

Research Article

# A New Approach to Evaluate Quality Adjusted Life Years using Proxy Utility Function - An Application to HIV/ AIDS Data

Vishal Deo<sup>1,2</sup>, Gurprit Grover<sup>2</sup>

<sup>1</sup>Department of Statistics, Ramjas College, University of Delhi, Delhi, India.

<sup>2</sup>Department of Statistics, Faculty of Mathematical Sciences, University of Delhi, Delhi, India.

DOI: <https://doi.org/10.24321/0019.5138.201919>

## I N F O

### Corresponding Author:

Vishal Deo, Department of Statistics, Ramjas College, University of Delhi, Delhi, India.

### E-mail Id:

[vishaaldeo@gmail.com](mailto:vishaaldeo@gmail.com)

### Orcid Id:

<https://orcid.org/0000-0003-3629-0962>

### How to cite this article:

Deo V, Grover G. A New Approach to Evaluate Quality Adjusted Life Years using Proxy Utility Function - An Application to HIV/ AIDS Data. *J Commun Dis* 2019; 51(3): 1-9.

Date of Submission: 2019-08-01

Date of Acceptance: 2019-10-19

## A B S T R A C T

Estimation of Quality Adjusted Life Years (QALYs) is pivotal towards economic evaluation and cost-effectiveness analysis of medical interventions. Most of the methods developed till date for calculating QALYs are based on multi-state structures where fixed utility values are assigned to each disease state and total QALYs are calculated on the basis of total lengths of stay in each state. In this article, we have presented a new proxy approach to define utility as a function of risk factors, which can be used to calculate QALY without defining discrete disease states. Retrospective survival data of HIV/ AIDS patients undergoing treatment at the Antiretroviral Therapy (ART) center of Ram Manohar Lohia hospital in New Delhi has been used to demonstrate implementation of the proposed methodology. Joint modelling, with a mixed effect longitudinal sub-model for CD4 count and a Cox proportional hazard survival sub-model with time dependent covariates, has been used to estimate risks associated with different factors and covariates. Using the proxy utilities, QALYs have been calculated for each individual for their lifetime time horizon, defined as the time since their registration in the ART till death or till their age reach average life expectancy of HIV/ AIDS patients in India. QALY results are consistent with findings of conventional cost-effectiveness studies on ART for HIV/ AIDS patients in India.

**Keywords:** Cost-Effectiveness Analysis, Health Economics, Joint Modelling, HIV/ AIDS, Utility Function

## Introduction

Cost-Effectiveness Analysis (CEA) is an important health economic technique to evaluate economic feasibility of treatment regimens for desired medical outcomes. CEA is hugely dependent on the calculation of gain in Quality Adjusted Life Years (QALYs) in patients after administration of the treatment regime under review. Quality adjustment

is achieved by assigning different utility values to different states of the disease on the basis of physical and psychological appraisal of the patients.

Most of the research works done on CEA till date have used either markov models or multi-state models to map disease progression in terms of transitions between pre-defined states. Survival models are used to fit the survival data and

hence, to estimate the transition probabilities between different states. The estimated transition probabilities are then used to estimate expected Total Length of Stay (TLOS) in each state. Discrete utility values are assigned to each state and these utility values are used as weights to calculate total expected QALYs. Briggs A and Sculpher M<sup>1</sup> laid down the fundamental structure of applying markov models in CEA. Parametric survival models have been most widely used for estimating transition probabilities between different disease states; see for instance.<sup>2-5</sup> Cox Proportional Hazards (PH) models have been used in the multi-state set up for estimating the transition probabilities by Mihaylova B et al.,<sup>6</sup> and Malehi AS et al.<sup>7</sup> among others. Bayesian parametric models have also been presented as alternatives for fitting multi-state transition models and hence, for estimating the gain in QALYs.<sup>8,9</sup>

Although the methods based on multi state models are uncomplicated and easy to comprehend, they are bound by certain theoretical restrictions. Defining few specific disease states can undermine the dynamic nature of disease progression within each state. For example, suppose we are carrying out CEA of Antiretroviral Therapy (ART) intervention for HIV/ AIDS patients using a multi-state model. The states of HIV/ AIDS patients are defined on the basis of pre specified cut off points of CD4 cell counts.<sup>10,11</sup> Suppose that state one is marked by a CD4 cell count of >500, and state two is marked by a CD4 cell count in the range 351-500. This implies that a patient with CD4 cell count of 499 will be classified in the state two, while a patient with CD4 count of 501 will be considered to be in state one. That is, a patient with CD4 cell count of 501 will be assigned much higher utility value than that of a patient having a CD4 cell count of 499. Also, although probably not significantly, but in the range of 351-500, different levels of CD4 cell count can be associated with different levels of health status. So, our assumption of assigning static utility values to the patients' health quality during the course of their stay in any particular state is also questionable. This simple example testifies to the ignorance of the dynamic nature and subtlety of disease progression in multi-state models for certain diseases or infections.

