



Published in final edited form as:

Stat Methods Med Res. 2017 December ; 26(6): 2667–2680. doi:10.1177/0962280215605107.

An ensemble survival model for estimating relative residual longevity following stroke: Application to mortality data in the chronic dialysis population

Milind A Phadnis¹, James B Wetmore^{2,3}, Theresa I Shireman^{4,5}, Edward F Ellerbeck⁴, and Jonathan D Mahnken¹

¹Departments of Biostatistics, University of Kansas School of Medicine, Kansas City, KS, USA

²Division of Nephrology, Hennepin County Medical Center, Minneapolis, MN, USA

³Chronic Disease Research Group, Hennepin County Medical Center, Minneapolis, MN, USA

⁴Preventive Medicine and Public Health, University of Kansas School of Medicine, Kansas City, KS, USA

⁵The Kidney Institute, University of Kansas School of Medicine, Kansas City, KS, USA

Abstract

Time-dependent covariates can be modeled within the Cox regression framework and can allow both proportional and nonproportional hazards for the risk factor of research interest. However, in many areas of health services research, interest centers on being able to estimate residual longevity after the occurrence of a particular event such as stroke. The survival trajectory of patients experiencing a stroke can be potentially influenced by stroke type (hemorrhagic or ischemic), time of the stroke (relative to time zero), time since the stroke occurred, or a combination of these factors. In such situations, researchers are more interested in estimating lifetime lost due to stroke rather than merely estimating the relative hazard due to stroke. To achieve this, we propose an ensemble approach using the generalized gamma distribution by means of a semi-Markov type model with an additive hazards extension. Our modeling framework allows stroke as a time-dependent covariate to affect all three parameters (location, scale, and shape) of the generalized gamma distribution. Using the concept of relative times, we answer the research question by estimating residual life lost due to ischemic and hemorrhagic stroke in the chronic dialysis population.

Reprints and permissions: sagepub.co.uk/journalsPermissions.nav

Corresponding author: Milind A Phadnis, Departments of Biostatistics, University of Kansas School of Medicine, 3901 Rainbow Boulevard, Kansas City, KS 66160, USA. mphadnis@kumc.edu.

Authors' Note

The data reported here have been supplied by the United States Renal Data System (DUA#2007-10 & 2009-19) and the Centers for Medicare & Medicaid Services (DUA#19707). The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy or interpretation of the U.S. government.

Conflict of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Keywords

Additive hazards; generalized gamma; relative times; residual median life; semi-Markov model

1 Introduction

Of the standard approaches available for the analysis of right censored survival data in biomedical applications, the semi-parametric Cox proportional hazards model¹ is the most popular. This is partially due to the fact that in such applications research interest often centers on being able to estimate the relative risk of an event (say death) for patients belonging to one group as compared to patients belonging to another group. Here, group membership can be attributed to the type of treatment received (e.g. treatment versus placebo) or in many cases to the presence or absence of a particular demographic (gender, race, BMI group) or biological condition (cancer versus no cancer). The main strength of this approach is that through maximization of a partial likelihood, the estimated regression coefficients can be exponentiated to accord direct clinical interpretation of hazards relative to group membership irrespective of how hazards in each group change as a function of time. Proportionality of hazards also implies that the survival functions in the groups are related to each other by means of a power function, and this allows plotting survival profiles for patients belonging to each group when a baseline survival function is modeled or assumed. The validity of the proportional hazards assumption for covariates of interest can be assessed using standard approaches discussed in literature.² When this assumption is violated, clinical interpretation of relative hazard is more involved than in the former case and alternative methods of analyses may be preferred.

Parametric methods of survival analyses permit the flexibility of modeling a variety of hazard shapes that can be described once the parameters specifying the particular distribution have been estimated by maximization of the overall likelihood function. With the advancement of statistical software, many parametric distributions can be used to model survival data thereby facilitating between-group comparison for individuals exposed to a condition as opposed to individuals not exposed to that condition through the concept of relative times (RT). Specifically, RT is the ratio of times that a given proportion (or percent) of individuals belonging to one group experience an event of interest as compared to individuals not belonging to that group. Parametric models, therefore, can have a bigger impact in those biomedical applications where comparative estimation of event-free survival times for individuals exposed to a medical condition (say group-1) versus individuals not exposed to this condition (group-2) is the focus of research interest. For example, in certain cancer studies researchers may be more interested in estimating the residual longevity of patients possessing certain characteristics compared to patients who do not possess those characteristics. The hazards in patients with and without these characteristics may change differentially across time making the quantification of such comparative estimation difficult—even with the allowance of time-dependent covariates in a Cox regression framework. In such applications, parametric models can be deployed to answer the research question provided reasonably accurate inference can be drawn about the various hazard shapes attributable to the estimated parameters of these distributions. Additionally, hazard functions

for the two groups of individuals can also be plotted over time by deriving these functions from the parametric survival function.

Our work is primarily motivated by research interest in estimating residual longevity in the chronic dialysis population due to two different types of stroke: hemorrhagic and ischemic. As the two types of strokes are physiologically different, it is hypothesized that they will confer completely different survival experience—both from each other and from those not experiencing either event. Here, stroke is a time-dependent covariate, that is, dialysis patients who enter the cohort do so with a “no stroke” status that may or may not change to one of the two “stroke” (hemorrhagic or ischemic) statuses at some point in the observation window. Note that some subjects may have incurred a stroke prior to dialysis initiation, but as Medicaid data is not available on all patients prior to dialysis initiation, we have focused our interest in studying the impact of ischemic and hemorrhagic strokes that occur while subjects are on dialysis. A priori, it is not known whether stroke affects patient survival through its time-dependent occurrence alone or its effect is accentuated by time spent since the stroke occurred. From a clinical perspective, doctors would like to inform patients as to the expected median residual life lost (RLL) due to the occurrence of each type of stroke, allowing also for adjustment for a host of demographic risk factors and comorbidities. Simultaneously, researchers would prefer to retain conventional hazard ratio (HR) interpretations for these accompanying risk factors. Another research interest is to study and compare the hazard ratios, which are potentially nonconstant of the two types of strokes over time so as to evaluate the differential relative risks that the two stroke types accord over time.

In order to answer the research question, we adopt an ensemble approach that combines various aspects of standard and advanced survival models. Section 2 summarizes information related to the motivating example. Section 3 provides model building justification for our model. In section 4, we briefly discuss the three-parameter generalized gamma (GG) distribution and its properties in the context of our example. Main results related to RT calculations are discussed in section 5. In section 6, we summarize the results and discuss potential extensions to our approach.

2 Motivating example

In this section, we briefly summarize background information related to this project; details of cohort construction are given in Wetmore et al.³ Briefly, a large cohort with 69,371 patients representing the chronic dialysis population in the time period 2000–2005 who were eligible for receiving both Medicare and Medicaid was constructed by linking the United States Renal Data System⁴ (USRDS) to Medicaid claims data. The research team developed clinical algorithms that allowed identification of incident hemorrhagic and ischemic strokes from Medicare claims data.^{5,6} Follow-up for this cohort began upon dialysis initiation (plus 90 days, as is standard with USRDS data). Subjects were followed until death, the outcome of interest, and subjects were right censored in the case they lost their Medicare or Medicaid coverage, received a kidney transplant, or were otherwise lost to follow-up (follow-up ended on 31 December 2005). Stroke and atrial fibrillation were treated as time-dependent covariates whereas all other risk factors were treated as time-independent baseline

covariates. Identification of stroke type (hemorrhagic or ischemic) was done using international classification of diseases-9 (ICD-9) codes, and for those with more than one stroke event, only the first event was used. This approach allowed identification of 2391 ischemic and 534 hemorrhagic strokes. As the main research question concerns estimation of residual lifetime lost due to the two stroke types, all subsequent section results will be discussed mainly with respect to stroke as a time-dependent covariate; the adjustment for other risk factors only being discussed in the model building part of sections 3 and 4.

