Paper 255-2010

Analysis of Survival Data with Recurrent Events Using SAS ®

Rena Jie Sun[1], Daniel Cotton[2]
[1]University of Michigan, Ann Arbor, MI
[2]Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, CT

ABSTRACT

This paper presents the application of survival analysis methods using SAS/STAT ® to a large clinical trial, which was designed to test the treatment effect on preventing recurrent stroke and cardiovascular events. In this paper, we present the complete analysis procedure of this case study, including the model assumption check, model selection, and the utility and discussion on the Cox proportional hazards model, marginal recurrent events model, robust sandwich variance estimator, and some conditional recurrent events models.

INTRODUCTION

Data used in this paper were obtained from a megatrial. Due to confidentiality, we have blinded the trial name and the treatment name, and only presented the detailed analysis and the corresponding SAS code. This study was designed to compare an antihypertensive treatment (drug Z) with placebo in the presence of background antiplatelet treatment in the prevention of recurrent stroke. The study focused on testing the effect of Z on preventing recurrent strokes, with time to recurrent stroke as outcome of interest.

METHODOLOGY

Survival analysis approaches were used to analyze the time to recurrent stroke endpoint. The most common and widely used approach is the Cox proportional hazards model. Checking the model assumptions before using the Cox model is important and essential for model selection later on. We also conducted analyses to examine the potential effect of treatment over time, using both marginal and conditional recurrent events models. At the end, we discussed alternative recurrent events models and some remedies while considering 'vascular death' as dependent censoring.

MODEL VERIFICATION AND EXPLORATORY ANALYSIS

The proportional hazards assumption is very important for the use of the Cox model, second to the non-informative censoring assumption. In a regression type of setting, for a dichotomous variable this means that the survival curves for two strata must have hazard functions that are proportional over time (i.e. parallel hazard curves). This assumption is especially important for the covariate of interest. Firstly, we examined the survival function estimate by the treatment Z assignment, without any adjustment of other covariates. Kaplan-Meier curves on survival function versus follow-up time and cumulative hazards versus follow-up time (Figure 1) are examined. Another popular plot is the log-log plot (the log of cumulative hazards versus the log of follow-up time), which is essentially the same as cumulative hazards plot versus time, but might give a clearer picture of these hazard curves. As we can see, two survival curves by treatment tangled together at the beginning and crossed at the 9th month after the treatment starting date (Figure 1).

```
proc lifetest data=master1t method=km outsurv=survpar1;
    time tstop*status(0);
    strata Z;
run;
* Tstop = min(failure time, censoring time), represents the observed follow-up time;

symbol1 color=blue I=join V=none l=1;
symbol2 color=red I=join V=none l=2;
axis1 label=(angle=90 'Survival Distribution Function Estimate')
      order=(0.80, 0.85, 0.90, 0.95, 1);
axis2 label=(angle=90 'Follow-up Time');
proc gplot data=survpar1;
```

```
    title 'K-M plot by treatment assignment';
    plot survival*tstop=Z/vaxis=axis1 haxis=axis2;
run;

data survpar2;
  set survpar1;
  cumhaz=-log(survival);
  logcumhaz=log(cumhaz);
  logt=log(tstop);
run;

symbol1 color=blue I=join V=none l=1;
symbol2 color=red I=join V=none l=2;
axis1 label=(angle=90 'Cumulative Hazard')
      order=(0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30);
axis2 label=(angle=90 'Follow-up Time');
proc gplot data=survpar2;
   title 'Cumulative hazard plot by treatment assignment';
   plot cumhaz*tstop=Z/vaxis=axis1 haxis=axis2;
run;
```

