Statistical Methods for Recurrent Event Analysis in Cohort Studies of CKD

Wei Yang,*† Christopher Jepson,† Dawei Xie,‡ Jason A. Roy,*† Haochang Shou,*‡ Jesse Yenchih Hsu,*‡
Amanda Hyre Anderson,*† J. Richard Landis,*† Jiang He,* Harold I. Feldman,*† and on behalf of the Chronic Renal Insufficiency Cohort (CRIC) Study Investigators

Abstract
Cardiovascular events, such as hospitalizations because of congestive heart failure, often occur repeatedly in patients with CKD. Many studies focus on analyses of the first occurrence of these events, and discard subsequent information. In this article, we review a number of statistical methods for analyzing ordered recurrent events of the same type, including Poisson regression and several commonly used survival models that are extensions of Cox proportional hazards regression. We illustrate the models by analyzing data from the Chronic Renal Insufficiency Cohort Study to identify risk factors for congestive heart failure hospitalizations in patients with CKD. We show that recurrent event analyses provide additional insights about the data compared with a standard survival analysis of time to the first event.

Motivating Example: Risk Factors for Recurrent CHF
The CRIC Study is a multicenter cohort study of individuals with CKD. The study enrolled 3939 individuals from seven clinical centers in the United States between 2003 and 2008. Mean age at study entry was 58 years (SD, 11.0), 55% of participants were men, 42% were white, 42% were black, and 13% reported Hispanic ethnicity. Nearly half had diabetes. The mean eGFR at study entry was 45 ml/min per 1.73 m² (SD, 16.8) and median 24-hour urine protein was 0.2 g (interquartile range, 0.07–0.91). Participants were followed up through annual clinic visits, with telephone interviews every 6 months in between. Clinical outcomes collected in the study included ESRD, cardiovascular diseases (including CHF, stroke, myocardial infarction, atrial fibrillation, and peripheral artery disease), and death. The cardiovascular disease events are adjudicated on the basis of reviews of hospitalization records. Details of the study design, participant characteristics, and events adjudication have been previously published (10–12). Prior analyses from the CRIC Study (13) have identified a number of novel risk factors for incident CHF, such as N-terminal pro-B-type natriuretic peptide (NT-proBNP) and high-sensitivity troponin-T (hsTnT). Given the high recurrence of CHF hospitalization in patients with CKD, it is also of interest to explore risk factors for recurrent hospitalizations for CHF and examine whether they differ from those for incident CHF. We utilized several different modeling approaches, reviewed below.

Introduction
In biomedical research, recurrent events refer to events of interest experienced repeatedly by a given individual. These events may all be of the same type, or different types. For example, patients with CKD are often hospitalized repeatedly for different cardiovascular disease events (e.g., congestive heart failure [CHF], myocardial infarction, or stroke). They may also be hospitalized repeatedly because of the same type of cardiovascular disease, such as CHF. Other examples of recurrent events include vascular access complications such as repeated thrombosis and infections in patients on dialysis (1), urinary tract infections, peritonitis, or episodes of transplant rejection.

A goal that is often pursued in clinical research is the identification of risk factors for an incident event, which can be addressed by examining the first occurrence of the event only. However, we may also be interested in understanding the factors associated with the repeated occurrence of events over time. We may hypothesize that the risk factors for the recurrence of events may differ from the risk factors for the first event. To answer these questions, we have to examine subsequent events using a set of analytical tools suited for this purpose.

There is a large body of statistical literature on how to model recurrent events (2–8), but few of these models have been used in nephrology literature. In this article, we focus on statistical models that are specifically designed to analyze recurrence of the same type of event. We review two types of models that are popular in clinical research: Poisson regression and several survival models that are extensions of Cox proportional hazards regression (9). We describe the assumptions underlying these different modeling strategies and how they differ from the analysis of time-to-first-event using Cox proportional hazards regression. The different methods are then illustrated using data from the Chronic Renal Insufficiency Cohort (CRIC) Study to examine risk factors for repeated hospitalizations because of CHF.