3 Semi-Markov model with additive hazard extension

3.1 Initial model building assessments

This section focusses on the model building procedure with reference to using a semi-Markov model with an additive hazard extension for our problem. As mentioned in section 1, a priori it is not known whether relative risk of death due to the occurrence of a stroke is affected by time at which stroke occurs. A Kaplan–Meier (KM) curve (see Figure 1) shows the survival profile of patients suffering hemorrhagic and ischemic stroke after occurrence of the corresponding stroke type (i.e. with time axis representing "time since stroke"). When stratified by year on dialysis of stroke occurrence (see Figure 2(a) for hemorrhagic, Figure 2(b) for ischemic), the corresponding log-rank tests suggest that stroke vintage is not associated with time to death following a stroke (p -values are 0.530 and 0.401, respectively). This further suggests that the predicted survival profiles following a stroke—regardless of when the stroke occurred—from our final model should somewhat resemble the shape of the KM curves in Figure 1 when the values of all covariates are adjusted to match the overall profile of the patients that experienced strokes in the cohort. The goodness of fit for any intermediate model choice can, therefore, be assessed diagnostically in terms of visual comparison of model-based predicted survival curves to the corresponding KM curve in Figure 1. Such an approach of diagnostic assessment in the time-independent covariate scenario has been adopted by Storer et al.⁷ for their Cox PH type model. To summarize to this point, we assumed time of stroke occurrence is not associated with subsequent mortality, and have identified "observed" plots to utilize for assessment of subsequent models built.

As a next step, we deployed the standard Cox PH model with stroke and atrial fibrillation incorporated as time-dependent covariates. Without loss of generality we describe how stroke was built into this phase of the modeling by including only parameters for stroke using

$$h(t|Z) = h_0(t) \exp\{\beta_1 Z_t + \beta_2 Z_t \cdot f(t) + \beta_3 X_1 + \beta_4 X_2 + \dots + \beta_k X_{k-2}\} \quad (1)$$

Here, X_1, X_2, \dots, X_{k-2} are the other risk factors including atrial fibrillation, $\beta_3, \beta_4, \dots, \beta_k$ are the regression coefficients associated with these other risk factors, Z_t is the time-dependent indicator variable for stroke (1 if stroke occurs, else 0), β_1 is the regression coefficient associated with Z_t , whereas $Z_t \cdot f(t)$ is the time-dependent term representing a functional interaction form between stroke and time, and β_2 is the regression coefficient associated with $Z_t \cdot f(t)$. Note that we have used the subscript t on Z to indicate the time-dependency of this variable. As we have assumed that time at which stroke occurs does not

affect survival, we can replace $Z_t \cdot f(t)$ by $Z_t \cdot f(t - s)$ where $s =$ time at which stroke occurs implying that post stroke, survival is affected only by: (1) the fact that stroke has occurred (Z_t) and (2) the time spent since occurrence of stroke $Z_t \cdot f(t - s)$.

For an indicator variable Z_t (equals 1 once stroke occurs, 0 otherwise), different functional forms of $Z_t \cdot f(t - s)$ can be modeled with $Z_t \cdot f(t - s) = Z_t \cdot (t - s)$ and $Z_t \cdot \log(t - s)$ being the most common. Hougaard⁸ has discussed how in the former case for $s = 0$, $h(t|Z) = h_0(t) \exp(\beta_1 Z_t + \beta_2 Z_t \cdot t)$ yields a relative risk function that has the same form as that of a Gompertz parametric model with different values of the shape parameter for the two groups (stroke versus no stroke). Likewise, in the latter case for $s = 0$, $h(t|Z) = h_0(t) \exp\{\beta_1 Z_t + \beta_2 Z_t \cdot \log(t)\}$ yields a relative risk function that has the same form as that of a Weibull model with different values of the shape parameter for the two groups (stroke versus no stroke). Substituting $s = 0$ and generating a predicted survival profile for patients who experience a stroke (this is done for both stroke types), we find that both functional forms prove inadequate in passing the graphical diagnostic check mentioned above. This implies that although stroke as a time-dependent covariate can be incorporated in the Cox regression framework, it is difficult to find the appropriate functional form that satisfies the diagnostic checking procedure. Thus, testing for the proportionality of hazard using the Therneau—Grambsch nonproportionality test⁹ alone is not enough as a significant result for this test might be indicative of a number of possible model failures. As noted by Keele,¹⁰ the correct interpretation of a significant test result on the nonproportionality test is not evidence that the hazards are not proportional but instead that if the covariate-dependent specification is correct, then there is evidence that hazards are not proportional. Thus, when the model is misspecified, the Therneau—Grambsch test may yield a false positive result.^{11,12} The latter paper also discusses importance of incorporating possible nonlinear covariate effects as well as considering other parametric survival models that result in nonproportional hazards. In our paper (in the next section), we therefore, use the GG family of distributions to model the nonproportional hazards part of our final model.

3.2 A state transition model building approach

One way to handle this problem is to move from a time-dependent covariate approach to an equivalent state-dependent transition approach. That is, we exploit the fact that a single time-dependent covariate (in this case stroke) can also be expressed as a single intermediate state transition model as represented in Figure 3. With reference to this figure, State-1 is the initial state representing entry into cohort, State-2 is the single intermediate transition state representing stroke occurrence, and State-3 is the single absorbing state signaling death. Entry into cohort takes place at time $t = 0$ from which an individual can progress to death at time t either directly (those with no strokes) or through the single transitional State-2 (a stroke) occurring at time s such that $t \geq s$. Individuals who do not experience the death are considered right censored at time t . Under the Markov assumption which states that what happens to an individual after he has visited a particular state depends on their current state and not on the preceding history, it can be shown¹³ that the transition hazard $h_{ij}(t)$ for moving from state i to state j can be modeled as

$$h_{i3}(t) = h_{3_0}(t) \exp \left\{ \beta'_{13} \mathbf{X} + (i-1) \Delta' \mathbf{X} + (i-1) \delta \right\} \quad i=1,2 \quad (2)$$

Here, $h_{3_0}(t)$ represents some baseline hazard attributable to moving into State-3, β_{13} represents the vector of regression coefficients for the other risk factors, \mathbf{X} is the vector of other risk factors, Δ represents the difference in covariate effects of these other risk factors for death after having stroke, compared to before having stroke, and δ represents the estimate of the stroke effect under the Markov assumption. This stringent Markov assumption, however, can be relaxed to form a semi-Markov model by allowing the sojourn times (incorporated through covariates) to depend on the history of the process through: (i) the current state itself and (ii) time spent in that state. Thus a semi-Markov model allows us the flexibility to incorporate the influence of time spent since stroke on the hazard of death for those that have had a stroke. It should be noted that though a semi-Markov model can be further extended to form an extended semi-Markov model (where in addition to the semi-Markov assumptions, entry time in a particular state can also affect hazard of death), such an extension is unnecessary in our case due to the assumption that the time at which stroke occurs does not play a role toward affecting patient survival from the time of the stroke until death.

In the context of Figure 3 and the fact that stroke is a time-dependent (and equivalently state-dependent) covariate, meaning the contribution of $(t-s)$ to the hazard of death is valid only for those had strokes (as those that have not had them do not pass through State-2), the semi-Markov model can be augmented with an additive hazards component in the following way

$$h_{i3}(t) = h_{13_0} \{ \min(t, s^-) \} \cdot \exp \left(\beta'_{13} \mathbf{X} \right) + h_{23}(t-s) \cdot I(t \geq s) \quad i=1,2 \quad (3)$$

where s^- is the instant just before the occurrence of a stroke at time $= s$ and I is the indicator function that equals 1 if $t \geq s$ and 0 otherwise.