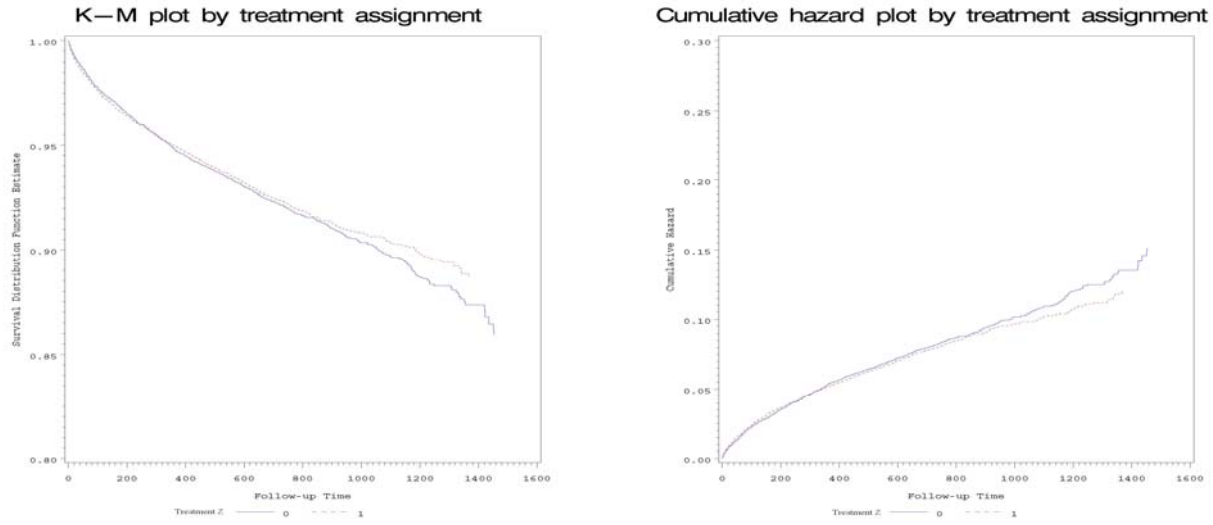

Figure 1: Kaplan-Meier plot and cumulative hazard plot by treatment assignment

One potential explanation of the crossed hazards above is the existence of the time dependent effect. We need to examine if this time dependency truly exists, or if it is due to the inadequate model specification.

We know that the usage of Angiotensin-Converting Enzyme (ACE) inhibitors reduce major cardiovascular events. Hence, assessing whether the drug Z would be effective in patients taking concomitant ACE inhibitors is important. Biologically, it is also believed the use of ACE has some unclear but very complicated interaction with treatment Z.

We first plotted a histogram by ACE and Z group, using PROC UNIVARIATE.

```
proc format;
    value trtfmt 1='Treatment'
                 0='Placebo';
run;
proc univariate data=master1t;
    where status=1;
    class acebase Z;
    histogram tstop / vscale=COUNT ;
    format acebase trtfmt.;
run;
```

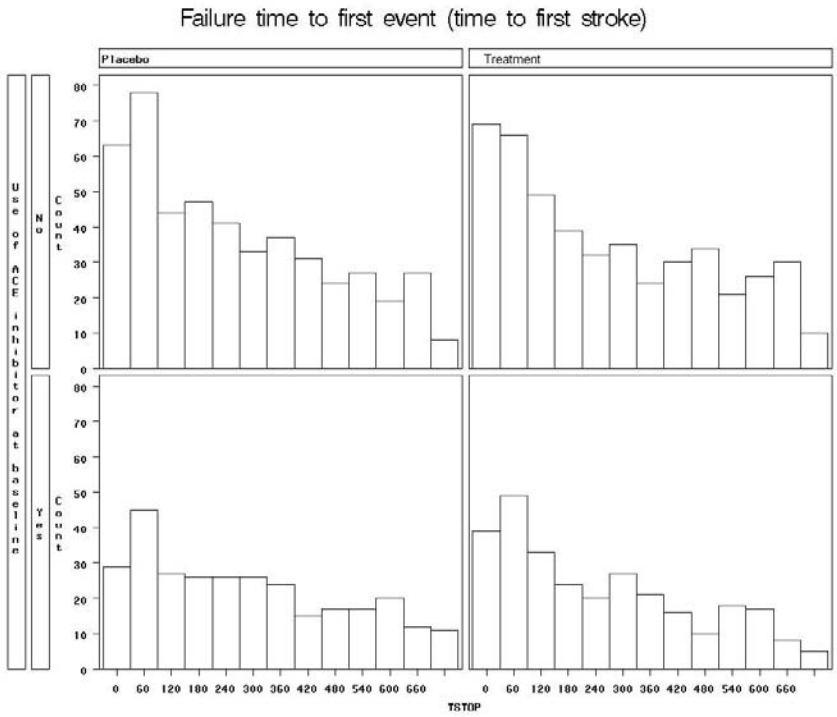

Figure 2: Histogram by treatment assignment and ACE inhibitor groups

From the raw plots (Figure 2), a discrepancy between ACE and non-ACE groups can be clearly seen, while the distributions by Z and non-Z are somewhat different. Due to the potentially unclear interaction between ACE and treatment Z, from the biological point of view, we decided to check the existence of the interaction. Stratification could potentially be used here. Note that stratification by ACE in the model can adjust for ACE without having to estimate its effect on the outcome, by modeling different baseline hazard by ACE categories. However, it assumes a constant effect of other covariates on the outcome across ACE groups. Due to this unclear difference between ACE and non-ACE groups, therefore, we decided to do separate analysis for ACE and non-ACE groups of patients.

