23.5 weeks (95 % CI : 22.9, 24.0), AA: 22.5 weeks (95 % CI : 21.4, 23.6)], the mean time to breakthrough adjusting for the symptom was almost identical between two races [CA: 19.5 (95 % CI : 19.0, 20.0), AA: 19.5 (95 % CI : 18.5, 20.5)]. This result was in contrast with the racial difference in the mean time to recurrence, where Caucasian Americans had a higher mean time to recurrence than African Americans regardless of adjusting for symptoms. With or without adjustment for the symptom scores, the mean time to recurrence was approximately 5 weeks higher for Caucasian Americans compared to African Americans. The results are likely due to the fact that Caucasian Americans suffered more from the treatment-related symptoms compared to African Americans but when treatments were withdrawn, these symptoms were less frequent in Caucasian Americans compared with African Americans (Figure 2b). This analysis result suggests that the mean time to breakthrough during the treatment after initial response between the two groups of patients should be compared by adjusting for symptoms.

6 Discussion

In this paper, we propose an estimator of the mean quality-adjusted lifetime for a continuous quality of life process using joint modeling of survival and longitudinal data. With an appropriately chosen utility function, we can model individuals' continuously changing health status by the trajectory of quality of life score. We present a proof of consistency and asymptotic normality of the proposed estimator under mild assumptions. We propose a variance estimator, and finite sample performance of this variance estimator is verified through simulations. The proposed method produces consistent estimates of the mean QAL, despite the need for estimating the underlying QOL process observed with errors, and matches the performance of an optimal estimator that uses

known QOL trajectories. Finally, a sensitivity analysis is developed to accompany the method to draw robust conclusions.

The consistency of our estimator depends on two assumptions, which might motivate future research. First, the censoring time is assumed to be independent of the survival time and the health experience. In situations where the censoring time depends on the survival time, one might consider sensitivity analysis to mimic the best and worst case scenarios and use the drop-out event as a study end-point. If the censoring depends on the health experience, theoretically we can model the actual distribution of the censoring time and substitute the Kaplan-Meier estimator for the censoring time in the denominator of our proposed estimator (Shih, 2002). Second, we adopt a parametric specification of the model. A variety of nonparametric and semiparametric estimation methods of QOL functions under censoring can be considered to allow more flexibility.

One limitation about the methods proposed in this manuscript is that the marginal survival model used is completely parametric with baseline hazard following a parametric form and with a parametric linear predictor part. Correct specification of these two components are necessary for any inferences to be valid.

Acknowledgments

The authors thank the anonymous referee for the valuable comments. Kang was supported by the University of Pittsburgh Central Research Development Fund (CRDF).

Appendix

A Proof of Theorem 1

When $\Delta_i = 1$, we observe $(T_i, W_{i\cdot}, X_i, V_i)$. On the event $E_n = \{\|\hat{\beta}_n - \beta_o\| < C,\ \sup_i \|X_i(t)\| < C\}$, the following inequalities hold by A3. By A1 and A2, $pr(E_n) = pr(\|\hat{\beta}_n - \beta_o\| < C) \to 1$.

$$
\begin{aligned}
|\mathrm{U}_i(\hat{\beta}_n) - \mathrm{U}_i(\beta_o)| &= \left| \int_0^{T_i} f(X_i(t), \hat{\beta}_n)\, dt - \int_0^{T_i} f(X_i(t), \beta_o)\, dt \right| \\
&\leqslant \int_0^{T_i} \|\dot{f}(X_i(t), \beta_o)\| \cdot \|\hat{\beta}_n - \beta_o\|\, dt + \frac{1}{2} \int_0^{T_i} \|\hat{\beta} - \beta_o\| \cdot \|\ddot{f}(X_i(t), \beta^*)(\hat{\beta}_n - \beta_o)\|\, dt\ .
\end{aligned}
$$

By A2, $\dot{f}(X_i(t), \beta_o)$ is continuous on $[0, C] \times \beta_0$, and thus bounded by some constant N. Also, by A3, $\|\ddot{f}_{X_i}(t, \beta^*)(\hat{\beta}_n - \beta_o)\| \leqslant M\|\hat{\beta}_n - \beta_o\|$, thus in probability,

$$
\sup_i |\mathrm{U}_i(\hat{\beta}_n) - \mathrm{U}_i(\beta_o)| \leqslant NL\|\hat{\beta}_n - \beta_o\| + \frac{1}{2} ML \|\hat{\beta}_n - \beta_o\|^2 =: W_n. \tag{A1}
$$

By the definition of our estimator $\hat{\mu}_n$, we have

$$
\hat{\mu}_n - \mu = \frac{1}{n} \sum_{i=1}^{n} \frac{\Delta_i U_i(\hat{\beta}_n)}{\hat{K}(T_i)} - \mu. \tag{A2}
$$

Plugging (A1), and

$$
\frac{1}{\hat{K}(T_i)} = \left(\frac{1}{K(T_i)} - \frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i) K(T_i)} \right)
$$

into (A2) we have

$$\hat{\mu}_n - \mu = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{K(T_i)} - \frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i)K(T_i)}\right)\Delta_i U_i(\hat{\beta}_n) - \mu. \tag{A3}$$

Therefore,

$$\begin{aligned}|\hat{\mu}_n - \mu| \leqslant & \left|\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i U_i(\beta_o)}{K(T_i)} - \mu\right| + \left|\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i U_i(\beta_o)(\hat{K}(T_i) - K(T_i))}{\hat{K}(T_i)K(T_i)}\right| \\ & + \left|\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i}{K(T_i)}\right)W_n\right| + \left|\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i(\hat{K}(T_i) - K(T_i))}{\hat{K}(T_i)K(T_i)}\right)W_n\right|\end{aligned} \tag{A4}$$

in probability.