Thus for those with no strokes, $i=1$ and $t < s$ resulting in

$$h_{13}(t) = h_{13_0}(t) \cdot \exp \left(\beta'_{13} \mathbf{X} \right) \quad (4)$$

and for those with strokes, $i=2$ and $t \geq s$ resulting in

$$h_{23}(t) = h_{13_0}(s^-) \cdot \exp \left(\beta'_{13} \mathbf{X} \right) + h_{23}(t-s) \quad (5)$$

Here, $h_{13_0}(t)$ is some baseline hazard associated with the transition from State-1 directly to State-3. However, for those who experience a stroke at time s , the corresponding part of the baseline hazard is only $h_{13_0}(s^-)$ because after the stroke they experience a change in their overall hazard owing to the effect of stroke, measured as a function of time spent since

stroke. This incremental hazard is given by $h_{23}(t-s)$ and it can be modeled in two ways (depending on how the covariates affect this hazard function), the discussion of which is postponed to section 4.4. In the next section, we show how we can use the three parameter GG distribution to identify $h_{130}(s^-)$ and $h_{23}(t-s)$ and how the properties of this distribution can be used to do the calculations related to RLL due to each stroke type using the concept of RT. More examples of multistate models with additive hazards extensions can be found in Klein et al.¹⁴ and in Shu et al.¹⁵ Variance estimation for some of these models can be found in Klein et al.¹⁶ Large-sample properties of three different additive hazards extension models can be found in Shu.¹⁷ Likewise, examples with a nonparametric semi-Markov model and a semi-Markov model incorporating covariate information are illustrated in Voelkel et al.¹⁸ and Andersen et al.,¹⁹ respectively. Estimation and prediction in a semi-Markov model in a Cox regression framework with baseline hazards and covariates depending on recurrence times has been illustrated in Dabrowska.²⁰

4 Using the GG distribution for calculating residual life lost due to stroke

4.1 The generalized gamma distribution as a family of distributions—The GG distribution²¹ is a three-parameter distribution family with location parameter μ , scale parameter σ , and shape parameter λ that generalizes the two-parameter gamma distribution. The general notation for specifying the distribution is $GG(\mu, \sigma, \lambda)$. Its density function is

$$f_{GG}(t) = \frac{|\lambda|}{\sigma t \Gamma(\lambda^{-2})} \left[\lambda^{-2} \{ \exp(-\mu)t \}^{\lambda/\sigma} \right]^{\lambda-2} \exp \left[-\lambda^{-2} \{ \exp(-\mu)t \}^{\lambda/\sigma} \right] \quad (6)$$

where $\sigma > 0$, $\mu \in (-\infty, \infty)$, $\lambda \in (-\infty, \infty)$, and $\Gamma(x) = \int_0^\infty m^{x-1} e^{-m} dm$ is the gamma function of x .

A complete taxonomy of the various hazard functions for the GG family has been explained in Cox et al.²² and only the relevant aspects are discussed here. Briefly, the GG family allows the flexibility of modeling different shapes of hazard functions such as increasing from 0 to ∞ , increasing from a constant to ∞ , decreasing from ∞ to 0, decreasing from ∞ to a constant, arc-shaped hazards, and bathtub-shaped hazards. Also many popular parametric distributions are special case members of this family. Thus, $\lambda = \sigma$ gives a two-parameter gamma distribution; $\mu = 0$; $\sigma = 1$ gives a standard gamma distribution for fixed λ ; $\lambda = 1$ gives a Weibull distribution; $\lambda = \sigma = 1$ gives an exponential distribution; the limiting case $\lambda = 0$ gives a lognormal distribution; $\lambda = -1$ gives an inverse Weibull distribution; $\lambda = -\sigma$ gives an inverse gamma distribution; $\lambda = 1/\sigma$ gives an ammag distribution; $\lambda = -1/\sigma$ gives an inverse ammag distribution; and a lognormal distribution with $\sigma' = 1.82\sigma$ approximates the loglogistic distribution.²² Maximum likelihood estimation using standard software can be used to obtain estimates of the GG distribution parameters with parsimonious reductions resulting in fitting of well-known distributions.

4.2 Quantiles of the GG distribution—As the main research interest is to quantify residual longevity, it is informative to study the quantiles of a $GG(\mu, \sigma, \lambda)$ distribution. Cox et al.²² show that

$$\begin{aligned} \log \{t_{GG(\mu, \sigma, \lambda)}(p)\} &= \mu + \sigma \cdot \log \{t_{GG(0, 1, \lambda)}(p)\} \\ &= \mu + \sigma \cdot g_{\lambda}(p) \end{aligned} \quad (7)$$

Here, $g_{\lambda}(p)$ is the logarithm of the 100 p th quantile ($0 < p < 1$) from the $GG(0, 1, \lambda)$ distribution. The location parameter μ acts as a time-multiplier and governs the values of the median for fixed values of σ and λ , resulting in an accelerated failure time (AFT) model. The scale parameter σ determines the interquartile ratio for fixed values of λ and independently of μ . The shape parameter λ determines the $GG(0, 1, \lambda)$ distribution. Together, σ and λ describe the type of hazard function for the $GG(\mu, \sigma, \lambda)$ distribution. Operational details of a standard AFT model can be found in Wei.²³

4.3 Relative times—The time by which 100 p % of the population experience an event can lead to a statistic called “relative times $RT(p)$ ”, which can be used to compare survival profiles of patients in different groups (e.g. those that had a stroke versus those that did not). Thus

$$RT(p) = \frac{t_1(p)}{t_0(p)} = \frac{S_1^{-1}(1-p)}{S_0^{-1}(1-p)} \quad (8)$$

where $S_i^{-1}(1-p)$ is the inverse survival function for group i ($i = 0, 1$).

The interpretation of $RT(p)$ is that the time required for 100 p % of individuals in the stroke group to experience death is $RT(p)$ times the time required for 100 p % of individuals in the no stroke group to experience death. Thus if $(\mu_0, \sigma_0, \lambda_0)$ and $(\mu_1, \sigma_1, \lambda_1)$ denote two different sets of GG parameter values, then

$$RT(p) = \exp\{(\mu_1 - \mu_0) + \sigma_1 \cdot g_{\lambda_1}(p) - \sigma_0 \cdot g_{\lambda_0}(p)\} \quad (9)$$

The manner in which covariates affect $RT(p)$ can be summarized as:

- (1) If $\lambda_1 = \lambda_0$ and $\sigma_1 = \sigma_0$ then we have a conventional AFT model resulting in nonproportional hazards, but proportional RT; that is, covariates affect μ only.
- (2) If only $\lambda_1 = \lambda_0$ then we have a model that results in nonproportional hazards and nonproportional $RT(p)$; i.e. covariates affect both μ and σ .
- (3) Full generalization is obtained by having covariates affect all three parameters.
- (4) Reduced parsimonious models result in the fitting of family members of the GG distribution.

4.4 Step-wise approach to identify parameters for the motivating example—

With reference to the model in (3), following step-wise approach can be utilized to answer the research question.

- (1) For pre-stroke times fit the Cox PH model given in (4) with right-censoring also occurring at the time of stroke (when applicable). Obtain estimates of β_{13} for all risk factors other than stroke. The Cox PH framework also allows estimation of risk due to AF, which is a time-dependent covariate.
- (2) Using maximum likelihood estimation method, parametrically estimate the parameters for the most parsimonious subfamily of GG distribution for these pre-stroke times (including subjects with no strokes prior to death) in the absence of any covariates.
- (3) Then generate baseline survival curve for pre-stroke times using the general relation

$$S_0(t) = \exp\left\{-\int_0^t h_0(t) dt\right\} \text{ using either step 1 or step 2.}$$

- (4) Generate a survival profile for pre-stroke times using the average values of all other covariates \mathbf{X} , using the (general) relation $S_{0,\bar{\mathbf{x}}}(t) = \{S_0(t)\}^{\beta_{13}'\bar{\mathbf{x}}}$
- (5) For any time s (of interest) for occurrence of a stroke, calculate $h_{130}(s^-)$. Adjust this for the effect of all other covariates \mathbf{X} .
- (6) The hazard $h_{23}(t-s)$ for those that had strokes as a function of “time since stroke” can be obtained in two ways.
 - (a) The first way is to have all the other covariates affect μ in the AFT model of (7) and then estimate parameters of the most parsimonious subfamily of the GG distribution. Software (e.g. the SAS procedure LIFEREG) can be used for this purpose. Then generate survival curves for those that had strokes as a function of $t-s$, and augment them to the survival curve for those that had not had a stroke at time of stroke $= s$. These survival curves can be generated for either particular values of covariates (such as African American males 65 years, say) or for an average profile of patients in the cohort as done in step 4 of this list.
 - (b) The second way is to consider $h_{23}(t-s) = h_{230}(t-s) \cdot \exp(\beta_{13}'\mathbf{X} + \mathbf{X})$ such that only the baseline hazard $h_{230}(t-s)$ as a function of “time since stroke” is due to the GG distribution and \mathbf{X} is the increment vector of regression coefficients β_{13} corresponding to the other covariates that indicates whether or not the effect of these risk factors on mortality change due to the occurrence of a stroke. As $\beta_{13}'\mathbf{X}$ now occurs on both sides of the + sign in (5), software such as the SAS procedure NLMIXED can be used to maximize the overall likelihood (see Cox et al.²² for sample code using NLMIXED) resulting from (5). Then, similar to step 6a above, the desired survival profiles can be generated. However, it should be noted that the PH assumption needs to be validated here.