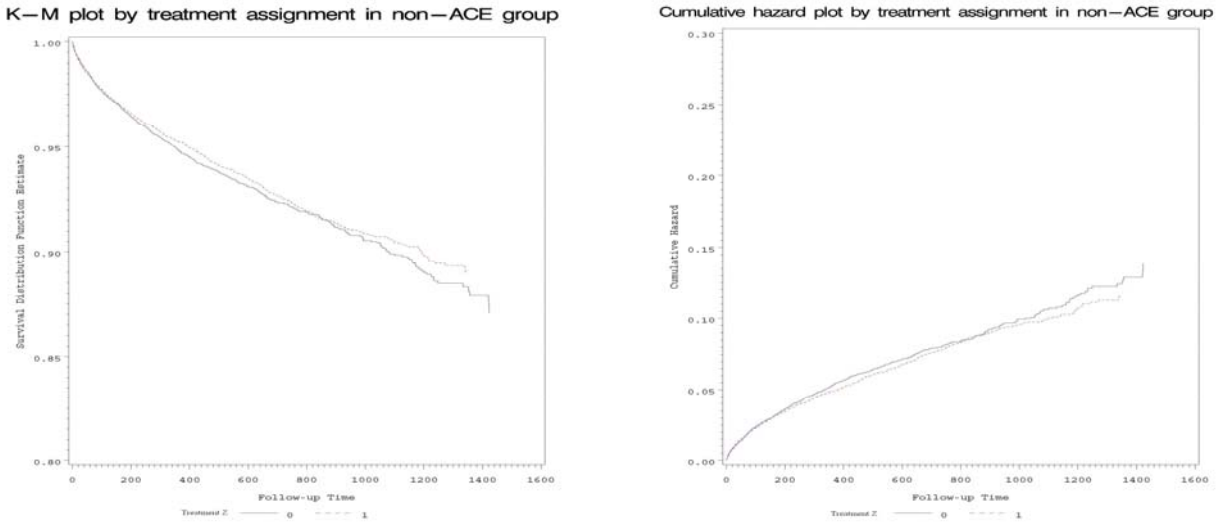

Figure 3: Kaplan-Meier plot and cumulative hazard plot by treatment assignment in non-ACE inhibitor group

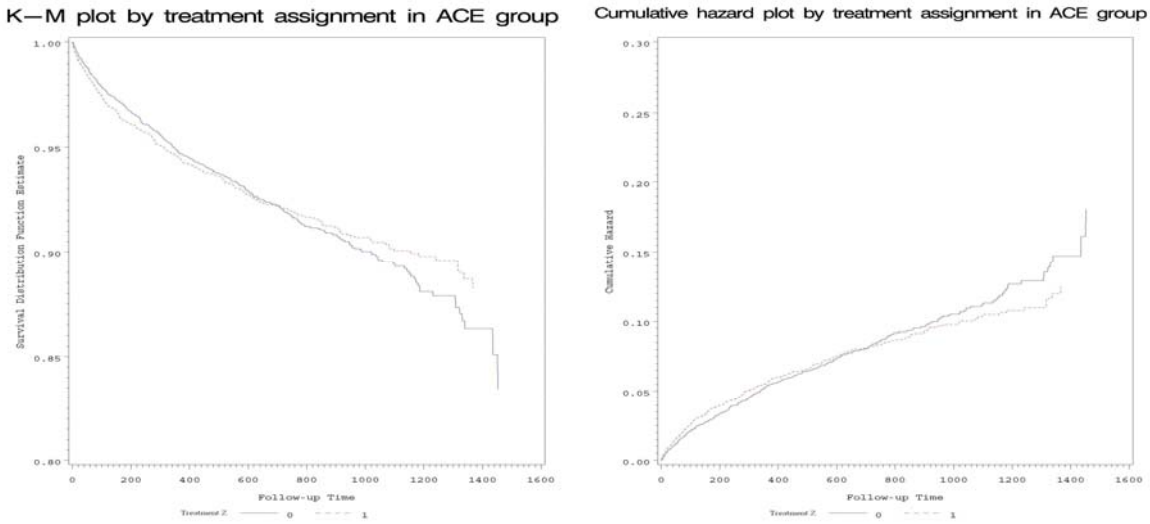

Figure 4: Kaplan-Meier plot and cumulative hazard plot by treatment assignment in ACE inhibitor group