By the Weak Law of Large Numbers, the first term on the right hand side (RHS) of (A4) converges to 0 in probability. The second term is bounded above by

$$\left(\frac{1}{n}\sum_{i=1}^{n}\Delta_i U_i(\beta_o)\right)\frac{\sup_{0\leqslant u\leqslant L}|\hat{K}(u) - K(u)|}{\hat{K}(L)K(L)}\ .$$

Since $\Delta_i = 1$ or 0 and $U_i(\beta_o)$ is bounded by L, $\{\sum_{i=1}^{n}\Delta_i U_i(\beta_o)\}/n \in [0, L]$. Also, by the uniform convergence of the KM estimator $\sup_{0\leqslant u\leqslant L}|\hat{K}(u) - K(u)|\to 0$ and $\hat{K}(L)K(L)\to K^2(L)$, both in probability. Therefore, the second term on the RHS converges in probability to 0. Similarly, the fourth term is $o_p(1)$. The third term is naturally bounded above by $W_n/K(L)$. Since W_n is a function of $\|\hat{\beta}_n - \beta_o\|$ and uniformly bounded, it is $o_p(1)$. Putting together, under assumptions A1-A3, $\hat{\mu}_n$ is a consistent estimator for μ.

B Proof of Theorem 2

We write the joint log-likelihood function $\mathrm{L}(\theta) = \frac{1}{n}\sum_{i=1}^{n} l_i(\theta)$, where $\mathrm{L}(\cdot)$ is the joint log-likelihood function and $l_i(\cdot)$ is the individual log-likelihood function for the ith subject in the sample. Under assumptions A1-A3, the first and second derivatives of individual likelihoods exist. Let us denote them by $\dot{l}_i$ and $\ddot{l}_i$, respectively. By manipulating the Taylor approximation of the score function around θ_o, we obtain

$$\hat{\beta}_n - \beta_o = \frac{1}{n}\sum_{i=1}^{n}\psi_i(\theta_o|X_i, Z_i, b_i, V_i) + o_p(n^{-\frac{1}{2}}), \tag{B1}$$

where $\psi_i(\theta_o|X_i, Z_i, b_i, V_i)$ is the β components of $E[\ddot{l}_i(\theta_o|X_i, Z_i, b_i, V_i)]^{-1}\{\dot{l}_i(\theta_o|X_i, Z_i, b_i, V_i)\}^T$, and thus has zero expectation. For some β^* between $\hat{\beta}_n$ and β_o,

$$\begin{aligned}U_i(\hat{\beta}_n) - U_i(\beta_o) &= \int_0^{T_i}\{f(X_i(t), \hat{\beta}_n) - f(X_i(t), \beta_o)\}\,dt \\ &= \int_0^{T_i}\{\dot{f}(X_i(t), \beta_o)(\hat{\beta}_n - \beta_o)\}\,dt + \frac{1}{2}\int_0^{T_i}(\hat{\beta}_n - \beta_o)^T\ddot{f}(X_i(t), \beta^*)(\hat{\beta}_n - \beta_o)\,dt\ .\end{aligned}$$

By A2 and A3, the second term on the RHS of above is bounded in probability by $\frac{1}{2}ML\|(\hat{\beta}_n - \beta_o)\|^2$, which is $o_p(n^{-1/2})$. Therefore,

$$U_i(\hat{\beta}_n) - U_i(\beta_o) = \phi_i^T(X_i(\cdot), T_i)(\hat{\beta}_n - \beta_o) + o_p(n^{-\frac{1}{2}}), \tag{B2}$$

where, for $\beta_o \in \mathbb{R}^{\mathrm{P}}$, $\phi_i(X_i(\cdot), T_i)$ is defined by

$$\phi_i(X_i(\cdot), T_i) := \left[\int_0^{T_i} \left\{ \frac{\partial f}{\partial \beta_1}\Big(X_i(t), \beta_o\Big) \right\} dt \quad \ldots \quad \int_0^{T_i} \left\{ \frac{\partial f}{\partial \beta_{\mathrm{P}}}\Big(X_i(t), \beta_o\Big) \right\} dt \right]^T .$$

From (A3), we have

$$\sqrt{n}(\hat{\mu}_n - \mu) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i U_i(\hat{\beta}_n)}{K(T_i)} - \mu \right) - \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i U_i(\hat{\beta}_n)(\hat{K}(T_i) - K(T_i))}{\hat{K}(T_i)K(T_i)} \right). \tag{B3}$$

Representing $U_i(\hat{\beta}_n)$ by $U_i(\beta_o)$ via (B2), the first term on the RHS of (B3) is written as

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o)}{K(T_i)} - \mu \right) + \left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i \phi_i(X_i(\cdot), T_i)}{K(T_i)} \right)^T \sqrt{n}(\hat{\beta}_n - \beta_o) + \left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i}{K(T_i)} \right) \cdot o_p(1).$$

Similarly the second term on the RHS of (B3) is

$$\begin{aligned} &- \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o)(\hat{K}(T_i) - K(T_i))}{\hat{K}(T_i)K(T_i)} \right) - \left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i (\hat{K}(T_i) - K(T_i))}{\hat{K}(T_i)K(T_i)} \right) o_p(1) \\ &- \left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i (\hat{K}(T_i) - K(T_i))\phi_i(X_i(\cdot), T_i)}{\hat{K}(T_i)K(T_i)} \right)^T \cdot \sqrt{n}(\hat{\beta}_n - \beta_o). \end{aligned}$$

Putting them together, we have the following result:

$$\begin{aligned} \sqrt{n}(\hat{\mu}_n - \mu) &= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o)}{K(T_i)} - \mu \right) - \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o)(\hat{K}(T_i) - K(T_i))}{\hat{K}(T_i)K(T_i)} \right) \\ &+ \left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i \phi_i(X_i(\cdot), T_i)}{K(T_i)} - \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i (\hat{K}(T_i) - K(T_i))\phi_i(X_i(\cdot), T_i)}{\hat{K}(T_i)K(T_i)} \right)^T \cdot \sqrt{n}(\hat{\beta}_n - \beta_o) \\ &+ \left(\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i}{K(T_i)} - \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i (\hat{K}(T_i) - K(T_i))}{\hat{K}(T_i)K(T_i)} \right) \cdot o_p(1). \end{aligned}$$