- (7) Residual longevity following a stroke can be readily calculated in terms of the 100 p th percentile (for median, $p = 0.5$) survival time after the occurrence of a stroke at time s from step 6 (a or b) above. With reference to Figure 4, this corresponds to calculating

$$t_{stroke}(p) - s = S_{stroke}^{-1} \{ (1 - p) \cdot S_{no\ stroke}(s) \} - s.$$

- (8) Residual longevity absent any strokes for any time s can be calculated by evaluating the 100 p th percentile of survival time from the conditional survival function $S_0(t|s)$. With reference to Figure 4, this corresponds to calculating

$$t_{no\ stroke}(p) - s = S_{no\ stroke}^{-1} \{ (1 - p) \cdot S_{no\ stroke}(s) \} - s.$$

- (9) Then $RT(p) = \frac{t_{no\ stroke}(p) - s}{t_{stroke}(p) - s}$ can be used to calculate the relative 100 p th percentile of time required for absent a stroke to experience death as compared to when a stroke occurs. Alternatively, the 100 p th percentile of RLL due to stroke, a concept more appealing in practice to the clinicians can be calculated using the relation given by

$$RLL(p) = S_{no\ stroke}^{-1} \{ (1 - p) \cdot S_{no\ stroke}(s) \} - S_{stroke}^{-1} \{ (1 - p) \cdot S_{no\ stroke}(s) \}$$

. Specifically, $RLL(0.25)$, $RLL(0.5)$, and $RLL(0.75)$ may prove to be quite informative in explaining the skew in the RLL due to stroke. These calculations are graphed in Figure 4. It should also be noted that though Figure 4 depicts the comparison of median (assuming $p = 0.5$) residual lifetime of those having a stroke (at say time s) to the median residual lifetime of those not having a stroke under the assumption that they will always maintain their “no-stroke” status, it can also be used for alternative calculations. Thus acknowledging that some patients in the “no-stroke” state will later go on to have a stroke, we can also use it to compare the RLL between those that have a stroke (at time s) and those who are stroke-free at time s but may experience a stroke a future time s' . For example, say $p=0.5$, $s=12$, and $s' = 36$, then the median RLL can now be calculated as

$$S_{stroke|s=36}^{-1} \{ (0.5) \cdot S_{no\ stroke}(12) \} - S_{stroke|s=12}^{-1} \{ (0.5) \cdot S_{no\ stroke}(12) \}.$$

- (10) Evaluation of the model fit can be performed using the graphical diagnostic checking procedure mentioned in section 3.1.

In the above 10-step procedure, we have explained how $RLL(p)$ and $RT(p)$ can be estimated with all covariate values adjusted at their average values (representing average survival profiles of patients in the stroke and no-stroke groups). We can also generate survival profiles for any particular combination of covariate values (say, white male, age 65 with hypertension and diabetes) and therefore estimate $RLL(p)$ and $RT(p)$ for those combination of covariate values.

5 Results

Descriptive statistics for risk factors related to patient demographics can be found in Wetmore et al.²⁴ Here, we discuss the HR estimates of these risk factors only in relevance to our model of (3). Table 1[A] displays the results representing (4) for no or pre-stroke times using a Cox PH model with AF treated as a time-dependent covariate. Baseline hazard generated from this model suggests a Weibull distribution as a parsimonious reduction of a GG distribution ($\lambda = 1$) for the times prior to stroke events, with a gradually decreasing hazard that slowly stabilizes over time with parameter estimates given by $\hat{\mu} = 4.385$, $\hat{\sigma} = 1.096$ (obtained from a linear fit of logarithm of the cumulative hazard versus logarithm of time). The survival curve shown in Figure 4 was generated using these parameters by adjusting the baseline survival curve for an average risk profile of patients in the cohort.

Table 1[B] displays the results corresponding to the additive part of the hazard outlined in (5) used to model the survival experience following a hemorrhagic stroke ($N = 534$ such events) using the first approach (step 6a in section 4.4), and other risk factors affect only μ in the AFT model of (7) using PROC LIFEREG). From this table, it can be seen that a GG distribution with parameters estimates given by $\hat{\mu} = -2.179$ (99% CI: $[-4.761, 0.403]$), $\hat{\sigma} = 1.258$ (99% CI: $[1.065, 1.485]$), and $\hat{\lambda} = -2.281$ (99% CI: $[-2.853, -1.710]$) provides the best fit (AIC 1681.89). Adjusting $\hat{\mu}$ to the average risk profile of patients allows generation of Figure 4 for hemorrhagic strokes (this adjusts the location estimate to $\hat{\mu} = -1.205$). We also fitted the PH model mentioned in step 6(b) in section 4.4 using PROC NLMIXED. On doing so, we found that only two risk factors belonging to the β vector (age and smoking status were statistically significant at the 0.01 level). However, the Cox PH assumption did not hold true for these two variables and hence this model was discarded. (Note that this model requires estimation of 44 regression coefficients in addition to three parameters each for the nonstroke and hemorrhagic stroke groups which may be unreliable due to there being only $N=409$ death events in hemorrhagic stroke group).

Table 1[C] displays the results corresponding to the additive part of the hazard outlined in (5) used to model the survival experience following ischemic strokes ($N=2381$) using the first approach (other risk factors affect only μ in the AFT model of (7) using PROC LIFEREG). A lognormal distribution with parameters estimates given by $\hat{\mu} = 3.041$ (99% CI: $[1.665, 4.416]$) and $\hat{\sigma} = 1.883$ (99% CI: $[1.793, 1.979]$) provides the best fit (AIC 7411.36 for lognormal as compared to AIC 7412.29 for GG). Adjusting $\hat{\mu}$ to the average risk profile of patients allows Figure 4 ischemic stroke results (this adjusts the location estimate to $\hat{\mu} = 2.333$). Here again, we fitted the PH model mentioned in response numbered 6b in section 4.4 using PROC NLMIXED and found that only two risk factors belonging to the β vector (ambulation and employment status were statistically significant at the 0.01 level). However, as the Cox PH assumption did not hold true for these two variables as evidenced by log—log survival plots, this model was discarded. It should be noted that though in our application the model using NLMIXED was not very useful, it can turn out to be the model of choice in other applications.

Tables 2 and 3 show calculations related to $RT(p)$ and $RLL(p)$ for different values of p when the two stroke types are separately compared to no stroke. The 99% CIs are generated using

the delta method. All comparisons are adjusted for the average demographic and risk profile of patients in the cohort. From this output, we see that for patients who experienced a hemorrhagic stroke right at the entry point into the cohort, $RLL(p)$ ranges from 2.291 months (5th percentile) to 53.453 months (60th percentile) with an estimated value of 39.903 months for median residual life lost. Likewise, for patients who experience a hemorrhagic stroke at 12 months into the cohort, the $RLL(p)$ ranges from 2.919 months (5th percentile) to 42.461 months for median residual life lost. As approximately 38% of the no stroke group do not die in the 60 month observation window, $RLL(p)$ calculations are restricted to the 60th percentile for $s=0$, 50th percentile for $s=12$, and 40th percentile for $s=24$. The increase in $RLL(p)$ values for different values of $s=0, 12, 24$ is due the fact that though the overall survival experience following a stroke remains the same, the Weibull hazard of the reference (no stroke) group decreases gradually over time. Thus there is greater loss in residual life for patients who experience a stroke late into the observation window (time on dialysis) as opposed to early. Analogously, for patients who experienced an ischemic stroke right at the entry point into the cohort, $RLL(p)$ ranges from 1.946 months (5th percentile) to 38.975 months (60th percentile) with an estimated value of 30.715 months for median residual life lost. Likewise, for patients who experience an ischemic stroke at 12 months into the cohort, the $RLL(p)$ ranges from 2.575 months (5th percentile) to 33.272 months for median residual life lost. From these numbers (and from Figure 4), we can conclude that a hemorrhagic stroke confers a higher disadvantage in terms of residual life lost than an ischemic stroke. For example, for $s=12$, median RLL due to a hemorrhagic stroke exceeds median RLL due to an ischemic stroke by approximately nine months.