By checking the Kaplan-Meier plots and cumulative hazard plots without adjusting for other covariates, two groups showed quite different characteristics. In the non-ACE group in Figure 3, the survival curves by treatment are very similar during the first couple of months. Then, there is a separation at the 6th months, but at the end of the second year, the curves come together then separate apart fairly quickly again. But for the ACE group, the survival curves crossed significantly at the end of the second year.

Now we conduct further exploratory analysis based on the adjusted results. From previous knowledge, we identified age, two dichotomous covariates (cov1 and cov2) and the usage of ACE inhibitor at baseline (acebase) as important factors that need to be adjusted for the model.

We then checked the time dependency, by inserting interactions between covariates and time and testing all interactions simultaneously by the TEST statement in PHREG. It gives the individual test results on each interaction, and a combined result through a generalized Wald test for a joint hypothesis on all interaction terms specified in the TEST statement. If the combination of the time dependent covariates is significant, then at least one of the covariates does not have a proportional effect on the hazard. It turned out that the proportionality assumption did not hold (p-value=0.024) in this setting.

```
proc phreg data=work.master1t;
    model tstop*status(0)= antip Z acebase age cov1 cov2 antipt Zt acebaset aget
                           cov1t cov2t/alpha=0.05 risklimits;
    antipt = antip*log(tstop);
    Zt = Z*log(tstop);
    acebaset = acebase *log(tstop);
    aget = age*log(tstop);
    cov1t = cov1*log(tstop);
    cov2t = cov2*log(tstop);
    test_proportionality: test antipt, Zt, acebaset, aget, cov1t, cov2t;
run;
```

The algorithm given above to check time dependency was also used in non-ACE and ACE inhibitor groups separately. No time dependent effect was found in the non-ACE group (p-value=0.43) meaning the proportionality assumption was not violated; while a significant time dependent effect exists in the ACE group (p-value=0.001).

Because the proportional hazards assumption holds for the non-ACE group, in the remainder of the paper, we only focus on this subgroup to demonstrate the usage of different survival models.

Correct functional forms (e.g. linear, quadratic, or log form) are also essential and would enhance the power of the model. The ASSESS statement is available in PROC PHREG ODS GRAPHICS, to plot the cumulative Martingale

residuals against the values of the covariate of interest in its specified functional form. The RESAMPLE option requests a Kolmogorov-type supremum test computed on a sample of 1000 simulated residual patterns and gives the p-value of the test. The CRPANEL option gives a plot of four panels with just a few paths from the default aggregate plot, for the simplicity of comparing simulated and observed paths visually. If the observed cumulative Martingale residual process for the covariate of interest is somewhat close to simulated realizations from the null distribution, with a p-value suggesting that the observed is not significantly different from the null, we can therefore conclude the functional form of the covariate of interest is appropriate. In our model, because only Age is continuous while other covariates are all dichotomous, we performed this check only on the Age variable.

```
ods graphics on;
proc phreg data= work.master1t;
    model tstop*status(0)= antip Z age cov1 cov2/alpha=0.05 risklimits;
    where acebase=0;
    assess var=(age) / resample=1000 seed=603708000 crpanel;
run;
ods graphics off;
```

STATISTICAL ANALYSIS

First, for the primary outcome of interest, we use the Cox model to test the treatment effect on the first recurrent stroke in the non-ACE patients. For a specific patient i, let $\lambda_i(t)$ represent the hazard function at time t, λ_0 represent the baseline hazard, and X_i be the covariate vector, so the Cox model has the form: $\lambda_i(t)= \lambda_0(t)\exp\{X_i\beta\}$.

```
proc phreg data=work.master1t;
    model tstop*status(0)= antip Z age cov1 cov2/alpha=0.05 risklimits;
    where acebase=0;
run;
```

It is worth noting here that cardiovascular related death would be dependent censoring for the stroke event in this case, which violates the basic independent censoring assumption of the Cox model. One remedy is to estimate the marginal survival function in the presence of dependent censoring. An alternative solution would be to look at the composite endpoint (first recurrent stroke or cardiovascular related death), instead of single endpoint of first recurrent stroke. Also, one can adjust the estimates by using an inverse weighting scheme. A few of these approaches were investigated, but the number of vascular related deaths before recurrent stroke was very small and equally distributed in the treatment and placebo groups, so it did not make a big difference. Thus, we will ignore this potential dependent censoring of vascular related death.