Common methods of obtaining health utility values, like EQ-5D and SF-6D, are based on questionnaires containing basic questions about physical and psychological status of the patients which are generally filled by the patients themselves. Parameters like mobility, self-care, usual activities, pain/discomfort, anxiety/depression, social functioning etc. form the basis of devising scores or measures based on the feedbacks.<sup>9,12,13</sup> Standardized utilities based on such methods can sometimes be misleading because of their inability to encompass individual level variation in the physical and mental status of patients in any particular state of a disease. Variability in the physical

and mental response of individual patients in any disease state can be either caused by associated factors like smoking, drinking, type of work, socio-economic status, age, existence of other unrelated medical conditions, pollution level in the region etc., or by practically unaccountable or immeasurable factors like personal demeanour, will power, family support and love, physical strength prior to onset of the disease etc. At the same time, convincing patients to fill out the questionnaires can also become challenging in case of certain diseases or for certain states of the diseases, especially when the disease under consideration is unfortunately perceived as a social stigma, like HIV/ AIDS, Leprosy, Dementia and many more. Thus, these standardized methods of calculating health utility values have three major limitations; a. assumption of fixed utility throughout the course of stay in a state of a disease, b. no account for patient to patient variability in health utility in any state of the disease and c. difficulty in reaching out to the patients to respond to the questionnaires when the patients are reluctant to be identified in public, or in situations when patients' judgment cannot be trusted.

In this paper our primary objective is to present an alternative method for calculating QALY without using the multi state model framework. The proposed method will be particularly useful in cases where the event time of interest is observed during follow-up; along with longitudinal covariates. The method involves using joint modelling approach to fit a Cox PH sub-model for the event process and a linear mixed effects sub-model for the associated repeated measures of the covariate (or biomarker). The fitted model is then used to project the longitudinal trajectory of the covariate for each censored case till the lifetime time horizon of the study. Utility is defined as a function of the covariate(s) with their weights calculated on the basis of the corresponding effects estimated through joint modelling. These utilities are case specific, are dynamic and they vary according to the changes in the measures on covariate (s) with time. Since the proposed utility function does not involve certain parameters like mental state, anxiety etc., which are included in the standard methods like EQ-5D and SF-6D, it is better to state such a utility function as a proxy utility function. These utility values are then used to calculate total QALY for each patient, since the time of registration in ART till the event of death (for observed cases) or till the lifetime time horizon (for censored cases).

We have discussed the joint modelling approach and the methodology proposed to calculate proxy utility values and QALYs. A demonstration of the proposed method for calculating QALY using proxy utility functions is presented in this article. Survival and longitudinal data of HIV/AIDS patients undergoing treatment at the Antiretroviral Therapy (ART) center of Ram Manohar Lohia hospital in New Delhi has been used for the analyses.

## Methodology

### Joint Modelling of Longitudinal and Time-to-Event Data

Many at times, longitudinal data of patients comprises of repeated measures on a time dependent covariate, apart from the time-to-event of interest. For example, in HIV studies patients' follow-up data includes time till the occurrence of the event of interest (death or development of AIDS) and repeated measurements on the biomarker (or covariate) CD4 lymphocyte count. Separate modelling of the longitudinal process and the time-to-event process fails in establishing association between the longitudinal progression of the covariate and the occurrence of the event. When our interest lies in assessing the impact of the time-dependent covariate on the time-to-event (or survival time), joint modelling serves as a better approach. Separate modelling does not take into account the fact that longitudinal measures on the time-dependent covariate are interrupted by the occurrence of the event, thus introducing bias in the estimation.<sup>14</sup> Joint modelling approach overcomes this problem by maximizing the log-likelihood corresponding to the joint distribution of the longitudinal and time-to-event outcomes. Lucid descriptions of this approach can be found in the works of Tsiatis AA, Davidian M,<sup>15</sup> Yu M et al.<sup>16</sup> and Rizopoulos D.<sup>17</sup>

### Joint Model Specification

#### List of notations used in specifying the sub-models

**n**: number of cases (patients).

**T<sub>i</sub>\***: True observed event time for i-th patient.

**C<sub>i</sub>**: Censoring time for i-th patient.

**T<sub>i</sub>**: event time for the i-th patient, where  $T_i = \min(T_i^*, C_i)$

**δ<sub>i</sub> = I(T<sub>i</sub>\* ≤ C<sub>i</sub>)**; is the event indicator which takes value 1 when the event has occurred and 0 when observation is censored.

**y<sub>i</sub>(t)**: value of the observed/ measured longitudinal outcome for patient i at time t.

**t<sub>ij</sub>**: j-th occasion (time point) at which longitudinal response variable is observed for i-th patient. (j = 1, 2, ..., n<sub>i</sub>) i.e. number of times, and the time points at which, longitudinal responses are recorded for a patient can differ from patient to patient.