In Theorem 1, we have shown that the summands of the terms in the last expression are bounded, and therefore the last term converges to 0 in probability. Therefore,

$$\sqrt{n}(\hat{\mu}_n - \mu) = H_1(\beta_o) + H_2^T(\beta_o) \cdot \sqrt{n}(\hat{\beta}_n - \beta_o) + o_p(1), \tag{B4}$$

where

$$\begin{aligned} H_1(\beta_o) &= n^{-\frac{1}{2}} \left[\sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o)}{K(T_i)} - \sum_{i=1}^{n} \Delta_i U_i(\beta_o) \left\{ \frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i)K(T_i)} \right\} - n\mu \right], \text{ and} \\ H_2^T(\beta_o) &= \left[\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i \phi_i(X_i(\cdot), T_i)}{K(T_i)} - \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i \{\hat{K}(T_i) - K(T_i)\}\phi_i(X_i(\cdot), T_i)}{\hat{K}(T_i)K(T_i)} \right]^T. \end{aligned}$$

Following the arguments derived in Zhao and Tsiatis (1997) and by letting $N_i^c(t) = I(Y_i \leqslant t, \Delta_i = 0)$, $H_1(\beta_o)$ can be decomposed into two uncorrelated sums (details could be found in the next section of the Appendix):

$$H_1(\beta_o) = n^{-\frac{1}{2}} \sum_{i=1}^{n} \{U_i(\beta_o) - \mu\} - n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^L \left\{ \frac{dM_i^c(u)}{K(u)} \right\} \{U_i(\beta_o) - G(u)\} + o_p(1), \tag{B5}$$

where $M_i^c(t) = N_i^c(t) - \int_0^t \lambda^c(u) \cdot I(Y_i \geqslant u)\, du$ is the martingale under the filtration $\mathcal{F}_t = \sigma\{N_i^c(s) \text{ for } 0 \leqslant s \leqslant t\,; X_i(h), Z_i(h), V_i, T_i \text{ for } 0 \leqslant h < \infty\,; i = 1, \cdots, n\}$. As $\phi_i(X_i(\cdot), T_i)$ are also $\mathcal{F}_0$-measurable variables, following the arguments in the previous section, we have $H_2(\beta_o) \to \Phi(\beta_o)$ in probability. Therefore, from (B4), we have

$$\sqrt{n}(\hat{\mu}_n - \mu) = n^{-\frac{1}{2}} \sum_{i=1}^{n} \left[\{U_i(\beta_o) - \mu\} - \int_0^L \left\{ \frac{dM_i^c(u)}{K(u)} \right\} \{U_i(\beta_o) - G(u)\} + \Phi^T \psi_i(\theta_o | X_i, Z_i, b_i, V_i) \right] + o_p(1). \tag{B6}$$

By the Central Limit Theorem, equation (B6) implies that $\sqrt{n}(\hat{\mu}_n - \mu)$ converges in distribution to normal as long as the summands have finite second moments. As $E[U_i(\beta_o)] = \mu$, $E[dM_i^c(u)|\mathcal{F}(u-)] = 0$, and $E(\psi_i) = 0$, the limiting distribution has a mean of zero. In order to check if the second moments are finite, it suffices to show that the expectations of the square terms are finite. As $0 \leqslant |U_i(\beta_o)| \leqslant L$ implies $E\left(U_i^2(\beta_o)\right) \leqslant L^2$, $\sigma_1^2 = E\left(U_i^2(\beta_o)\right) - \mu^2$. Also, $|\frac{(U_i(\beta_o) - G(u))}{K(u)}| \leqslant \frac{L}{K(L)} + \frac{L}{S(L)K(L)}$, and therefore,

$$\sigma_2^2 = E\left[\int_0^L \left(\frac{dM_i^c(u)}{K(u)} \right) \cdot \{U_i(\beta_o) - G(u)\} \right]^2 = \int_0^L \frac{\lambda^c(u)}{K(u)} E\left[I(T_i \geqslant u) \{U_i(\beta_o) - G(u)\}^2 \right] du$$
$$\leqslant \left\{ L + \frac{L}{S(L)} \right\}^2 \cdot K(L) \int_0^L \lambda^c(u) du.$$

The second moment of the third term in (B6) is $\sigma_3^2 = \Phi^T(\beta_o) \mathrm{I}^{-1}(\theta_o) \Phi(\beta_o)$. To derive the explicit form of the asymptotic variance of $\hat{\mu}_n$, we first show the cross terms in (B6) have zero expectation. As $U_i(\beta_o)$ is $\mathcal{F}(0)$ measurable, the first two terms in (B6) are uncorrelated. For other cross terms, we define

$$\sigma_{13} = E\left[(U_i(\beta_o) - \mu)\, \Phi^T \psi_i(\theta_o | X_i, Z_i, b_i, V_i) \right], \quad \text{and}$$
$$\sigma_{23} = E\left[\Phi^T \psi_i(\theta_o | X_i, Z_i, b_i, V_i) \int_0^L \left(\frac{dM_i^c(u)}{K(u)} \right) \cdot (U_i(\beta_o) - G(u)) \right].$$