The analysis based on the first recurrence times cannot be used to examine if there is a treatment effect over time from Z on multiple events. So, recurrent event models were used in addition to time to first event models, to explore the treatment effect on the number of occurrences of events over time.

Firstly, Anderson and Gill (AG) intensity model with model-based variance (1982) was utilized to evaluate the treatment effect in the non-ACE group. AG model is a generalization of the Cox proportional hazards model and relates the intensity function of event recurrences to the covariates multiplicatively. It is a counting process approach, by treating each subject as a multi-event counting process with essentially independent increments. For a specific patient i, let $\lambda_{ik}(t)$ represent the hazard function for the k^{th} event at time t, λ_0 represent the common baseline hazard for all events/subjects, so AG model has the form: $\lambda_{ik}(t)= \lambda_0(t)\exp\{X_i\beta\}$.

```
proc summary data= freqtb NWAY;
    var count;
    output out=_m1 max=max_ct;
run;

data _null_;
    set _m1;
    call symput('max_ct', put(max_ct,2.));
run;

data master1;
    set master;
    by ptno;
    retain k lagtime;  /* k is # of reccurent events */
    if first.ptno then k=1;
    else k+1;
```

```
     if k eq 1 then tstart=0;
     else tstart=lagtime;
     lagtime=time;
     tstop=time;
     status=event;
     OUTPUT;
     l=k+1;
     if last.ptno AND k ne &max_ct then do i=l to &max_ct;
       if i=l then do; k=i; tstart=lagtime; tstop=fuptime; status=0; end;
       else do; k=i; tstart=fuptime; tstop=fuptime; status=0; end;
       OUTPUT;
     end;
     drop i l randt evtdt time fuptime randdt fpcdt event;
run;

proc phreg data=work.master1 COVM;
    model (tstart,tstop)*status(0)=antip Z age cov1 cov2/alpha=0.05 risklimits;
    ID ptno;
    where tstart < tstop and acebase=0;
run;
```

To take into account the within subject correlation, we also used the proportional means model with sandwich variance estimate, which provides the robust sandwich variance estimators for standard errors of coefficients, and do not require specification of the correlation matrix. Alternatively, one could construct a frailty model, essentially an AG model with a random effect, to take within subject correlation into account.

```
proc phreg data=work.master1 COVM COVS(aggregate);
    model (tstart,tstop)*status(0)= antip Z age cov1 cov2 /alpha=0.05 risklimits;
    ID ptno;
    where tstart < tstop and acebase=0;
run;
```

Next, we utilized the PWP (Prentice, Williams, Peterson; 1981) total time model with common effects and the PWP gap time model with common effects to further explore the temporal treatment effect on recurrent events.

The PWP models are conditional models, which specify that the hazard function at time t for the k^{th} recurrence of a patient, conditional on the filtration history prior to time t. In other words, a subject is assumed not to be at risk for $(k+1)^{th}$ event until the k^{th} event has occurred and terminated. This produces a proportional hazards model with time dependent strata, where the dependence between event times is handled by stratifying by the prior number of failures.

Depending on how the starting point of the risk interval is set, there are two variations of PWP models: PWP total time model and PWP gap time model. Here the 'total time' means the time from the start of treatment, and 'gap time' is the time from the prior event. The PWP total time model is similar to the counting process model (AG model) but stratified by event. Similar to the notation introduced for AG model, we also let λ_{0k} be the event-specific baseline hazard for the k^{th} event. PWP total time model has the form of $\lambda_{ik}(t)= \lambda_{0k}(t)\exp\{X_i\beta\}$ and PWP gap time model has the form of $\lambda_{ik}(t)= \lambda_{0k}(t-t_{k-1})\exp\{X_i\beta\}$.

```
/*create PWP dataset*/
data master2;
    retain lstatus;
    set master1;
    by ptno;
    if first.ptno then lstatus=1;
    if status=0 and lstatus=0 then delete;
    lstatus=status;
    gaptime=tstop-tstart;
run;

proc phreg data=master2;
    model tstop*status(0)= Z age cov1 cov2/alpha=0.05 risklimits;
    strata k;
    where acebase=0;
run;

proc phreg data=master2;
    model gaptime*status(0)= Z age cov1 cov2/alpha=0.05 risklimits;
```

```
    strata k;
    where acebase=0;
run;
```