**y<sub>ij</sub>**: value of the longitudinal outcome for i-th patient at t<sub>ij</sub>.

**m<sub>i</sub>(t)**: true (unobserved) value of the longitudinal outcome for i-th patient at time t. It is different from y<sub>i</sub>(t) as it is assumed to be free of any measurement error which is present in y<sub>i</sub>(t).

Longitudinal sub-model: Linear mixed effects model.

Using the notations defined in the list of notations above, linear mixed effects model for i-th subject can be defined as follows:<sup>17</sup>

$$y_i(t) = m_i(t) + \epsilon_i(t) \\ = x_i^T(t)\beta + z_i^T(t)b_i + \epsilon_i(t), \epsilon_i(t) \sim N(0, \sigma^2) \dots (1)$$

Where,  $\beta$  denotes the vector of the unknown fixed effects parameters,  $b_i$  denotes a vector of random effects,  $x_i(t)$  and  $z_i(t)$  denote row vectors of the design matrices for the fixed and random effects respectively and  $\epsilon_{-i}(t)$  denotes the measurement error term which follows Normal distribution and is assumed to be independent of  $b_i$ . As a standard choice, the random effects are assumed to follow multinomial normal distribution.

Survival sub-model: Time dependent Cox PH

Again, using the notations mentioned in above, the survival sub-model can be defined as follows:<sup>17</sup>

$$h_i(t|m_i(t), w_i) = h_0(t)\exp(\gamma^T w_i + \alpha m_i(t)) \dots (2)$$

Where,  $h_0(t)$  denotes the baseline risk function,  $w_i$  is a vector of baseline covariates and  $\gamma$  is a vector of coefficients corresponding to these baseline covariates.  $\alpha$  measures the effect of the longitudinal outcome variable,  $m_i(t)$ , on the risk of occurrence of the event and is known as the association parameter. The name 'association parameter' emphasizes on the fact that this parameter measures the extent of association between the course of the longitudinal outcome variable and the risk of the event of interest. Association parameter helps in structural implementation of the idea of joint modelling of the longitudinal and survival processes.

Parameters of both sub-models are estimated jointly using maximum likelihood estimation by maximizing the joint likelihood function of the longitudinal and survival components. The joint log-likelihood contribution for i-th patient can be stated as:

$$\log p(T_i, \delta_i, y_i; \theta) = \log \int p(T_i, \delta_i | b_i; \theta_t, \beta) [\prod_j p\{y_i(t_{ij}) | b_i; \theta_y\}] p(b_i; \theta_b) db_i \dots (3)$$

Likelihood contribution  
from the survival

Likelihood contribution  
from the longitudinal

Where,  $\theta = (\theta_t^T, \theta_y^T, \theta_b^T)^T$  denotes the vector of unknown parameters, with  $\theta_t$  denoting the parameters for event time outcome,  $\theta_y$  denotes the fixed parameters for longitudinal outcomes, and  $\theta_b$  denotes the unique vector of parameters of the random-effects covariance matrix. Here  $p(\cdot)$  denotes an appropriate probability density function. The likelihood contribution of the survival component can be written as,

$$p(T_i, \delta_i | b_i; \theta_t, \beta) = \{h_i(T_i | m_i(T_i); \theta_t, \beta)\}^{\delta_i} S_i(T_i | m_i(T_i); \theta_t, \beta) \dots \quad (4)$$

Where, the first term represents likelihood for the observed cases and the second term represents contribution for the censored cases. R package JM, given by Rizopoulos D, provides a hybrid optimization procedure for maximizing the joint log-likelihood function given in equation (3).<sup>17</sup> This procedure, under the method “spline-GH-PH”, starts with Expectation-Maximization algorithm, and if convergence is not attained, switches to quasi-Newton algorithm until convergence is achieved.

### Projection of Expected Longitudinal Response using the Fitted Model

Next step involves projecting progression of the longitudinal covariate at regular time intervals for censored cases, preferably at each time unit, until the Lifetime Time Horizon (LTH) of the study. LTH has a significant role in CEA as its choice can impact the estimation of overall gain in QALY and Incremental Cost Effectiveness Ratio (ICER) corresponding to the treatment under consideration. In studies involving a disease which has high rate of occurrence of the event, or for which it is possible to use quantile of the survival function to estimate the number of years after which most of the patients will experience the event of interest, time point corresponding to a quantile below 10% (say) of survivability can be taken as the LTH of the study. For instance, in a cancer study if we know that after 15 years of chemotherapy 95% of the patients are expected to be dead, we can choose the LTH for CEA as 15 years since the start of the treatment. For diseases where rate of occurrence of the event of interest is low making it difficult to calculate lower quantiles of survivability, life expectancy of the patients as reported by credible published works in the region of study can be taken as a parameter for fixing the LTH of the study. In such cases, CEA for each patient will be carried out for the period since the start of their treatment till the time their age reaches the average life expectancy; the period being taken as the LTH. We will discuss the implementation of the second case here. Ideally, sensitivity of CEA towards the choice of LTH should be verified by comparing CEA results at different possible values of LTH.<sup>18</sup>