The $RLL(p)$ estimates are complemented by the $RT(p)$ estimates in that they provide an interpretation of the multiplicative factor that the residual longevity would have increased for the 100th percentile of patients who experience a stroke at time s had they continued on dialysis absent that stroke. Thus, for patients who experience a hemorrhagic stroke at $s=12$, $RT(p)$ increases from 25.204 for $p=0.05$ to 54.025 for $p=0.25$ before gradually dropping down to 39.271 at $p=0.5$. Analogously for $s=12$, $RT(p)$ for those that had ischemic stroke increases from 6.538 at $p=0.05$ to 6.814 at $p=0.1$ before decreasing gradually to 4.231 at $p=0.5$ suggesting that hemorrhagic stroke is a more drastic event than ischemic stroke. These estimates increase for patients who experience a stroke later in the observation window. Keeping p fixed, the confidence intervals are wider for both $RT(p)$ and $RLL(p)$ for high values of s (smaller risk set) as compared to small values of s (larger risk set).

6 Discussion

By using a semi-Markov model with an additive hazard extension, we were able to estimate relative residual longevity due to the two types of strokes. Estimates of $RT(p)$ and $RLL(p)$ obtained from this framework suggest that both stroke types are major, watershed events that radically change the survival profile of those who experience them. In fact, these stroke events affected all three parameters of the GG distribution model, which was then used to estimate the survival trajectory of patients who experienced strokes, which is essential to calculating residual longevity. This would not have been possible in a standard Cox PH model with stroke treated as a time-dependent covariate because in that case stroke would affect only the magnitude of the hazard in a proportional manner without affecting the shape.

Even if the general Cox model framework can incorporate an interaction term of stroke with some functional form of time to account for the nonproportionality of hazard, it would be very difficult to reliably find this functional form. On the other hand, proportionality of hazard functions in the GG family is only limited to the special case of conventional Weibull regression models where the covariate of interest affects only the location parameter. Instead when the covariate can affect two or all three parameters of the GG family, our modeling framework allows the flexibility of using commonly available standard statistical software to model the nonproportionality of hazards and/or the nonproportionality of relative times in order to generate the survival profiles of interest (for any combination of covariate values).

Though we have focused on the nonproportional $RT(p)$ and $RLL(p)$ in this paper, it is also possible to estimate the nonproportional hazards due to stroke, and this has been done in Wetmore et al.²⁴ as clinicians are more accustomed to the hazard ratio interpretation of risk factors. It is the motivating example, however, that has driven our current approach in the sense that it makes it possible for researchers to inform doctors and patients who experienced a stroke while on dialysis, the estimated median (or any other 100 p th percentile) residual life lost due to the stroke. This is a paradigm shift from the usual approach in which only the mortality hazard ratio is reported.

Another interesting feature of our model is revealed by observing the exponentiated location estimates in Table 1[B] and [C]. The drastic nature of a hemorrhagic stroke is revealed by the fact that none of the other risk factors play a significant role in altering the risk of mortality in its presence. This is understandable given that median time to death after hemorrhagic stroke is only 0.84 months. On the other hand, in case of ischemic stroke, advancing age and white race still continue to significantly alter the mortality risk. In the context of the AFT extension of (5), every one year increase in age decreases the median time to death by a factor of 0.96, and race (black versus white) increases the median time to death by a factor of 1.75. A similar effect of race and age is observed in the no or pre-stroke time, and this can be verified using the relation between parameter estimates from a Cox PH model and an AFT Weibull model for no/pre-stroke survival given by $\hat{\beta}_{PH} = \hat{\beta}_{AFT-Weibull} / \sigma$. This suggests that after an ischemic stroke, race and age continue to retain their effect on mortality whereas the other risk factors are shadowed by the drastic nature of an ischemic stroke.

One limitation of our study is that the observation window is limited to a maximum possible follow-up time of 69 months resulting in a Weibull distribution with a gradually decreasing hazard for the nonstroke subjects. A larger follow-up window would result in the hazard for nonstroke subjects increasing gradually over time owing to the fact that this cohort represents a chronic dialysis population. In that case, a Weibull distribution perhaps, would not be the best choice and the new choice would affect the $RT(p)$ and $RLL(p)$ calculations accordingly. It is also possible that in that case we would have to consider the complicated extended semi-Markov model with time at which stroke occurs also playing a key role in the subsequent calculations. Another minor limitation of our study is that it uses claims-based data (all dates have been calculated using clinically approved algorithms rather than the exact event dates).

Our modeling framework also provides the flexibility to extend our approach to analysis for chronic dialysis cohort observed after 2005 and treated with exposure to new treatments (say, warfarin) with current cohort survivors left truncated at time of entry into the “new treatment-exposure” cohort. In this context, it may be interesting to investigate the effect of new treatment by studying $RT(p)$, $RLL(p)$, and nonproportional hazards across different time periods. Perhaps, the more involved model using PROC NLMIXED (discussed in section 4.4) could provide more useful in this regard. We envisage that our ensemble modeling framework will provide real life interpretations of research questions where the primary focus lies in drawing meaningful conclusions related to the survival and hazard trajectory of cohort participants using multi-parameter distributions. To this end, our model provides a reasonable solution.

Acknowledgments

Funding

This work was supported by National Institute of Diabetes and Digestive and Kidney Diseases (R01 DK080111); the National Kidney Foundation Young Investigator Award and the Sandra A. Daugherty Foundation Grant (K23 DK085378-01).

References

1. Cox DR. Regression models and life tables (with discussion). *J Roy Stat Soc B.* 1972; 34:187–220.
2. Klein, JP., Moeschberger, ML. *Survival analysis - Techniques for censored and truncated data.* 2nd ed.. Springer; New York: 2003. p. 304
3. Wetmore JB, Ellerbeck EF, Mahnken JD, et al. Atrial fibrillation and risk of stroke in dialysis patients. *Ann Epidemiol.* 2013; 23:112–118. [PubMed: 23332588]
4. United States Renal Data System. *USRDS 2006 Annual Data Report: Atlas of end-stage renal disease in the United States.* National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Disease; 2006.
5. Wetmore JB, Ellerbeck EF, Mahnken JD, et al. Stroke and the ‘stroke belt’ in dialysis: contribution of patient characteristics to ischemic stroke rate and its geographic variation. *J Am Soc Nephrol.* 2013; 24:2053–2061. [PubMed: 23990675]
6. Wetmore JB, Phadnis MA, Mahnken JD, et al. Race, ethnicity, and state-by-state geographic variation in hemorrhagic stroke in dialysis patients. *Clin J Am Soc Nephrol.* 2014; 9:756–763. [PubMed: 24458073]
7. Storer BE, Gooley TA, Jones MP. Adjusted estimates for time-to-event endpoints. *Lifetime Data Anal.* 2008; 14:484–495. [PubMed: 18791867]
8. Hougaard, P. *Analysis of multivariate survival data.* Springer; New York: 2001. p. 87
9. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika.* 1994; 81:515–526.
10. Keele L. Proportionality difficult: Testing for nonproportional hazards in Cox models. *Political Anal.* 2010; 18:189–205.
11. Therneau TM, Grambsch PM, Fleming TR. Martingale based residuals for survival models. *Biometrika.* 1990; 77:147–160.
12. Therneau, TM., Grambsch, PM. *Modeling survival data: Extending the Cox model.* Springer Verlag; New York: 2000.
13. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. *Stat Med.* 2007; 26:2389–2430. [PubMed: 17031868]
14. Klein JP, Shu YY. Multi-state models for bone marrow transplantation studies. *Stat Meth Med Res.* 2002; 11:117–139.