We have the following bounds:

$$-B_1 \Phi^T E\left[\psi_i(\theta_o | X_i, Z_i, b_i, V_i)\right] \leqslant \sigma_{13} \leqslant B_1 \Phi^T E\left[\psi_i(\theta_o | X_i, Z_i, b_i, V_i)\right],$$
$$-B_2 \Phi^T E\left[\psi_i(\theta_o | X_i, Z_i, b_i, V_i)\right] \leqslant \sigma_{23} \leqslant B_2 \Phi^T E\left[\psi_i(\theta_o | X_i, Z_i, b_i, V_i)\right],$$

where $B_1 = \mu + L$ and $B_2 = (\frac{L}{K(L)} + \frac{L}{S(L)K(L)})(1 + \int_0^L \lambda^c(u) du)$. As $E\left[\psi_i(\theta_o | X_i, Z_i, b_i, V_i)\right] = 0$, $\sigma_{13} = \sigma_{23} = 0$. Then $\sqrt{n}(\hat{\mu}_n - \mu) \to N(0, \sigma^2)$ in law, where

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2\,.$$

C Derivation of $H_1(\beta_o)$ of Eq. B5

Recall that

$$H_1(\beta_o) = n^{-\frac{1}{2}} \left[\sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o)}{K(T_i)} - \sum_{i=1}^{n} \Delta_i U_i(\beta_o) \left\{ \frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i) K(T_i)} \right\} - n\mu \right].$$

We define the following filtration:

$$\mathcal{F}_t = \sigma\{N_i^c(s) \text{ for } 0 \leqslant s \leqslant t \,;\, X_i(h), Z_i(h), V_i, T_i \text{ for } 0 \leqslant h < \infty \,;\, i = 1, \cdots, n\}\,,$$

where $N_i^c(t) = I(Y_i \leqslant t\,, \Delta_i = 0)$ and $Y_i = \min(C_i, T_i)$. The compensated censoring process is then expressed by

$$M_i^c(t) = N_i^c(t) - \int_0^t \lambda^c(u) \cdot I(Y_i \geqslant u)\, du.$$

By summing over all the individual, we have

$$M^c(t) = N^c(t) - \int_0^t \lambda^c(u) \cdot Y(u)\, du,$$

where $M^c(t) = \sum_i M_i^c(t)$, $N^c(t) = \sum_i N_i^c(t)$ and $Y(t) = \sum_i I(Y_i \geqslant t)$. Note that $M^c(t)$ is $\mathcal{F}_t$ martingale. Let $S(\cdot)$ and $K(\cdot)$ denote the marginal survival function of T_i and C_i, respectively. The following equality holds from Robins and Rotnitzky (1992):

$$\frac{\Delta_i}{K(T_i)} = 1 - \int_0^\infty \frac{dM_i^c(u)}{K(u)} = 1 - \int_0^L \frac{dM_i^c(u)}{K(u)}\,.$$

Then the first and third terms in $H_1(\beta_0)$ can be written as

$$n^{-\frac{1}{2}} \sum_{i=1}^n (U_i(\beta_o) - \mu) - n^{-\frac{1}{2}} \sum_{i=1}^n \int_0^L \left(\frac{dM_i^c(u)}{K(u)}\right) U_i(\beta_o)\,. \tag{C1}$$

Gill (1980) showed that

$$\frac{\hat{K}(T_i) - K(T_i)}{K(T_i)} = -\int_0^{T_i} \frac{\hat{K}(u^-)}{K(u)} \frac{dM^c(u)}{Y(u)}\,,$$
$$n^{-1} Y(u) = \hat{K}(u^-)\hat{S}(u^-)\,,$$

which implies

$$\frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i)K(T_i)} = -\int_0^{T_i} \left(\frac{dM^c(u)}{K(u)}\right) \left(\frac{1}{n\hat{S}(u^-)\hat{K}(T_i)}\right)\,.$$

Note that

$$\begin{aligned}
\sum_{i=1}^n \Delta_i U_i(\beta_o) \left\{\frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i)K(T_i)}\right\} &= -\sum_{i=1}^n \int_0^{T_i} \left(\frac{dM^c(u)}{K(u)}\right) \left(\frac{\Delta_i U_i(\beta_o)}{n\hat{S}(u^-)\hat{K}(T_i)}\right) \\
&= -\sum_{j=1}^n \sum_{i=1}^n \int_0^{T_i} \left(\frac{dM_j^c(u)}{K(u)}\right) \left(\frac{\Delta_i U_i(\beta_o)}{n\hat{S}(u^-)\hat{K}(T_i)}\right) \\
&= -\sum_{j=1}^n \sum_{i=1}^n \int_0^{L} \left(\frac{dM_j^c(u)}{K(u)}\right) \left(\frac{\Delta_i U_i(\beta_o) I(T_i \geqslant u)}{n\hat{S}(u^-)\hat{K}(T_i)}\right) \\
&= -\sum_{j=1}^n \int_0^{L} \left(\frac{dM_j^c(u)}{K(u)}\right) \left(\sum_{i=1}^n \frac{\Delta_i U_i(\beta_o) I(T_i \geqslant u)}{n\hat{S}(u^-)\hat{K}(T_i)}\right)\,.
\end{aligned}$$

Therefore we have

$$\sum_{i=1}^{n} \Delta_i U_i(\beta_o) \left\{ \frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i)K(T_i)} \right\} = -\sum_{j=1}^{n} \int_0^L \left(\frac{dM_j^c(u)}{K(u)} \right) \left(\sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o) I(T_i \geqslant u)}{n\hat{S}(u^-)\hat{K}(T_i)} \right). \tag{C2}$$

From (C1) and (C2), we have

$$H_1(\beta_o) = n^{-\frac{1}{2}} \sum_{i=1}^{n} (U_i(\beta_o) - \mu) - n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^L \left(\frac{dM_i^c(u)}{K(u)} \right) \cdot \left(U_i(\beta_o) - \hat{G}(u) \right), \tag{C3}$$

where

$$\hat{G}(u) = \frac{1}{n\hat{S}(u^-)} \sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o) I(T_i \geqslant u)}{\hat{K}(T_i)} \tag{C4a}$$

$$= \frac{\hat{\mu}}{\hat{S}(u^-)} - \frac{1}{n\hat{S}(u^-)} \sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o) I(T_i < u)}{\hat{K}(T_i)}. \tag{C4b}$$