Let  $L$  be the average life expectancy of patients undergoing

the treatment under consideration and let the baseline age of  $i$ -th patient at the time of start of the treatment be  $AGE_i$ . Then, the total length of time for  $i$ -th patient, since the start of treatment, for which QALY has to be calculated is given as:

$$E_i = L - AGE_i \dots \quad (5)$$

Here,  $E_i$  can be regarded as the LTH of the  $i$ -th patient. Let,  $R_i = E_i - T_i$ , where  $T_i$  is defined in the list of notations. Then  $R_i$  denotes the total length of time beyond  $T_i$  (observed period) for which the longitudinal covariate has to be projected for the  $i$ -th patient. To incorporate both censored and observed cases, we define  $R_i$  as follows;

$R_i = E_i - T_i$ ; if the case is censored i.e. the patient is still alive (hasn't experienced the event).

$T_i$ ; if the case is observed i.e. the patient has experienced the event or if  $L \leq AGE_i \dots \quad (6)$

### Steps for Projection/ Prediction

- Let  $\widehat{\theta}_y$  be the vector of estimates of the parameters of the fixed effects and  $\widehat{\mu}_b$  be the vector of estimates of the mean component of the random effects.
- Use  $\widehat{\theta}_y$  and  $\widehat{\mu}_b$  to estimate expected values (mean values) of the longitudinal response  $y_{ij}$  (given  $\widehat{\mu}_b$  for each censored patient  $i$  at each time point  $j$  using the following equation.

$$\widehat{m}_i(t) = x_i^T(t) \widehat{\theta}_y + z_i^T(t) \widehat{\mu}_b; t = T_i+1, T_i+2, \dots, R_i \dots \quad (7)$$

### Proxy Utility Function

The proposed proxy utility function has been defined as a function of changes in the longitudinal measures of the time dependent covariate. Effect of change in longitudinal measure on the utility has been derived from the estimated value of the association parameter  $\alpha$  included in the survival sub-model in equation (2). Since the survival sub-model considered in our study is a time dependent Cox-PH model, exponential of the estimated coefficients represents extent of change in odds in favor of the event of interest (death in our case) per unit change in the covariate value. Thus, the association parameter represents the effect of changes in longitudinal covariate on the risk of the event (death). If the estimate is positive, increase in the covariate value results in increase in the odds in favor of the event (death), and if it is negative, odds in favor of the event decreases as the covariate value increases. Also, higher risk of death is associated with lower quality of life. So, inverse of exponential of the association parameter estimate has been taken as the proxy effect of change in the value of time dependent covariate on utility of the patient. Using these arguments, the utility function has been defined as follows.

$$U_i(t) = U_{0i} + \exp(-\widehat{\alpha}) \cdot \sum_t \left[ \frac{y_i(t) - y_i(t-1)}{K} \right] = U_{0i} + \exp(-\widehat{\alpha}) \left[ \frac{y_i(t) - y_{0i}}{K} \right] \dots \quad (8).$$

Where,

$U_i(t)$ : Utility value of  $i^{\text{th}}$  patient at time  $t$ .

$U_{oi} = \frac{y_{oi}}{K}$ , where  $y_{oi}$  is the observed value of the time dependent covariate at baseline.

$K$ : Cut-off value or reference value of the longitudinal covariate beyond which it is considered as to be in medically normal range. If no standardized clinical value is available, it can also be taken as equal to the average covariate values of people free from the disease/ infection under consideration.

$y_i(t)$ : Observed or predicted value of the longitudinal covariate at time  $t$  for the patient  $i$ .

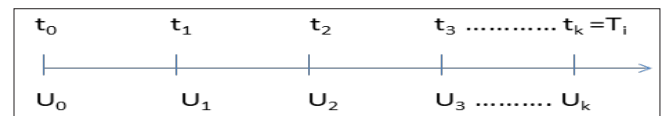
$\hat{\alpha}$  Estimate of the association parameter obtained from joint modelling.