15. Shu YY, Klein JP. Additive hazards Markov regression models illustrated with bone marrow transplant data. *Biometrika*. 2005; 92:283–301.
16. Klein, JP., Qian, C. Modeling multistate survival illustrated in bone marrow transplantation; *Proceedings of the Biometrics Section of the American Statistical Society*; 1996; p. 93-102.
17. Shu, YY. Multistate survival models: Theory and applications. Division of Biostatistics, The Medical College of Wisconsin; USA: 2001. PhD Dissertation
18. Voelkel JG, Crowley J. Nonparametric inference for a class of semi-Markov processes with censored observations. *Ann Stat*. 1984; 12:142–160.
19. Andersen PKA, Esbjerg S, Sorensen TIA. Multistate models for bleeding episodes and mortality in liver cirrhosis. *Stat Med*. 2000; 19:587–599. [PubMed: 10694738]
20. Dabrowska DM, Sun G, Horowitz MM. Cox regression in a Markov renewal model: An application of bone marrow transplant data. *J Am Stat Assoc*. 1994; 89:867–877.
21. Stacy EW, Mihram GA. Parameter estimation for a generalized gamma distribution. *Technometrics*. 1965; 7:349–358.
22. Cox C, Chu H, Schneider MF, et al. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Stat Med*. 2007; 26:4352–4374. [PubMed: 17342754]
23. Wei LJ. The accelerated failure time model: A useful alternative to the Cox regression model in survival analysis. *Stat Med*. 1992; 11:1871–1879. [PubMed: 1480879]
24. Wetmore JB, Phadnis MA, Ellerbeck EF, et al. Stroke as a watershed event: Impact of first acute stroke on mortality in dialysis patients. *Clin J Am Soc Nephrol*. 2014; 10:80–89. [PubMed: 25318759]

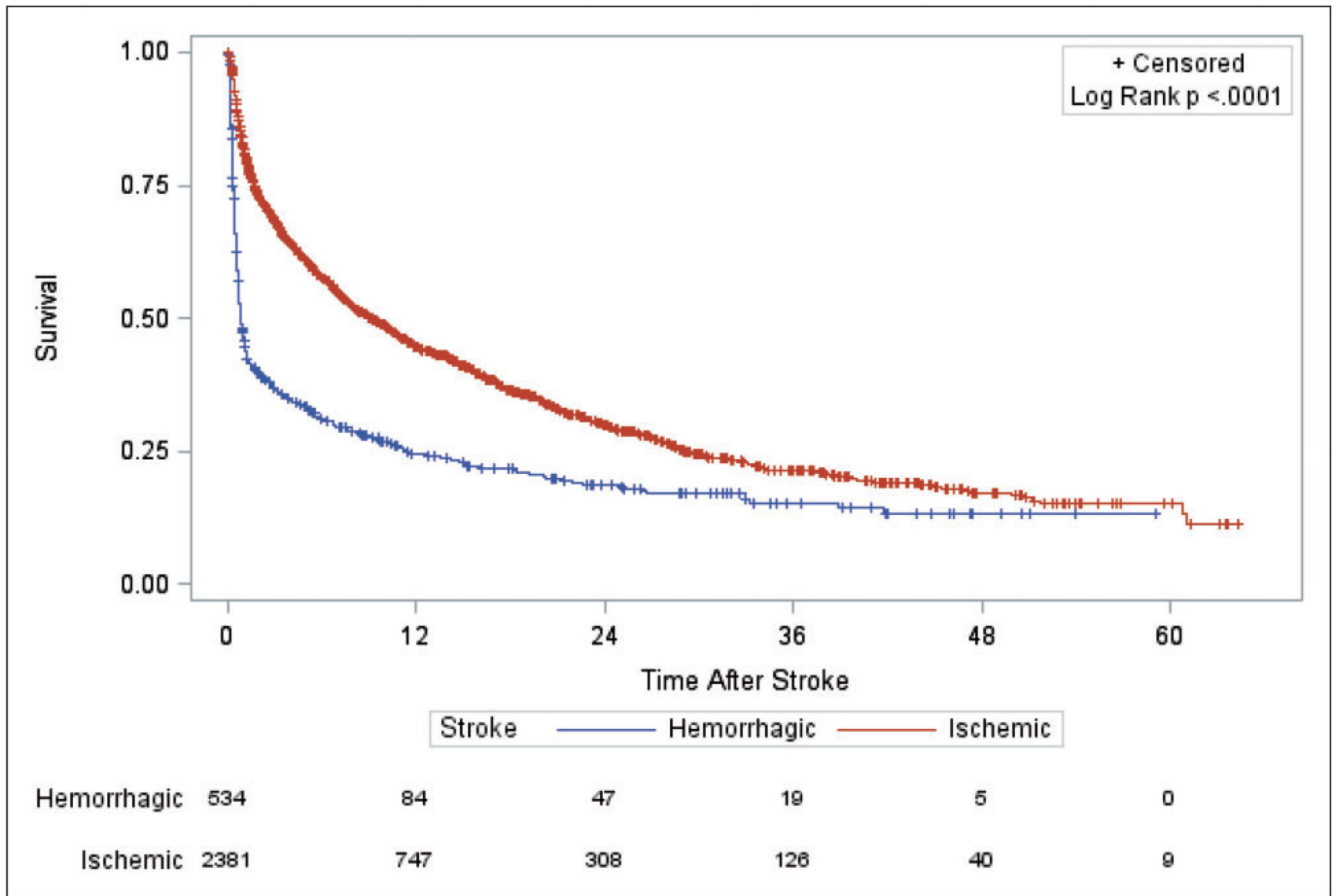


Figure 1. Kaplan–Meier curve illustrating survival after an ischemic or hemorrhagic stroke.

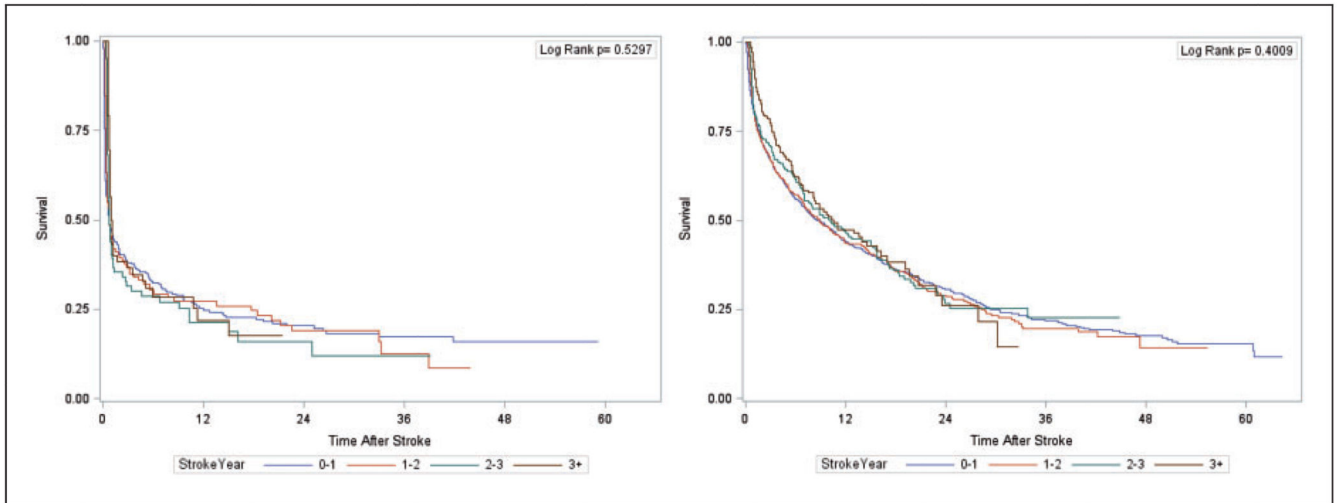


Figure 2. Kaplan–Meier curve stratified by year of occurrence for survival after first stroke: (a) hemorrhagic, (b) ischemic.

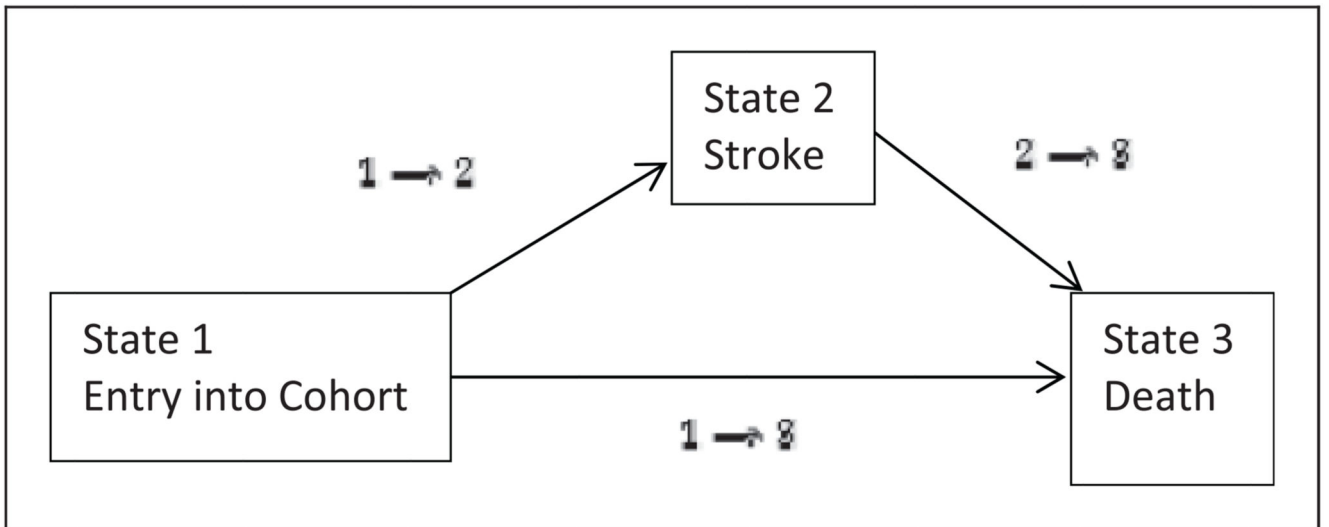


Figure 3. Single-state transition model for mortality with stroke as an intermediate state.