Correspondence:
Dr. Wei Yang, Department of Biostatistics, Epidemiology, and Informatics and 
†Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; and 
‡Department of Epidemiology, Tulane University, New Orleans, Louisiana

Description of Recurrent Event Data

A description and visualization of the data are important first steps in analysis. For example, in standard survival analyses of a single event, the Kaplan–Meier curve (14) is often used to examine the distribution of survival times in the study population. Kaplan–Meier curves look only at the first occurrence of the event of interest and are inadequate for describing recurrent events. One can describe recurrent event data by estimating the so-called mean cumulative function (MCF) (15–17), which is the average number of cumulative events experienced by an individual in the study at each point in time since the start of follow-up. As an illustration of this, Figure 1 shows the MCF of CHF hospitalization for individuals in the different race categories from the CRIC Study. For example, the MCF value for non-Hispanic blacks is about 0.4 at year 5, which means that a non-Hispanic black individual experienced, on average, 0.4 CHF hospitalizations over the first 5 years of follow-up in the study.

Statistical Models for Recurrent Event Data

Poisson Regression

The Poisson regression model is very popular in modeling recurrent data in the literature (18–20). It is a technique that models the number of times that an event of interest has occurred, for example, the number of hospitalizations that have occurred over 10 years among patients with CKD. Different individuals, of course, have different numbers of events; some people had no hospitalizations, whereas others had many. However, these different numbers of events that are observed across different individuals tend to follow a certain pattern that can be described mathematically using a Poisson distribution. Poisson regression assumes that the outcome (i.e., the number of events of interest that happen in a given interval) follows a Poisson distribution with fixed rate of event occurrence over time. Like all other forms of regression, it attempts to explain the variation in this outcome that occurs across individuals, on the basis of certain known characteristics of the individuals (e.g., age, body mass index [BMI], BP, etc.). To account for the differential length of follow-up between individuals, one can include an offset term denoting the length of follow-up.

Figure 2A provides an example of three hypothetical individuals with different patterns of events. Patient A had two events before censoring at time point 8. Patient B experienced one event before censoring at time point 7. Patient C had no event through the end of follow-up at time point 9. Table 1 shows how the data for these three patients would be used when fitting a Poisson regression. In the dataset, each individual contributes one record, which includes the number of events as the outcome and length of follow-up as the offset term.

Cox Proportional Hazards Model

The three survival models that we review next for recurrent event data are extensions of the Cox proportional hazards model, which allows for the incorporation of covariates that may affect the hazard rate. These models include the Cox–Snell model, the Fine–Gray model, and the Andersen–Gill model. Each of these models has its own strengths and weaknesses, and the choice of model will depend on the specific research question and data available.
hazards model (9). Thus, we shall begin with a brief description of the standard Cox regression, which is a technique for modeling the hazard rate of a survival outcome. The hazard rate at a given time, $t$, is an instantaneous risk of experiencing an event, given that they have not experienced the event since the beginning of follow-up. The function describing how the hazard rate changes over time is called the hazard function. The survival data records the event time for those who experience an event and the end of follow-up time for those who never experience the event during follow-up, referred to as being censored. The model uses the data from each individual up until the time that the individual either experiences the event of interest or is censored. Thus, the parameters in the model are estimated on the basis of a subset of participants that differs from one time point to the next. This subset, i.e., the subset of participants who are still in the study and still at risk for the event of interest at a given point in time, is known as the risk set.

In Cox regression, the hazard of the outcome is modeled as a function of covariates. The model does not specify a particular form of baseline hazard function (i.e., the hazard when all covariates are set to their reference levels, similar to the intercept term in standard regression) and allows the baseline hazard to go up and down over time. It does, however, require the ratio of two hazards (for example, the hazard for men versus women) to remain constant over time (i.e., a constant difference when hazard is measured on the log scale, and a proportional difference on the exponential scale).

The Proportional Intensity Model

Andersen and Gill (4) proposed a model that generalizes the Cox proportional hazards model to allow for analysis of recurrent data, referred to as the proportional intensity model or Andersen–Gill model. The outcome in the proportional intensity model is the intensity function, which is conceptually similar to the hazard function, but does not require the individual to have not experienced the event before time $t$. In the proportional intensity model, as in standard Cox regression, the baseline intensity function is not assumed to have a specific parametric form, making this approach more flexible than Poisson regression (which, as described earlier, assumes that outcome follows a Poisson distribution).