Clearly, in the formula for calculating proxy utility represents the baseline proxy utility value at the time of registration and  $\exp(-\hat{\alpha}) \left[ \frac{y_i(t) - y_{oi}}{K} \right]$  represents the change in proxy utility value as a result of the change in the value of the longitudinal covariate. Using the function defined in equation (8), utility values of each patient can be calculated at every time point at which measurement on the time dependent covariate is either available or predicted. At any point in time, if the calculated utility value exceeds 1, it is taken as equal to 1. The utility function defined in (8) depicts the effect of relative change in the longitudinal covariate, with respect to its reference value, on utility. It should be noted that the functional form of the proxy utility does not assign equal utility values at equal values of the longitudinal covariate for two different subjects. Rather, it indirectly incorporates the association of relative changes in covariate values over time with the effectiveness of the treatment. If patients enter the study at different time points, and have different baseline covariates and different disease status (in terms of severity) at the time of registration, expecting equal utility values for two individuals only because of same CD4 count after a fixed time since registration will not be justified. The way in which the proxy utility function has been defined reflects on the quantum of effect of treatment on the subject through the inclusion of relative changes in CD4 cell count. Since our ultimate objective is to quantify the effectiveness of treatment over time, such definition of proxy utility shall lead us to logical conclusions.

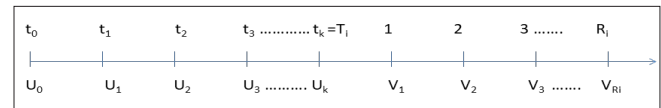
Equation (8) can be extended for more than one-time dependent covariates also. For instance, in a study on cost effectiveness of ART for HIV/ AIDS patients, changes in both CD4 count and body weight measured at each observation point can be used to define the proxy utility function. In such cases, the vector of weights of the variables included in the function can be normalized to ensure that the utility values remain in the range [0,1].

## Calculation of QALY

Utility values can be classified into two categories; the ones calculated on the basis of observed covariate values and the ones calculated on the predicted values of the covariate. Let us denote the first category by  $U$  and the second category by  $V$ . Suppose that for censored cases, value of the time dependent covariate has been predicted at every time unit beyond the last observation time till the average life expectancy  $L$  i.e. for  $R_i$  time units after  $T_i$ ; refer list of notations. The timeline of utility values for observed and censored cases can be visualized in the figures 1 and 2 respectively.



**Figure 1.** For patients who have experienced the event or whose baseline age  $\geq L$



**Figure 2.** For patients who have not experienced the event till  $T_i$

Based on these utility values, QALY for each patient is calculated as follows:

For observed cases:

$$QALY = \frac{1}{12} \sum_{i=0}^{k-1} \{U_i(t_{i+1} - t_i)(1 + d)^{-t_{i+1}}\} \dots (9)$$

Here  $d$  is the discount rate per annum converted per unit time, say monthly.

For censored cases:

$$QALY = \frac{1}{12} \sum_{i=0}^{k-1} \{U_i(t_{i+1} - t_i)(1 + d)^{-t_{i+1}}\} + U_k(1 + d)^{-t_k} + \sum_{j=1}^{R_i-1} V_j(1 + d)^{-(t_k+j+1)} \dots (10)$$

## Implementation of the Proposed Methodology

As a demonstration, the methodology proposed in section 2 has been employed to calculate QALY for HIV/ AIDS patients undergoing treatment at the ART center at Ram Manohar Lohia hospital in New Delhi. Retrospective follow-up data on 3465 patients, enrolled for ART treatment during the period April 2004 to November 2009 and followed till December 2010, has been obtained from register records of the hospital. All the records are in terms of patient id and in no way whatsoever the privacy and anonymity of patients is compromised. Cases with incomplete or missing information on factors and covariates included in the model like age, sex, Mode Of Transmission (MOT), smoking habit, drinking habit, hemoglobin, baseline weight etc., have been excluded from the analysis. Only adults of age 18 and above are retained in the study. Cases with no subsequent visit after the date of registration into ART are also not considered

for the analysis. After applying all exclusion criteria, only 1520 cases qualified to be included in the study. Dataset contains information on status (dead or alive), time of death, age at the time of registration (baseline age), sex, smoking habit, alcohol consumption, baseline hemoglobin, mode of transmission of HIV (MOT), time of subsequent visits, CD4 count at each visit, and weight measured at each visit. Use of this data complies with the ethical guidelines defined in the National Ethical Guidelines for Biomedical and Health Research - 2017 for administrative and secondary data.

### Joint Modelling

For the survival sub-model, the event of interest is occurrence of death. Longitudinal measurements on CD4

count act as a significant indicator for assessing progression of HIV infection. So, the longitudinal sub-model is defined to assess the longitudinal progression of CD4 count of patients with respect to various factors and covariates. Since there are repeated measures on CD4 count for each patient, random effects are introduced through a linear mixed effects model to account for the hierarchical structure of the data. Random effects of observation time and intercept has been included in the longitudinal sub-model. The Cox-PH survival sub-model in (2) and the linear mixed effects longitudinal sub-model in (1) are jointly modelled using joint Model () function of the JM package in R. Results of the joint model are provided in the Table 2 and 3. Independent variables included in the models are defined in Table 1.