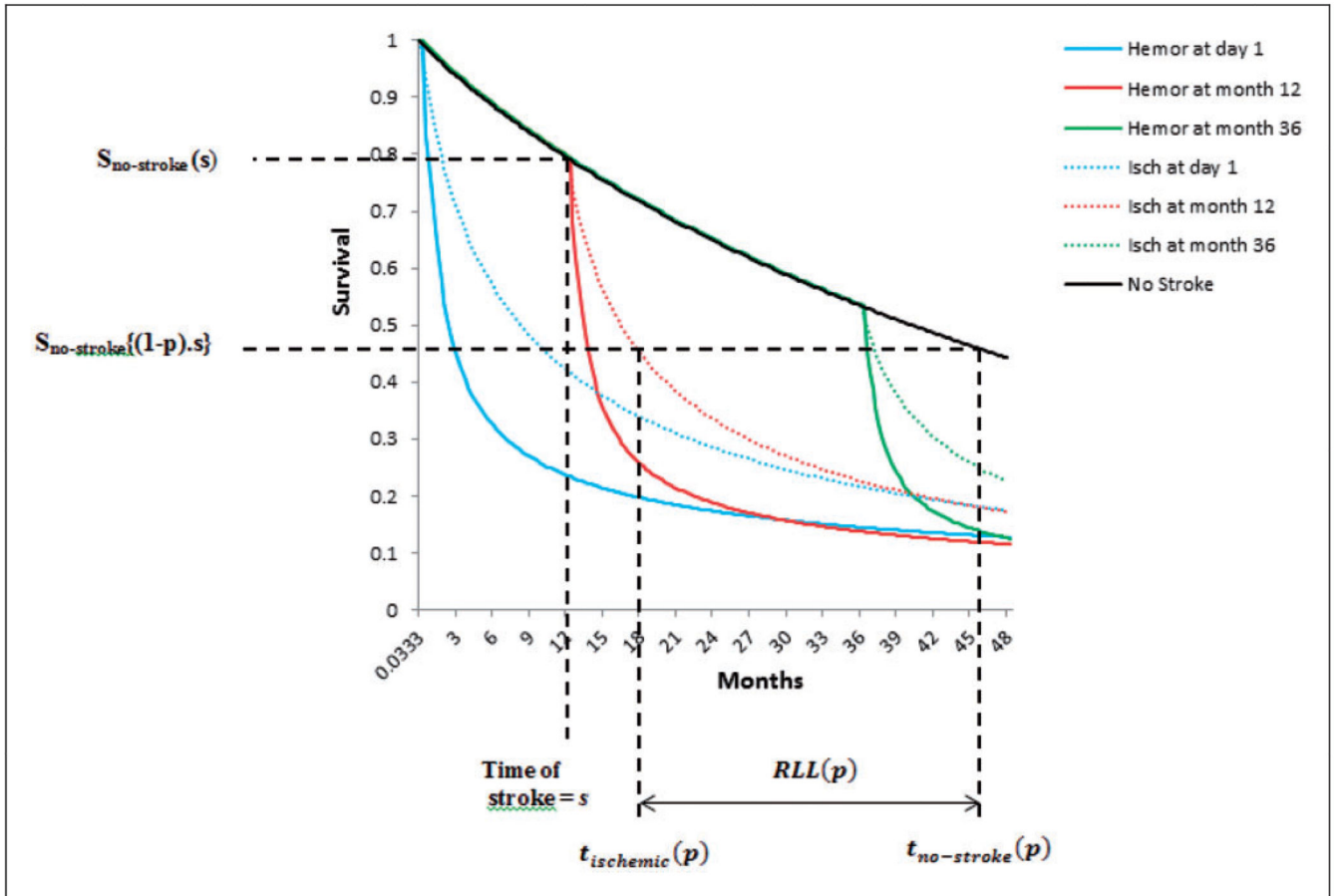


Figure 4. Calculating residual longevity following stroke occurrence at time = s.

Table 1

Modeling results using equation (4) for [A] no or pre-stroke times and equation (5) for times following [B] hemorrhagic and [C] ischemic strokes.

| Risk factors (covariates) | [A] Nonstrokes (Cox PH) | | | [B] Hemorrhagic strokes (AFT extension) | | | [C] Ischemic strokes (AFT extension) | | |
|---------------------------------------|-------------------------|--------------|---------|---|----------------|---------|---|---------------|---------|
| | HR Estimate | 99% CI | p-Value | Exponentiated location estimate $\exp(\hat{\mu})$ | 99% CI | p-Value | Exponentiated location estimate $\exp(\hat{\mu})$ | 99% CI | p-Value |
| Intercept | - | - | - | 0.113 | 0.009, 1.497 | 0.0297 | 20.918 | 5.288, 82.756 | <0.0001 |
| Race: Black vs White | 0.779 | 0.750, 0.809 | <0.0001 | 1.292 | 0.862, 1.937 | 0.1035 | 1.750 | 1.355, 2.589 | <0.0001 |
| Race: Hispanic vs White | 0.673 | 0.641, 0.707 | <0.0001 | 1.469 | 0.902, 2.390 | 0.0421 | 1.802 | 1.306, 2.485 | <0.0001 |
| Race: Others vs White | 0.585 | 0.541, 0.633 | <0.0001 | 1.328 | 0.676, 2.609 | 0.2801 | 2.673 | 1.507, 4.743 | <0.0001 |
| Age | 1.030 | 1.028, 1.031 | <0.0001 | 1.003 | 0.992, 1.014 | 0.4794 | 0.959 | 0.948, 0.968 | <0.0001 |
| Sex: F vs M | 0.944 | 0.912, 0.977 | <0.0001 | 0.861 | 0.603, 1.230 | 0.2796 | 1.068 | 0.837, 1.362 | 0.4886 |
| No ambulation: No vs Yes | 0.686 | 0.639, 0.736 | <0.0001 | 0.821 | 0.291, 2.317 | 0.6242 | 1.082 | 0.660, 1.777 | 0.6804 |
| No transfer: No vs Yes | 0.808 | 0.726-0.899 | <0.0001 | 2.752 | 0.309, 24.498 | 0.2329 | 1.410 | 0.657, 3.028 | 0.2461 |
| Dialysis: In-center vs Self | 0.962 | 0.887, 1.044 | 0.2232 | 0.808 | 0.353, 1.847 | 0.5057 | 1.208 | 0.728, 2.004 | 0.3360 |
| Cardiac failure: No vs Yes | 0.887 | 0.840, 0.936 | <0.0001 | 0.791 | 0.439, 1.423 | 0.3031 | 1.485 | 1.026, 2.149 | 0.0059 |
| Previous stroke: No vs Yes | 0.875 | 0.832, 0.920 | <0.0001 | 0.887 | 0.550, 1.432 | 0.5200 | 1.059 | 0.773, 1.449 | 0.6388 |
| Vascular disease: No vs Yes | 0.898 | 0.857, 0.941 | <0.0001 | 0.994 | 0.587, 1.684 | 0.9781 | 0.946 | 0.688, 1.300 | 0.6504 |
| Diabetes: No vs Yes | 0.979 | 0.936, 1.024 | 0.2248 | 0.744 | 0.478, 1.158 | 0.0852 | 1.160 | 0.845, 1.592 | 0.2289 |
| CAD: No vs Yes | 0.959 | 0.921, 0.999 | 0.0078 | 1.234 | 0.809, 1.882 | 0.1998 | 1.201 | 0.923, 1.562 | 0.0734 |
| Hypertension: No vs Yes | 1.323 | 1.266, 1.382 | <0.0001 | 0.952 | 0.594, 1.526 | 0.7869 | 0.814 | 0.586, 1.131 | 0.1064 |
| BMI: Normal vs (<25) | 0.764 | 0.725, 0.806 | <0.0001 | 1.082 | 0.666, 1.756 | 0.6761 | 1.024 | 0.687, 1.523 | 0.8796 |
| BMI: (25-30) vs (<25) | 0.644 | 0.609, 0.681 | <0.0001 | 1.031 | 0.593, 1.791 | 0.8886 | 1.152 | 0.770, 1.723 | 0.3656 |
| BMI new: 30+ vs (<25) | 0.592 | 0.560, 0.626 | <0.0001 | 1.317 | 0.767, 2.264 | 0.1898 | 1.296 | 0.864, 1.943 | 0.0992 |
| Employed: Yes vs No | 0.588 | 0.491, 0.703 | <0.0001 | 0.464 | 0.155, 1.383 | 0.0700 | 0.480 | 0.163, 1.415 | 0.0801 |
| Smoker: No vs Yes | 0.915 | 0.854, 0.980 | 0.0009 | 1.167 | 0.651, 2.091 | 0.4962 | 1.145 | 0.689, 1.902 | 0.4908 |
| Substance abuser: No vs Yes | 0.822 | 0.742, 0.910 | <0.0001 | 1.068 | 0.536, 2.129 | 0.8061 | 1.207 | 0.510, 2.855 | 0.5739 |
| Liulindex: ^d (3-6) vs (<3) | 0.812 | 0.744, 0.887 | <0.0001 | 1.619 | 0.678, 3.862 | 0.1538 | 0.917 | 0.507, 1.659 | 0.7063 |
| Liulindex: ^d (> 6) vs (<3) | 0.954 | 0.897, 1.014 | 0.0464 | 1.451 | 0.768, 2.741 | 0.1316 | 1.013 | 0.669, 1.534 | 0.9358 |
| Atrial fibrillation: Yes vs No | 1.181 | 1.098, 1.269 | <0.0001 | Scale, $\hat{\sigma} = 1.258$ | 1.065, 1.485 | | Scale, $\hat{\sigma} = 1.883$ | 1.793, 1.978 | |
| Atrial fibrillation*Time | 1.006 | 1.003, 1.009 | <0.0001 | Shape, $\hat{\lambda} = -2.281$ | -2.853, -1.710 | | | | |