Note that for all $u \in [0, L]$, in probability $\hat{G}(u) \to G(u) := \{1/S(u^-)\} E\left(U_i(\beta_o) I(T_i \geqslant u)\right)$ by uniform convergence of the KM estimator and weak law of large numbers. By replacing $\hat{\mu}_n$ with μ in (C4b), similar to Zhao and Tsiatis (1997), we define the $\mathcal{F}_t$ predictable process as follows:

$$\begin{aligned} G^*(u) &= \frac{\mu}{\hat{S}(u^-)} - \frac{1}{n\hat{S}(u^-)} \sum_{i=1}^{n} \frac{\Delta_i U_i(\beta_o) I(T_i < u)}{\hat{K}(T_i)} \\ &= \hat{G}(u) + \frac{1}{\hat{S}(u^-)} (\mu - \hat{\mu}_n). \end{aligned} \tag{C5}$$

Now we plug $G(u)$ and (C5) into (C3) and we have

$$\begin{aligned} H_1(\beta_o) &= n^{-\frac{1}{2}} \sum_{i=1}^{n} (U_i(\beta_o) - \mu) - n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^L \left(\frac{dM_i^c(u)}{K(u)} \right) \cdot \left(U_i(\beta_o) - \hat{G}(u) \right) \\ &= n^{-\frac{1}{2}} \sum_{i=1}^{n} (U_i(\beta_o) - \mu) - n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^L \left(\frac{dM_i^c(u)}{K(u)} \right) \cdot (U_i(\beta_o) - G(u)) - \\ n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^L \left(\frac{dM_i^c(u)}{K(u)} \right) \cdot (G(u) - G^*(u)) &+ n^{-\frac{1}{2}} (\mu - \hat{\mu}_n) \sum_{i=1}^{n} \int_0^L \left(\frac{dM_i^c(u)}{K(u)\hat{S}(u^-)} \right). \end{aligned}$$

As $U_i(\beta_o)$ is $\mathcal{F}_0$-measurable, all the stochastic integrands above are $\mathcal{F}_t$ predictable processes. Moreover, the first and second terms are uncorrelated. Note that $u \in [0, L]$, $G^*(u) \to G(u)$ in probability, thus the integrand of the third term above converges to 0 in probability. Clearly, by Slusky's theorem and the martingale central limit theorem, the last term also converges to 0 in probability. Then $H_1(\beta_o)$ could be decomposed into the two uncorrelated sums as follows:

$$H_1(\beta_o) = n^{-\frac{1}{2}} \sum_{i=1}^{n} (U_i(\beta_o) - \mu) - n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^L \left(\frac{dM_i^c(u)}{K(u)} \right) \cdot (U_i(\beta_o) - G(u)) + o_p(1).$$

D More Simulation Results

In this section, we provide the results of simulation for the scenario where two subgroups have the different strength but the same direction in QOL trend. In summary, the finite sample performance of the proposed method was similar to the optimal IPW estimator. Table D1 shows the results of simulations for two subgroups with QOL trends in the same direction with function Q chosen to be (2.2). Bias, Monte Carlo standard error (MCSE), Monte Carlo mean of standard error estimate (ESE) and coverage probability of the 95% confidence interval were summarized over 1,000 Monte Carlo samples. Both the optimal and the proposed estimators had negligible absolute biases ($\leq$ 0.005) and comparable variabilities and empirical coverage probabilities. The coverage probabilities of the 95% confidence intervals for the proposed estimator were between 93.2% and 97.2%, compared to 93.5% to 96.2% of those from the optimal estimator. The standard error estimates for the proposed estimator closely matched the MC standard errors. The accuracy and the precision of the mean estimates improved with larger sample sizes and smaller censoring proportions, as evidenced by the decrease bias and difference between MCSE and ESE. We observed similar results when Q was taken to be the simple linear function that maps from the range of QOL to [0, 1] (Table D2).

Table D1: Simulation results for subgroups with QOL in the same direction using the inverse logit Q ($\mu = 3.461$). We compare two estimators: The proposed mean QAL estimator ($\hat{\mu}$) and the estimator assuming the true QOL trajectories are known ($\hat{\mu}_{ZT}^{(o)}$). Absolute bias (Bias), the standard error of estimates in Monte Carlo samples multiplied by $\sqrt{n}$ (MCSE), Monte Carlo mean of the variance estimates multiplied by $\sqrt{n}$ (SES), and coverage probability of the 95% confidence interval (95% CP) are reported.

Censoring	Sample size	Estimator	Bias	MCSE	ESE	95% CP
20%	200	$\hat{\mu}$	0.003	1.265	1.209	94.1%
		$\hat{\mu}_{ZT}^{(o)}$	0.004	1.244	1.195	94.1%
	400	$\hat{\mu}$	0.003	1.282	1.212	93.2%
		$\hat{\mu}_{ZT}^{(o)}$	<0.001	1.258	1.197	93.5%
	800	$\hat{\mu}$	0.003	1.181	1.212	95.7%
		$\hat{\mu}_{ZT}^{(o)}$	<0.001	1.148	1.197	95.4%
40%	200	$\hat{\mu}$	0.003	1.385	1.408	94.8%
		$\hat{\mu}_{ZT}^{(o)}$	0.004	1.364	1.319	93.8%
	400	$\hat{\mu}$	0.005	1.411	1.408	94.4%
		$\hat{\mu}_{ZT}^{(o)}$	0.002	1.388	1.321	93.5%
	800	$\hat{\mu}$	0.003	1.302	1.405	97.2%
		$\hat{\mu}_{ZT}^{(o)}$	<0.001	1.272	1.320	96.2%

Table D2: Simulation results for subgroups with QOL in the same direction using the linear Q ($\mu = 3.283$). We compare two estimators: The proposed mean QAL estimator ($\hat{\mu}$) and the estimator assuming the true QOL trajectories are known ($\hat{\mu}^{(o)}_{ZT}$). Absolute bias (Bias), the standard error of estimates in Monte Carlo samples multiplied by $\sqrt{n}$ (MCSE), Monte Carlo mean of the variance estimates multiplied by $\sqrt{n}$ (SES), and coverage probability of the 95% confidence interval (95% CP) are reported.