**Table 1. Description of independent variables included in the models**

S. No.	Variable name	Label	Levels (if a factor)
1.	obstime	Time of observation in months calculated from the date of registration (a time dependent covariate)	NA
2.	SEX	A factor variable with two levels, Male and Female	0=Female 1=Male
3.	HB	Hemoglobin level observed at baseline	NA
4.	D1_MOT	Dummy variable for mode of transmission of HIV	1=sexually transmitted 0=other modes
5.	D1_smoking	Dummy variable for current status of smoking	1=currently smoke 0=otherwise
6.	D2_smoking	Dummy variable for past status of smoking	1=smoked in past 0=otherwise
7.	D_alcohol	Dummy variable for status of alcohol consumption	1=Yes 0=No
8.	Age	Age at baseline	NA
9.	Wight	Weight measured at each visit (time dependent)	NA

**Table 2. Results of the joint model**

Factor/Covariate	Longitudinal Sub-model		Survival sub-model	
	Coefficient Estimates	p-value	Coefficient estimates	p-value
(Intercept)	108.4910	<0.0001***	-----	
obstime	12.1841	<0.0001***	-----	
SEX (male)	-33.6147	<0.0001***	-0.0329	0.9134
HB	11.5416	<0.0001***	-0.1183	0.0224**
D1_MOT (sexual)	-17.2330	0.0368**	-0.4348	0.1379
D1_smoking (currently)	8.2168	0.3860	0.4751	0.2413
D2_smoking (past)	-1.5441	0.8427	0.5077	0.0795*
D_alcohol (yes)	-5.4702	0.4546	0.0318	0.9105
AGE	-1.0397	0.0007***	0.0096	0.4013
wt	-----		-0.0854	<0.0001***
Association	-----		-0.0269	<0.0001***

\*\*\* Significant at 1% level of significance; \*\* Significant at 5% level of significance; \* Significant at 10% level of significance.

**Table 3. Estimates of variance component of random effects in the longitudinal sub-model**

	Standard Deviation
Intercept	61.4849
obstime	6.8313

From Table 2, we can see that the estimate of association parameter between the longitudinal measures on CD4 count and the survival sub-model is highly significant and its negative value implies that increase in the values of CD4 count results in decrease in the hazard of death.

### Projecting CD4 Count for Censored Cases Using the Fitted Longitudinal Sub-Model

#### Steps

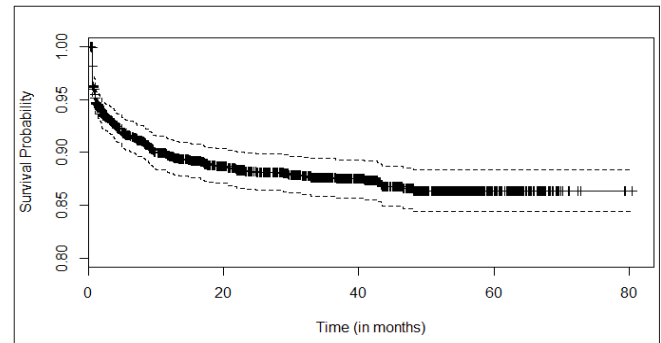
##### Life Expectancy

It is not possible to calculate median survival time or any p-th percentile survival time for  $p \leq 0.5$  from the Kaplan-Meier (KM) curve (figure 3). Also, survival times of each censored patient cannot be directly predicted using Cox-PH survival sub-model. As an alternative, as discussed in section 2.2, we have calculated LTH for each patient on the basis of average life expectancy of HIV/AIDS patients in India. There is a dearth of studies on forecasting life expectancy of HIV patients in India. However, since the interest of this section lies mainly in demonstrating the implementation of the proposed method, we haven't presented any comprehensive scientific review for our choice of average life expectancy of HIV patients under ART in India. But we have also refrained from arbitrary choice of average life expectancy and have employed the following arguments to support our choice. According to the global scenario of life expectancies reported in various studies like May MT et al.,<sup>19</sup> Egger M & Johnson LF,<sup>20</sup> Katz IT & Maughan-Brown B<sup>21</sup> and many others, life expectancy of HIV positive people under ART are getting closer to the HIV negative people across the world, while the closeness may differ from country to country. According to a meta-analysis study done by Teeraananchai S et al.<sup>22</sup> for low/ middle income countries, average life expectancy after starting ART at ages 20 years and 35 years are 28.3 years and 25.6 years respectively. The average baseline age of the patients included in our study, i.e. the average age at which the patients were enrolled into ART, is 34.48 years. Also, it should also be noted that the average life expectancy at birth in India during 2009 was reported to be around 65 years. Based on this information, we have taken the average life expectancy of HIV positive patients undergoing ART for our study as 60 years, i.e.  $L = 60$  years.

1. For each patient we calculate the TMH, i.e.  $E_i$ , using equation (5).  $E_i$  is further used to calculate  $R_i$  using equation (6).