In a standard Cox regression, which looks only at the time to first event, each individual is at risk for the event until he or she has experienced the event. In a recurrent events analysis, an individual is at risk for the same event throughout the follow-up period, regardless of whether an event has occurred or not. The Supplemental Appendix provides the risk set definition for the three hypothetical individuals in Figure 2 under different survival models.

![Figure 2. Diagrams of three hypothetical individuals with recurrent events.](image)

Table 1. Data structure for Poisson regression for the three hypothetical individuals illustrated in Figure 2A

<table>
<thead>
<tr>
<th>Patient</th>
<th>No. of Events</th>
<th>Length of Follow-Up Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>9</td>
</tr>
</tbody>
</table>

Number of events is used as the outcome and the length of follow-up time is used as an offset term in the model. Each individual contributes one record in the data.
The data structure for fitting the proportional intensity model is illustrated in Table 2. The column labeled “sequence number” represents the sequence of time intervals for each individual. The columns labeled “start time” and “end time” represent the start and end time of each interval in the original time scale, respectively. The “event indicator” represents whether an event occurs at the end of the time interval (with a value of one indicating an interval that ends with an event). As this table shows, each individual can contribute multiple records in the dataset. Because of the correlation of multiple events from the same individual, robust variance methods (21) may be used for variance estimation in the model.

**The Prentice, Williams, and Peterson Total Time Model**

In the proportional intensity model, the baseline intensity function is assumed to be the same across all events, regardless of their order. This means that, for example, the model would treat an individual’s third CHF hospitalization as being identical in all important respects to their first one. Prentice et al. (22) proposed a model considering each sequential event (first, second, etc.) separately. The model is often referred to as the Prentice, Williams, and Peterson (PWP) total time model (in contrast to the PWP gap time model introduced in the next section) because the time scale used in this model is the time from study entry. The model is stratified by the event sequence so that the baseline hazard function can differ between the sequential events. One can conceptualize it as if a separate model is fit for each event. Thus, the PWP total time model is essentially a stratified proportional hazards model. In principle, one can create as many strata as the maximum number of events experienced by any individual in the study. However, because each succeeding stratum has (by definition) fewer events than the previous one, the uppermost strata may contain very small numbers of events. In this case, some events can be combined to define the strata if the baseline hazards remain similar.

The PWP total time model allows any covariate to have different associations with different sequential events. A formal test of the covariate by stratum interaction can be conducted to check whether these associations are the same. If there is no evidence of heterogeneous associations, one can fit a simpler model with a common set of associations for all events. In the special case where the parameters for the first event are specified differently from those for the subsequent events, the coefficients of the covariates for the first events are identical to the standard Cox proportional hazards model for time to the first event.

The PWP total time model considers the events in order when defining the risk set. To construct a risk set for the $k^{th}$ event, an individual is included only if the individual has experienced $k - 1$ events and is still in the study. This ensures that the comparison at the time when a risk set is defined only happens among individuals who have had the same number of prior events and survived the same length of follow-up since study entry. The data structure used in the PWP total time model is the same as in the proportional intensity model (Table 2). An additional variable, however, is required to define the strata. The column labeled sequence number in Table 2 is unique to each sequential event and can be used to define the strata.

**The PWP Gap Time Model**

Prentice et al. (22) proposed another model that is similar to the PWP total time model, but differs in that the time is reset to zero whenever an event occurs; this model is referred to as the PWP gap time model. The rationale of resetting time is to match on the length of follow-up since the most recent event (or study entry if no event has happened) instead of study entry when constructing a risk set. The consequence is that all individuals in the same risk set have had the same number of prior events and survived the same length of follow-up since the most recent event. The start time and end time used in the PWP gap time model are also different (Table 2). The start time for all events is zero because of the time reset. The end time is the time from the most recent event, also called the gap time. In addition, the sequence number column is used to define the strata. Similar to the PWP total time model, the PWP gap time model also allows different associations with different events for any covariate.