Censoring	Sample size	Estimator	Bias	MCSE	ESE	95% CP
20%	200	$\hat{\mu}$	<0.001	1.915	1.803	92.7%
		$\hat{\mu}^{(o)}_{ZT}$	0.004	1.729	1.792	95.6%
	400	$\hat{\mu}$	<0.001	1.891	1.803	93.6%
		$\hat{\mu}^{(o)}_{ZT}$	<0.001	1.748	1.794	95.8%
	800	$\hat{\mu}$	<0.001	1.757	1.806	95.0%
		$\hat{\mu}^{(o)}_{ZT}$	<0.001	1.758	1.795	95.4%
40%	200	$\hat{\mu}$	0.002	2.020	1.938	93.3%
		$\hat{\mu}^{(o)}_{ZT}$	0.004	1.830	1.894	95.3%
	400	$\hat{\mu}$	<0.001	1.977	1.937	93.7%
		$\hat{\mu}^{(o)}_{ZT}$	0.001	1.856	1.895	95.3%
	800	$\hat{\mu}$	<0.001	1.870	1.940	94.6%
		$\hat{\mu}^{(o)}_{ZT}$	<0.001	1.851	1.896	95.8%

E Sensitivity Analysis for Virahep-C

Following section on sensitivity analysis in the main text, mean symptom adjusted survival times were estimated for a range of q. More specifically, we considered q values from 0.6 to 0.99 with an increment of 0.01, and from 1 to 5 with an increment of 0.1. Mean adjusted time for African Americans and Caucasian Americans were plotted against $1/q$ in FigureE1 (left panel) with the corresponding threshold utility functions (right panel). For all utilities considered, Caucasian Americans had higher symptom-adjusted mean recurrence compared with African Americans, with difference in mean symptom-adjusted-time to recurrence ranging from 0.95 to 7.01 in weeks. The 95% CIs of two groups were mostly non-overlapping except for $1.1 \leqslant q \leqslant 2.2$. The utility functions corresponding to both threshold qs are given in Figure E1b.

For mean symptom-adjusted-time to recurrence, difference between Caucasian and African Americans includes 0, ranging from -4.45 to 1.75 in weeks. The 95% CIs of the two groups were overlapping for $0.76 < q < 1.3$. The utility functions corresponding to both threshold qs are given in Figure E1d.

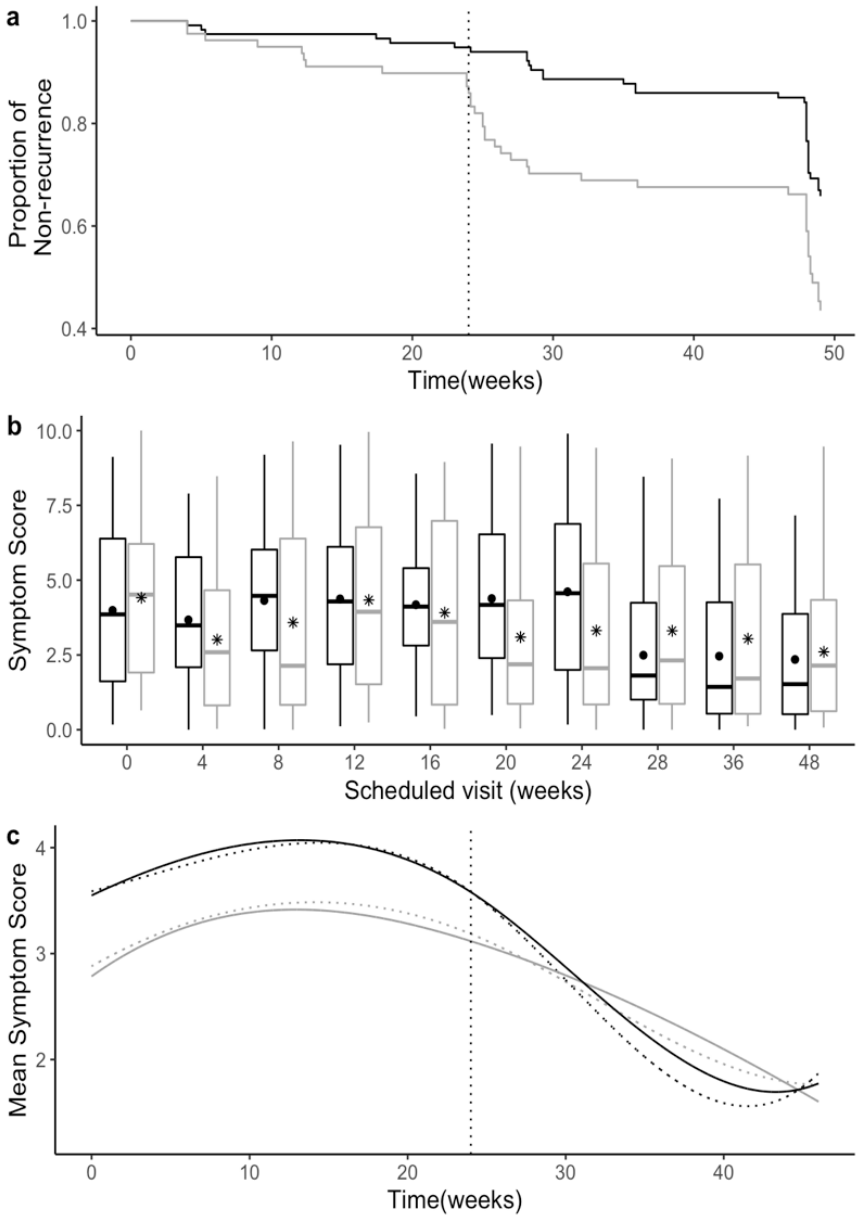

Figure 2: Time to breakthrough/recurrence and symptom scores in the Virahep-C study by race: African Americans (grey) and Caucasian Americans (black). a) Kaplan-Meier survival curves for the time to hepatitis recurrence restricted to 50 weeks; b) Box-plot of symptom scores at each visit (CA means by dots, AA means by stars); c) Mean symptom score curves estimated by using linear mixed model (dotted lines) and the proposed joint modeling (solid lines). The dotted vertical line at week 24 indicates the time when the treatment ended.