2. Based on the results reported in table 2, expected CD4 count (conditional over estimated random effects) for i-th patient at j-th month is projected using the following function.

$$CD4_{ij} = 108.4910 + 12.1841 \cdot obstime_{ij} - 33.6147 \cdot SEX_i + 11.5416 \cdot HB_i - 17.2330 \cdot D1\_MOT_i + 8.2168 \cdot D1\_smoking_i - 1.5441 \cdot D2\_smoking_i - 5.4702 \cdot D\_alcohol_i - 1.0397 \cdot AGE_i$$



**Figure 3. KM curve for survival of the HIV/AIDS patients**

#### Utility Calculations

On examining the findings of Rungta A et al.,<sup>23</sup> Shahapur PR et al.,<sup>24</sup> Ramalingam S et al.<sup>25</sup> and Uppal SS et al.,<sup>26</sup> the reference value of CD4 count for a healthy person in India is taken as 800 on an average (rounded off), i.e.  $K = 800$ . Let  $y_{0i}$  be the baseline CD4 count of the i-th patient. Then we define

$$U_{0i} = \frac{y_{0i}}{800}$$

Using equation (8), the proxy utility function for i-th patient at time  $t$  can be written as,

$$U_i(t) = U_{0i} + \exp(-0.0269) \cdot \sum_t \left[ \frac{y_i(t) - y_i(t-1)}{800} \right] = U_{0i} + \exp(-0.0269) \left[ \frac{y_i(t) - y_{0i}}{800} \right] \dots (11).$$

For all patients, proxy utility values are calculated at each observation time (visit time) till  $T_i$  using equation (11). For censored cases, proxy utility values are further calculated at each month from  $T_i + 1$  till  $R_i$  using the same function. It should be noted that till  $T_i$ , between any two successive time points at which observed values of CD4 counts are available, utility values are assumed to be constant.

#### QALY Calculations

A discount rate of 3% per annum has been used for the calculation of QALY. Converting this rate into effective monthly rate, we have  $d = .03 / 12 = 0.0025$ . Using this value of  $d$ , and the utility values calculated in the previous section in the formulae (9) and (10), we get the total QALY gained by each patient after getting registered in the ART centre. For instance, Total QALY gained by first ten patients came out as 9.802356, 12.657515, 9.425577, 10.003628, 11.416873, 14.433912, 14.662325, 13.053493, 15.310978, and 14.126491 years respectively. The average of total QALYs of all patients is calculated to get the mean QALY

gain due to ART treatment. The mean QALY has come out to be 11.18045 years.

### Programming

Algorithms for all calculations described in the sections 3.2, 3.3, and 3.4 have been indigenously implemented in R without using any pre-existing package.

### Discussion and Conclusion

Based on the results of section 3.4, we can conclude that, on an average, patients gain 11.18045 QALYs, discounted at 3% per annum, after getting enrolled in ART till the time they reach the average life expectancy of HIV patients. This result is consistent with, or at least not much different from, the findings of CEA of ART in India done by Bender MA et al.<sup>27</sup> Bender MA et al.<sup>27</sup> found that for different combinations of drugs in ART, the average gain in QALY discounted at 3% per annum ranged from 115.5 months (9.625 years) to 125.8 months (10.48 years). It should be noted that the authors have used state specific utility values taken from studies in United States, which is a developed country as opposed to India. Also, calculations of QALY in their study are based on a state transition model or a multi-state model.

The method proposed in this paper is specifically applicable when the survival data is accompanied with longitudinal measurements on certain covariates. In such cases, this method incorporates the advantages of joint modelling of survival and longitudinal models over separate modelling. Further, it eliminates the need to define different states of the disease and to assume that utility values of all patients are equal and fixed within each state. Thus, it provides the scope of introducing patient level variability in the course of utilities, even within any particular state of the disease. Since the utility values of each patient can be calculated at each time point, we can calculate total QALY gained by each patient separately, which is not possible in the case of multi state model-based CEA. Apart from these benefits, proxy utility function allows us to calculate utilities without spending extra resources on conducting surveys for using standard methods like EQ-5D and SF-6D. Also, proxy utility function provides much more flexibility and dynamics by avoiding standardization of state specific utility values across different regions, socio-economic strata, ages, gender, opportunistic risks and many such factors of patient level variability.

The fact that psychological state of the patients cannot be taken into account while defining the proxy utility function, can be considered as a limitation of the proposed method. To include psychological state in the calculation of proxy utilities some measures like scores of psychometric tests will be required. Although possible, this exercise will have to be carried out during each visit of the patients, contrary to the objective of the proposed method.

### Acknowledgement

We are extremely grateful to Dr. A. K. Gadpayle, Ram Manohar Lohia Hospital, New Delhi, for allowing us to use the retrospective data for demonstrating the implementation of our proposed methodology. Our team has published significant number of research papers in the past based on the same dataset.