$L_{i,j}$ index represents modified Liu-Comorbidity Score measured on an ordinal scale of 0 to 25 (but categorized in three groups in our analyses).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2Comparison of hemorrhagic stroke versus no stroke using $RT(p)$ and $RLL_{months}(p)$.

| p | $s = 0$ | | $s = 12$ months | | $s = 24$ months | |
|------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| | RT [99%CI] | RLL [99%CI] | RT [99%CI] | RLL [99%CI] | RT [99%CI] | RLL [99%CI] |
| 0.05 | 19.993 [16.633–23.354] | 2.291 [2.198–2.198] | 25.204 [20.311–30.096] | 2.919 [2.602–3.236] | 26.290 [20.100–32.480] | 3.050 [2.514–3.586] |
| 0.10 | 32.432 [26.739–38.124] | 5.115 [4.960–5.271] | 38.588 [31.531–45.645] | 6.117 [5.745–6.489] | 40.447 [32.515–48.380] | 6.420 [5.812–7.027] |
| 0.15 | 40.711 [33.082–48.340] | 8.251 [8.039–8.464] | 46.523 [37.637–55.409] | 9.459 [9.027–9.891] | 49.208 [39.489–58.926] | 10.017 [9.330–10.705] |
| 0.20 | 45.885 [36.720–55.051] | 11.684 [11.413–11.955] | 51.842 [41.372–62.312] | 13.235 [12.730–13.740] | 53.791 [42.708–64.874] | 13.743 [12.967–14.518] |
| 0.25 | 48.550 [38.259–58.842] | 15.425 [15.090–15.760] | 54.025 [42.487–65.563] | 17.201 [16.613–17.789] | 56.056 [43.927–68.185] | 17.860 [16.981–18.738] |
| 0.30 | 49.127 [38.112–60.143] | 19.496 [19.087–19.904] | 53.743 [41.625–65.861] | 21.365 [20.683–22.048] | 56.023 [43.274–68.771] | 22.289 [21.294–23.283] |
| 0.35 | 47.959 [36.593–59.324] | 23.930 [23.433–24.427] | 51.989 [39.336–64.642] | 25.984 [24.421–27.547] | 53.819 [40.921–66.717] | 26.917 [25.793–28.040] |
| 0.40 | 45.350 [33.971–56.729] | 28.772 [28.168–29.376] | 48.784 [36.500–61.067] | 31.000 [30.073–31.926] | 50.461 [37.688–63.233] | 32.088 [30.812–33.364] |
| 0.45 | 41.591 [30.490–52.692] | 34.075 [33.338–34.811] | 44.444 [32.549–56.334] | 36.469 [35.386–37.553] | – | – |
| 0.50 | 36.966 [26.397–47.536] | 39.903 [39.000–40.807] | 39.271 [28.017–50.525] | 42.461 [41.188–43.733] | – | – |
| 0.55 | 31.760 [21.942–41.577] | 46.335 [45.214–47.457] | – | – | – | – |
| 0.60 | 26.253 [17.385–35.122] | 53.453 [52.029–54.878] | – | – | – | – |

Table 3

Comparison of ischemic stroke versus no stroke using $RT(p)$ and $RLL_{months}(p)$.

| p | $s = 0$ | | $s = 12$ months | | $s = 24$ months | |
|------|------------------------|---------------------------|------------------------|---------------------------|------------------------|---------------------------|
| | RT [99%CI] | RLL [99%CI] | RT [99%CI] | RLL [99%CI] | RT [99%CI] | RLL [99%CI] |
| 0.05 | 5.186 [4.260–6.112] | 1.946 [1.825–2.067] | 6.538 [5.209–7.866] | 2.575 [2.248–2.901] | 6.819 [5.162–8.476] | 2.706 [2.164–3.248] |
| 0.10 | 5.727 [4.829–6.625] | 4.356 [4.148–4.565] | 6.814 [5.689–7.938] | 5.358 [4.961–5.755] | 7.142 [5.860–8.425] | 5.661 [5.034–6.284] |
| 0.15 | 5.785 [4.944–6.626] | 6.997 [6.701–7.293] | 6.611 [5.619–7.602] | 8.205 [7.726–8.683] | 6.992 [5.886–8.098] | 8.763 [8.045–9.480] |
| 0.20 | 5.660 [4.876–6.444] | 9.834 [9.442–10.227] | 6.394 [5.488–7.300] | 11.385 [10.805–11.965] | 6.635 [5.657–7.613] | 11.892 [11.067–12.718] |
| 0.25 | 5.447 [4.715–6.179] | 12.858 [12.353–13.363] | 6.061 [5.232–6.891] | 14.634 [13.935–15.333] | 6.289 [5.401–7.178] | 15.293 [14.337–16.249] |
| 0.30 | 5.188 [4.502–5.875] | 16.065 [15.424–16.706] | 5.676 [4.913–6.439] | 17.935 [17.093–18.777] | 5.916 [5.100–6.732] | 18.858 [17.747–19.968] |
| 0.35 | 4.903 [4.257–5.549] | 19.456 [18.647–20.264] | 5.315 [4.605–6.025] | 21.509 [20.490–22.528] | 5.502 [4.752–6.253] | 22.442 [21.150–23.734] |
| 0.40 | 4.604 [3.994–5.213] | 23.030 [22.011–24.049] | 4.952 [4.289–5.616] | 25.258 [24.020–26.496] | 5.122 [4.423–5.582] | 26.346 [24.828–27.863] |
| 0.45 | 4.296 [3.719–4.872] | 26.786 [25.500–28.072] | 4.590 [3.968–5.213] | 29.181 [27.669–30.693] | – | – |
| 0.50 | 3.982 [3.437–4.528] | 30.715 [29.084–32.345] | 4.231 [3.646–4.816] | 33.272 [31.411–35.132] | – | – |
| 0.55 | 3.667 [3.151–4.182] | 34.793 [32.713–36.874] | – | – | – | – |
| 0.60 | 3.349 [2.862–3.835] | 38.975 [36.293–41.656] | – | – | – | – |