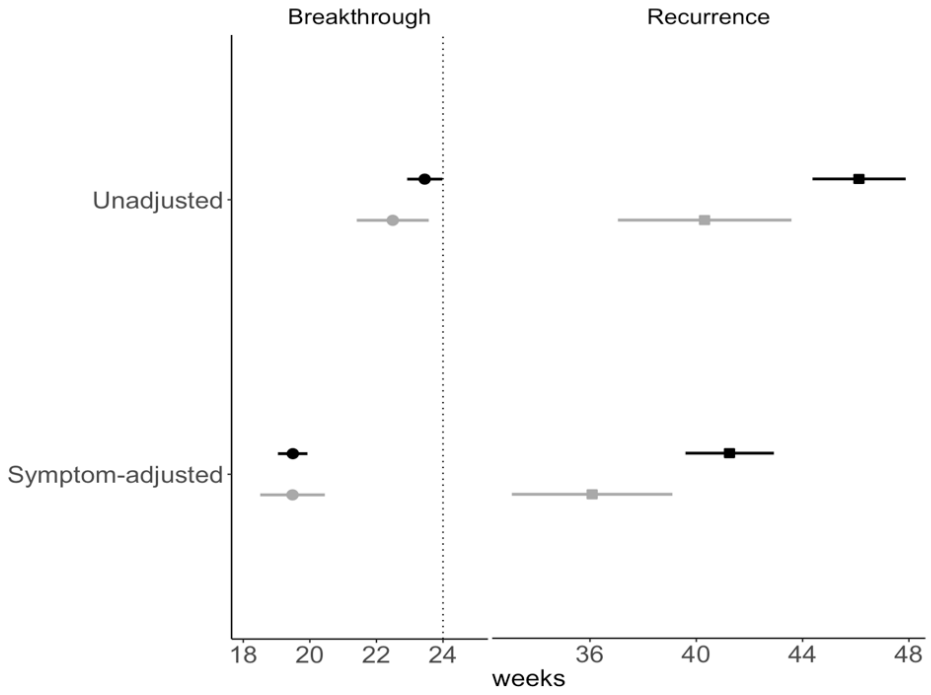

Figure 3: Mean time to recurrence/breakthrough and their 95% confidence intervals in the Virahep-C study by race: African Americans (grey) and Caucasian Americans (black). The dotted vertical line at week 24 indicates the time when the treatment ended. Unadjusted/symptom-adjusted time to breakthrough (means denoted by dots, and 95% CIs denoted by horizontal lines) lies to the left of the dashed line; time to recurrence (means denoted by squares, and 95% CIs denoted by horizontal lines) lies to the right of the dashed line.

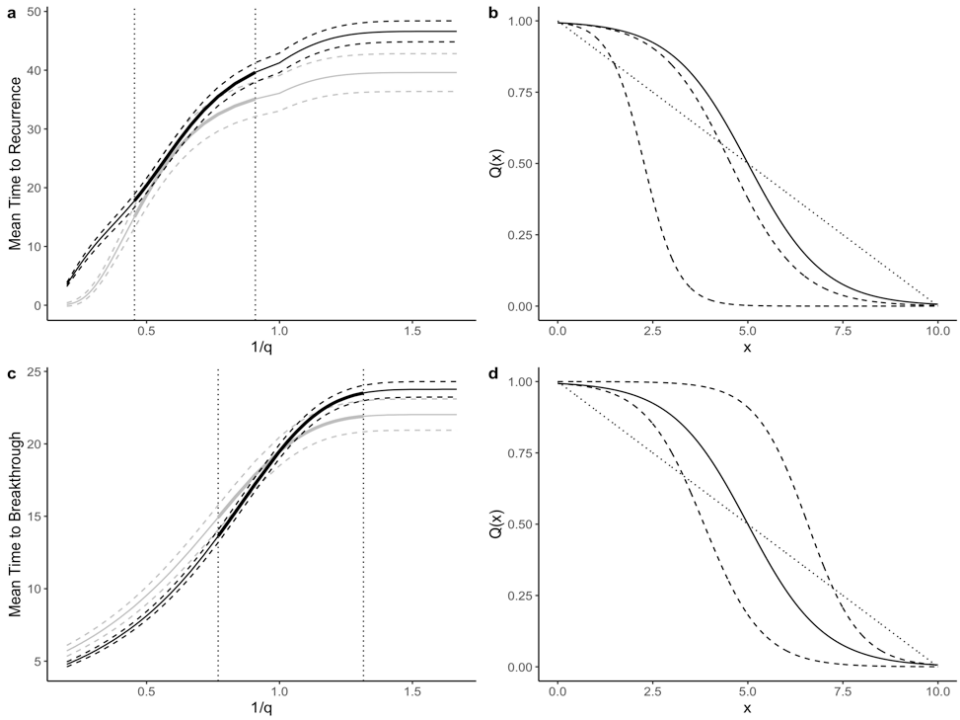

Figure E1: Sensitivity analysis for the choice of utility function. **a**) Mean time to recurrence in weeks (solid lines) along with the 95% CI (dashed lines) for different choice of utility functions by race: African Americans (grey) and Caucasian Americans (black). The bold lines between two dotted lines indicates the utility functions with overlapping CIs for African Americans and Caucasian Americans. **b**) Utility functions. Dashed curves have the threshold utility parameters, which is indicated by vertical dotted lines to the left. The black curve corresponds to the basic choice (q=1). **c**) Mean time to breakthrough in weeks (solid lines) along with the 95% CI (dashed lines) for different choice of utility functions by race: African Americans (grey) and Caucasian Americans (black). The bold lines between two dotted lines indicates the utility functions with overlapping CIs for African Americans and Caucasian Americans.**d**) Utility functions. Dashed curves have the threshold utility parameters, which is indicated by vertical dotted lines to the left. The black curve corresponds to the basic choice (q=1).