**Conflict of Interest:** None

### References

1. Briggs A, Sculpher M. An introduction to markov modelling for economic evaluation. *Pharmacoeconomics* 1998; 13(4): 397-409.
2. Speight PM, Palmer S, Moles DR et al. The cost-effectiveness of screening for oral cancer in primary care. *Health Technol Assess* 2006; 10(14): 1-144.
3. Coon JT, Hoyle M, Green C et al. Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation. *Health Technol Assess* 2010; 14(2): 1-184.
4. Wu B, Li T, Cai J et al. Cost-effectiveness analysis of adjuvant chemotherapies in patients presenting with gastric cancer after D2 gastrectomy. *BMC Cancer* 2014; 14: 984.
5. Diaby V, Adunlin G, Montero AJ. Survival modeling for the estimation of transition probabilities in model-based economic evaluations in the absence of individual patient data: a tutorial. *Pharmacoeconomics* 2014; 32(2): 101-108.
6. Mihaylova B, Briggs A, O'Hagan A et al. Review of statistical methods for analysing healthcare resources and costs. *Health Economics* 2011; 20: 897-916.
7. Malehi AS, Pourmohammadi F, Angali KA. Statistical models for the analysis of skewed healthcare cost data: a simulation study. *Health Economics Review* 2015; 5-11.
8. Welton NJ, Ades AE. Estimation of Markov Chain Transition Probabilities and Rates from Fully and Partially Observed Data: Uncertainty Propagation, Evidence Synthesis, and Model Calibration. *Medical Decision Making* 2005; 25(6): 633-645.
9. Baio G. *Bayesian Methods in Health Economics*. Chapman & Hall/CRC Biostatistics Series, 2013.
10. Grover G, Gadpayle AK, Swain PK et al. A Multistate Markov Model Based on CD4 Cell Count for HIV/AIDS Patients on Antiretroviral Therapy (ART). *International Journal of Statistics in Medical Research* 2013; 2: 144-151.
11. Shoko C, Chikobvu D. Time-homogeneous Markov process for HIV/AIDS progression under a combination treatment therapy: cohort study, South Africa. *Theor Biol Med Model* 2018; 15: 3.
12. Kind P, Brooks R, Rabin R. *EQ-5D Concepts and Methods: A Developmental History*. Springer, Dordrecht. 2005.



13. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 2002; 21(2): 271-292.
14. Tseng YK, Hsieh F, Wang JL. Joint modelling of accelerated failure time and longitudinal data. *Biometrika* 2005; 92(3): 587-603.
15. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* 2004; 14: 809-834.
16. Yu M, Law NJ, Taylor JMG et al. Joint Longitudinal-Survival-Cure Models and Their Application to Prostate Cancer. *Statistica Sinica* 2004; 14: 835-862.
17. Rizopoulos D. JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software* 2010; 35(9).
18. Jackson C, Stevens J, Ren S et al. Extrapolating survival from randomized trials using external data: a review of methods. *Medical Decision Making* 2017; 37: 377-390.
19. May MT, Gompels M, Delpech V et al. Impact on life expectancy of HIV-1 positive individuals of CD4R cell count and viral load response to antiretroviral therapy. *AIDS* 2014; 28(8): 1193-1202.
20. Egger M, Johnson LF. Estimating trends in life expectancy in HIV-positive individuals. *The Lancet Global Health* 2015; 3: e122-23.
21. Katz IT, Maughan-Brown B. Improved life expectancy of people living with HIV: who is left behind? *The Lancet HIV* 2017; 4: e324-6.
22. Teeraananchai S, Kerr SJ, Amin J et al. Life expectancy of HIV-positive people after starting combination antiretroviral therapy: a meta-analysis. *HIV Medicine*, 2016.
23. Rungta A, Hooja S, Vyas N et al. Enumeration of CD4 and CD8 T lymphocytes in healthy HIV seronegative adults of northwest India: A preliminary study. *Indian Journal of Pathology and Microbiology* 2008; 51(1): 127-129.
24. Shahapur PR, Bairy I, Shivananda PG. CD4 and CD8 reference counts in normal healthy south indian adults. *Indian Journal of Medical Microbiology* 2008; 26(3): 280-290.
25. Ramalingam S, Kannangai R, Zachariah A et al. CD4 counts of normal and HIV-infected south Indian adults: Do we need a new staging system? *The National Medical Journal of India* 2001; 14: 335-339.
26. Uppal SS, Verma S, Dhot PS. Normal Values of CD4 and CD8 lymphocyte subsets in healthy indian adults and the effects of sex, age, ethnicity, and smoking. *Cytometry Part B: Clinical Cytometry* 2003; 52B: 32-36.
27. Bender MA, Kumarasamy N, Mayer KH et al. Cost-effectiveness of tenofovir as first-line antiretroviral therapy in India. *Clinical Infectious Diseases* 2010; 50: 416-425.