**Analyses of the CRIC Study Data**

Over an average follow-up of 6.6 years in the CRIC Study, 638 participants (16%) experienced at least one CHF hospitalization. The average number of these hospitalization events
was 0.39 per person with an event rate of 6.0% per year. The MCF of CHF hospitalizations by different race categories is shown in Figure 1. Over time, the crude cumulative rates in non-Hispanic black and Hispanic race/ethnicity categories were higher than in the non-Hispanic white and other categories.

We fit four models to identify the risk factors for recurrent CHF hospitalizations (Poisson regression, proportional intensity model, PWP total time model, and PWP gap time model). Figure 3 shows the parameter estimates from the four models for risk factors including demographics and traditional and novel cardiovascular risk factors, adjusted for medication use. The results from the three Cox-type models were qualitatively similar. Older age was associated with increased risk of CHF. There was no difference by sex or kidney function, as measured by eGFR and 24-hour urine protein. Non-Hispanic blacks had higher risk of CHF hospitalizations compared with non-Hispanic whites. Diabetes and smoking, as well as higher BMI, LDL, fibroblast growth factor 23, NT-proBNP, and hsTnT, were all associated with increased risk of CHF hospitalization, whereas alcohol consumption and higher HDL were associated with decreased risk.

We further explored the risk factor associations with either the first or subsequent CHF events using the PWP total and gap time models. Note that the parameter estimates for the first event in the two PWP models are identical (Figure 4), which are also the same as in a standard Cox proportional hazards model for the time to first event. The two models also provided similar results (Figure 4) for subsequent events. A few predictors were found to have different relationships with the first and subsequent CHF events. For example, participants in the “other race” category had a higher risk of a first CHF hospitalization, but a lower risk of subsequent CHF hospitalizations, compared with non-Hispanic whites. Diabetes and higher BMI were associated with increased risk of first CHF only. Higher proteinuria was associated with increased risk of first CHF and decreased risk of subsequent CHF. Higher levels of NT-proBNP were associated with both first and subsequent CHF, although its association with the first CHF was much stronger. Higher levels of hsTnT were associated with the first CHF but not subsequent events.

Summary

We reviewed two classes of statistical models that are often used in the literature to analyze recurrent event data: Poisson regression and three Cox-type survival models. Poisson regression is a fully parametric model where the count of events is assumed to follow the Poisson distribution. The Cox-type survival models are extensions of the Cox proportional hazards model of time to a single event.

There are close connections between the two types of models. If survival time follows an exponential distribution, Poisson regression is equivalent to a parametric survival model assuming exponential distribution. It provides smaller confidence intervals than the Cox-type survival models. However, if survival time is not exponentially distributed and the event count does not follow a Poisson distribution, the inference made from Poisson regression may be biased. The CRIC Study data analysis showed empirically that some Poisson regression estimates are quantitatively different from the three Cox-type survival models. In general, caution should be used when using Poisson regression to analyze recurrent events (16) because of the strong parametric assumptions.

The selection among the three different Cox-type survival models is often related to the event mechanism. The proportional intensity model may be useful if the baseline risk does not change as a result of previous events, an assumption that is often unsatisfied. For example, a previous infection may change the risk of having additional infections. Because of this, the proportional intensity model is rarely used in practice. The two PWP models assume event-specific baseline hazard and are consequently often preferred to the proportional intensity model. When choosing which PWP model to use in analysis, an important consideration is whether to match on the time from study entry or since the last event when constructing the risk set.

The PWP total time model takes into account the time from study entry so that patients in the same risk set have the same time from study entry to when the risk set is defined. By contrast, in the PWP gap time model, patients in the same risk set have the same time from the last event (or study entry if no event has occurred) to the time at which the risk set is defined. One can also take into account both time from study entry and from the most recent event (22). For example, in the PWP total time model, one can include an interaction term between the time since the most recent event and an exposure of interest to evaluate whether the exposure-outcome association changes over the time from the most recent event. A similar approach can be taken in the PWP gap time model. In the CRIC Study example, the two PWP models gave similar results, suggesting that the results are not sensitive to the time scale being used and corresponding risk set definitions.