References

Andrei, A.-C. and Murray, S. (2007), "Regression Models for the Mean of the Quality-of-Life-Adjusted Restricted Survival Time Using Pseudo-Observations," *Biometrics*, 63, 398–404.

Bonkovsky, H. L., Snow, K. K., Malet, P. F., Back-Madruga, C., Fontana, R. J., Sterling, R. K., Kulig, C. C., Di Bisceglie, A. M., Morgan, T. R., Dienstag, J. L., et al. (2007), "Health-related quality of life in patients with chronic hepatitis C and advanced fibrosis," *Journal of Hepatology*, 46, 420–431.

Chong, C. A., Gulamhussein, A., Heathcote, E. J., Lilly, L., Sherman, M., Naglie, G., and Krahn, M. (2003), "Health-state utilities and quality of life in hepatitis C patients," *The American Journal of Gastroenterology*, 98, 630.

Cole, B. F., Gelber, R. D., and Goldhirsch, A. (1993), "Cox regression models for quality adjusted survival analysis," *Statistics in Medicine*, 12, 975–987.

Conjeevaram, H. S., Fried, M. W., Jeffers, L. J., Terrault, N. A., Wiley-Lucas, T. E., Afdhal, N., Brown, R. S., Belle, S. H., Hoofnagle, J. H., Kleiner, D. E., et al. (2006), "Peginterferon and ribavirin treatment in African American and Caucasian American patients with hepatitis C genotype 1," *Gastroenterology*, 131, 470–477.

Gelber, R. D., Cole, B. F., Gelber, S., and Goldhirsch, A. (1995), "Comparing treatments using quality-adjusted survival: the Q-TWiST method," *The American Statistician*, 49, 161–169.

Gelber, R. D., Goldhirsch, A., and Cavalli, F. (1991), "Quality-of-life-adjusted evaluation of adjuvant therapies for operable breast cancer," *Annals of Internal Medicine*, 114, 621–628.

Gill, R. D. (1980), "Censoring and stochastic integrals," *Statistica Neerlandica*, 34, 124–124.

Goldhirsch, A., Gelber, R. D., Simes, R. J., Glasziou, P., and Coates, A. S. (1989), "Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis." *Journal of Clinical Oncology*, 7, 36–44.

Huang, Y. and Louis, T. A. (1999), "Expressing estimators of expected quality adjusted survival as functions of Nelson-Aalen estimators," *Lifetime Data Analysis*, 5, 199–212.

Hwang, J.-S., Tsauo, J., and Wang, J. (1996), "Estimation of expected quality adjusted survival by cross-sectional survey," *Statistics in Medicine*, 15, 93–102.

Korn, E. L. (1993), "On estimating the distribution function for quality of life in cancer clinical trials," *Biometrika*, 80, 535–542.

Laan, M. J. and Hubbard, A. (1999), "Locally Efficient Estimation of the Quality-Adjusted Lifetime Distribution with Right-Censored Data and Covariates," *Biometrics*, 55, 530–536.

Pradhan, B. and Dewanji, A. (2009), "Parametric estimation of quality adjusted lifetime (QAL) distribution in progressive illness–death model," *Statistics in Medicine*, 28, 2012–2027.

Rizopoulos, D. (2010), "JM: An R package for the joint modelling of longitudinal and time-to-event data," *Journal of Statistical Software (Online)*, 35, 1–33.

— (2012), *Joint models for longitudinal and time-to-event data: With applications in R*, CRC Press.

Robins, J. M. and Rotnitzky, A. (1992), "Recovery of information and adjustment for dependent censoring using surrogate markers," in *Aids Epidemiology*, Springer, pp. 297–331.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994), "Estimation of regression coefficients when some regressors are not always observed," *Journal of the American Statistical Association*, 89, 846–866.

Shih, W. J. (2002), "Problems in dealing with missing data and informative censoring in clinical trials," *Current Controlled Trials in Cardiovascular Medicine*, 3, 4.

Tsiatis, A. A. and Davidian, M. (2004), "Joint modeling of longitudinal and time-to-event data: an overview," *Statistica Sinica*, 14, 809–834.

Wang, H. and Zhao, H. (2006), "Regression analysis of mean quality-adjusted lifetime with censored data," *Biostatistics*, 8, 368–382.

Wulfsohn, M. S. and Tsiatis, A. A. (1997), "A joint model for survival and longitudinal data measured with error," *Biometrics*, 53, 330–339.

Zeng, D., Cai, J., et al. (2005), "Asymptotic results for maximum likelihood estimators in joint analysis of repeated measurements and survival time," *The Annals of Statistics*, 33, 2132–2163.

Zhao, H. and Tsiatis, A. A. (1997), "A consistent estimator for the distribution of quality adjusted survival time," *Biometrika*, 84, 339–348.

— (2000), "Estimating mean quality adjusted lifetime with censored data," *The Indian Journal of Statistics, Series B*, 62, 175–188.

Zhao, Y. and Wang, H. (2008), "Empirical likelihood inference for the regression model of mean quality-adjusted lifetime with censored data," *Canadian Journal of Statistics*, 36, 463–478.

Received: October 1, 2019

Accepted: February 3, 2020