Alternatively, one can consider a regression model for Poisson process, if one assumes the process of event occurrence is a homogeneous Poisson process. The number of cardiovascular events for each patient over time t can be seen as a Poisson random variable with mean λ_t, where λ is the rate of occurrence. PROC GENMOD can be implemented to carry out such analysis.

RESULTS

Table 1 presents the summary of results under various modeling schemes for the non-ACE inhibitor group of patients. The last two columns list the parameter estimate and the p-value of treatment Z versus placebo under the time to first event model (Cox PH model) and the time to recurrent events models (AG; AG with robust sandwich estimator; PWP total time model; PWP gap time model). Although treatment Z has fewer first recurrent strokes than placebo, the difference is not enough to make the treatment effect significant in the Cox PH model. For strokes after the first, treatment Z actually has more events than placebo; correspondingly, the treatment effect is washed out in the recurrent events model, with much higher p-values. Although all models suggest a numerical advantage for treatment Z compared to placebo, the difference is not statistically significant at a level of 5%.

Table 1: Number of recurrent strokes by treatment and analysis results from using Cox PH model, AG model, AG model with robust sandwich variance estimator, PWP total time model and PWP gap time model

# of recurrent strokes	Treatment Z		Placebo		Time to first recurrent stroke (Estimate; P-Value)	Recurrent Events Models (Estimate; P-Value)
	# of events	Average time to event	# of events	Average time to event		
1	548	371.3	573	378.2	Cox: -0.05; 0.38	AG: -0.03; 0.56
2	72	198.1	69	225.7		AG Robust: -0.03; 0.60
3	13	138.2	12	290.3		PWP Total: -0.02; 0.68
4	4	288.5	1	60		PWP Gap: -0.03; 0.62

CONCLUSION

It is important to check model assumptions before carrying out analysis using any statistical model. The proportional hazards assumption is especially important for the use of the Cox model, and is commonly or intentionally ignored. Through this case study we showed various ways to check the PH assumption graphically and analytically. We also present what we did in our case for the non-proportionality issue. A careful model selection procedure should not only involve much statistical evidence but also reasonable biological support. Additionally, we demonstrated that ASSESS statement in proc PHREG can be used to determine the best functional form of each covariate in the model.

In addition to time to the single event model (e.g. Cox model), a recurrent events model can gain many insights on the drug effects, if such recurrent outcomes exist. We discussed AG model and PWP models in detail. Although no significant treatment effect was shown in any of the recurrent events model in our particular dataset, such analyses are very important to conduct to give investigators a comprehensive understanding of the interventions being studied.

REFERENCES

[1] Cox D. Regression models and life-tables. Journal of Royal Statistical Society 1972; Ser B 34: 187-220.
[2] Anderson P.K. and Gill R.D. Cox's regression model for counting processes: a large sample study. Annals of Statistics 1982; 10:1100-1120.
[3] Lin D.Y. and Wei L.J. The robust inference for the Cox proportional hazards model. J Am Stat Assoc 1989; 84: 1074-1078.
[4] Prentice R.L., Williams B.J. and Peterson A.V. On the regression analysis of multivariate failure time data. Biometrika 1981; 68: 373-379.
[5] Guo Z., Gill T.M. and Allore H.G. Modeling repeated time-to-event health conditions with discontinuous risk

intervals: an example of a longitudinal study of functional disability among older persons. Methods Inf Med. 2008; 47(2): 107-116

ACKNOWLEDGMENTS

We would like to thank Boehringer Ingelheim Pharmaceuticals Inc. for providing the dataset introduced in the paper, and David Hall, Ph.D. for providing biostatistical guidance.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Rena J. Sun
1420 Washington Heights
University of Michigan
Ann Arbor, MI 48109
Email: renajsun@umich.edu