Recurrent event analysis provides additional insights compared with analyses of time to first event. In the CRIC Study example, the two PWP models provided evidence that some risk factors may have quite different associations with the first versus subsequent events. Further thought needs to be given to interpret these results. For example, there are at least two reasons that may explain the attenuated association of hsTnT with subsequent CHF events. One explanation is that hsTnT is not predictive of CHF among the prevalent population. Another possible explanation is that hsTnT is an acute marker of CHF in the sense that it is not predictive of CHF if it happens too far in the future. An analysis with updated hsTnT information over time may provide additional insights on the interpretation of these results.

In summary, we reviewed several statistical models for analyzing recurrent event data and illustrated them through the analysis of CRIC Study data. We demonstrated that recurrent event analyses can provide additional insights about the data compared with a standard survival analysis of time to the first event. Given the lack of distributional assumptions and flexible specification of baseline hazard functions, we suggest that both of the two PWP models we presented are a robust option for analyses of risk factors for recurrent events.
Figure 3. Parameter estimates from the four models for risk factors including demographics and traditional and novel cardiovascular risk factors, adjusted for medication use. Log hazards ratios and 95% confidence intervals for (A) demographic variables, (B) cardiovascular risk factors, and (C) novel risk factors from different models. The x-axis represents different risk factors and the y-axis represents the parameter estimates from the different models. The parameters are log rate ratio in the Poisson regression and log hazards ratio in the Cox-type models. For continuous predictors, the parameter estimates are per one SD increase in the independent predictor. Each predictor has four coefficients estimated from four different models. The dotted horizontal line represents no association between a predictor and outcome. The reference race category is non-Hispanic white. BMI, body mass index; FGF23, fibroblast growth factor 23; hsTnT, high-sensitivity troponin-T; NT-proBNP, N-terminal pro-B-type natriuretic peptide; PTH, parathyroid hormone; PWP, Prentice, Williams, and Peterson; SBP, systolic BP.
Figure 4. Further exploration of the risk factor associations with either the first or subsequent congestive heart failure (CHF) events using the Prentice, Williams, and Peterson (PWP) total and gap time models. Log hazards ratio and 95% confidence intervals for (A) demographic variables, (B) cardiovascular risk factors, and (C) novel risk factors stratified by first versus subsequent CHF events from the PWP total time and gap time models. The lines with a square shape in the middle represent the associations between risk factors and first CHF event. The lines with a circle in the middle represent the associations between risk factors and subsequent CHF events. Black lines represent results from the PWP total time model. Red lines represent results from the PWP gap time model. The numbers below the figure are the P values for the test of difference of the two associations for the first and subsequent CHF events. For continuous predictors, the parameter estimates are per one SD increase in the predictor. The reference race category is non-Hispanic white.
Acknowledgments

The authors thank the referees and editors for extensive and insightful comments, which led to a significantly improved article.

Funding for the Chronic Renal Insufficiency Cohort Study was obtained under a cooperative agreement from the National Institute of Diabetes and Digestive and Kidney Diseases (grants U01DK060990, U01DK060984, U01DK061022, U01DK061021, U01DK061028, U01DK060980, U01DK060963, and U01DK060902). In addition, this work was supported in part by the Perelman School of Medicine at the University of Pennsylvania Clinical and Translational Science Award (CTSA); the National Institutes of Health (NIH), National Center for Advancing Translational Sciences (NCATS) (UL1TR000003); Johns Hopkins University (UL1TR-000424); University of Maryland (The General Clinical Research Center M01 RR-16500); the Clinical and Translational Science Collaborative of Cleveland; the NCATS component of the NIH and NIH Roadmap for Medical Research (UL1TR000439); Michigan Institute for Clinical and Health Research (UL1TR000433); University of Illinois at Chicago CTSA (UL1RR029879); Tulane Center of Biomedical Research Excellence for Clinical and Translational Research in Cardiometabolic Diseases (P20 GM109036); and Kaiser Permanente NIH/Neurobehavioral Research Core for Rehabilitation Research University of California, San Francisco Clinical & Translational Science Institute (UL1 RR-024131).

Disclosures

None.

References


Published online ahead of print. Publication date available at www.cjasn.org.

This article contains supplemental material online at http://cjasn.asnjournals.orglookup/suppl/doi:10.2215/CJN.12841216/-/DCSupplemental.