

Paper SD-39

Introducing Two-Way and Three-Way Interactions into the Cox Proportional Hazards Model Using SAS®

Seungyoung Hwang, Johns Hopkins University Bloomberg School of Public Health

ABSTRACT

The Cox proportional hazards model to explore the effect of explanatory variables on survival is by far the most popular and powerful statistical technique. It is used throughout a wide variety of types of clinical studies. However, special techniques are required when multiple interaction terms are introduced into the Cox model. This paper provides an in-depth analysis, with some explanation of the SAS code. It examines two-way and three-way interaction terms into the Cox proportional hazards model using SAS. Examples of using the PHREG procedure are drawn from the recently accepted article in the Journal of American Geriatrics Society (JAGS) (1).

CASE STUDY: DEPRESSION, DIABETES, AND MORTALITY

The Prevention of Suicide in Primary Care Elderly: Collaborative Trial (PROSPECT) was a cluster-randomized controlled trial designed to compare an algorithm-based intervention with usual care to reduce major risk factors of suicide (e.g., depression) for older primary care patients (2, 3).

The PROSPECT data contain survival times for 1,226 patients in 20 primary care practices. The median follow-up time was 98 months and ranged from 1 to 116 months. The twenty practices were randomly assigned to intervention or usual care. The variables are as follows:

time	is the time in months from the baseline interview to death.
death	1 = dead; 0 = censored.
intervention	1 = intervention practices; 0 = usual care practices.
practice	is the practice number for each patient.
major	1 = major depression; 0 = no depression.
minor	1 = minor depression; 0 = no depression.
diabetes	1 = diabetes; 0 = no diabetes.
age	is the age in years at baseline.
gender	1 = female; 0 = male.
education	is the number of years of schooling completed.
marital	1 = married; 0 = other.
smoking	1 = smoker; 0 = non-smoker.
cognition	is a measure of the cognitive impairment, ranging from 18 to 30 (lower scores indicating greater cognitive impairment).
suicidal	1 = suicidal ideation; 0 = no suicidal ideation.
charlson	is the Charlson comorbidity index at baseline.

I explored mortality risk for depression (major and minor) among primary care patients with diabetes in intervention practices and in usual care practices (results are shown below in **Table 1**).

Table 1. Mortality risk for major and minor depression compared to no depression for patients with diabetes within intervention or usual care, adjusted hazard ratios and 95% confidence intervals. Data from PROSPECT (1999-2008).

Medical condition	Depression status	Intervention practices	Usual care practices
		Hazard ratio (95% CI)	Hazard ratio (95% CI)
Diabetes	Major depression	0.79 (0.42 to 1.46)	2.41 (1.36 to 4.26)
	Minor depression	1.40 (0.67 to 2.90)	2.25 (1.12 to 4.53)
	No depression	1.00	1.00

Notes: Hazard ratios are from Cox proportional hazards models. Adjusted models include terms for characteristics identified by their association ($P < 0.05$) with time to death: baseline age, gender, education, marital status, smoking, cognition, suicidal ideation, and Charlson comorbidity index.

Among patients with diabetes, persons with major depression in usual care were significantly more likely to die during the course of follow-up than were persons in the same practices without depression (adjusted hazard ratio = 2.41, 95% confidence interval [1.36, 4.26]). In contrast, the mortality risk was attenuated in the intervention practices comparing persons with major depression to no depression (confidence interval includes the null). Similar patterns were identified for minor depression (adjusted HR = 2.25, 95% CI [1.12, 4.53] in usual care; adjusted HR = 1.40, 95% CI [0.67, 2.90] in intervention).

More importantly, I investigated whether a depression management program (i.e., intervention) improves mortality among older adults with diabetes (results are shown below in **Table 2**).

Table 2. Mortality risk for intervention compared to usual care for patients with diabetes within depression status, adjusted hazard ratios and 95% confidence intervals. Data from PROSPECT (1999-2008).

Medical condition	Depression status	Intervention practices	Usual care practices
		Hazard ratio (95% CI)	Hazard ratio (95% CI)
Diabetes	Major depression	0.47 (0.24 to 0.91)	1.00
	Minor depression	0.89 (0.48 to 1.62)	1.00

Notes: Hazard ratios are from Cox proportional hazards models. Adjusted models include terms for characteristics identified by their association ($P < 0.05$) with time to death: baseline age, gender, education, marital status, smoking, cognition, suicidal ideation, and Charlson comorbidity index.

I found older primary care patients with diabetes and major depression in practices implementing depression care management were less likely to die over the course of a 8-year interval than patients with diabetes and major depression in usual care practices (adjusted HR = 0.47, 95% CI [0.24, 0.91]). Stated differently, patients with diabetes and major depression were 53% less likely to die over follow-up if they had received the PROSPECT intervention. No statistically significant effects were observed for patients with diabetes and minor depression.

STATISTICAL ANALYSIS USING SAS

Here is the code for estimating the Cox regression models with PROC PHREG for the PROSPECT data:

```
PROC PHREG DATA=prospect COVS(AGGREGATE);
  MODEL time*death(0) = major minor intervention diabetes
                        inter1 inter2 inter3 inter4 inter5 inter6 inter7
                        age gender education marital smoking cognition
                        suicidal charlson;

  inter1 = major*intervention;
  inter2 = minor*intervention;
  inter3 = major*diabetes;
  inter4 = minor*diabetes;
  inter5 = intervention*diabetes;
  inter6 = major*intervention*diabetes;
  inter7 = minor*intervention*diabetes;

  X1_vs_X3: TEST major + inter1 + inter3 + inter6 / PRINT;
  X2_vs_X3: TEST minor + inter2 + inter4 + inter7 / PRINT;
  X7_vs_X9: TEST major + inter3 / PRINT;
  X8_vs_X9: TEST minor + inter4 / PRINT;

  X1_vs_X7: TEST intervention + inter1 + inter5 + inter6 / PRINT;
  X2_vs_X8: TEST intervention + inter2 + inter5 + inter7 / PRINT;

  ID practice;
  ODS OUTPUT TestPrint1 = output1;
RUN;

DATA output2;
  SET output1 (RENAME=(Col1=variance Col2=estimate));
  Hazard = exp(estimate);
  Lower_CL = exp(estimate - 1.96*sqrt(variance));
  Upper_CL = exp(estimate + 1.96*sqrt(variance));
RUN;

PROC PRINT DATA=output2 NOOBS;
  VAR Hazard Lower_CL Upper_CL;
  FORMAT Hazard Lower_CL Upper_CL 4.2;
RUN;
```

The COVS(AGGREGATE) option is specified to compute the robust sandwich covariance matrix estimate (4). The ID statement is specified in order to adjust standard errors for within-practice clustering in the Cox proportional hazards regression (5).

The analysis introduced three-way interactions (i.e., major*intervention*diabetes, minor*intervention*diabetes) into the Cox model, in addition to main effects (i.e., major, minor, intervention, diabetes), and the corresponding two-way interactions (i.e., major*intervention, minor*intervention, major* diabetes, minor* diabetes). Based on the above Cox three-way interaction model, hazard ratios for all-cause mortality, and corresponding 95% confidence intervals, were stratified by twelve groups defined by diabetes, depression status, and intervention condition as follows.

Medical condition	Depression status	Intervention practices	Usual care practices
Diabetes	Major depression	X1	X7
	Minor depression	X2	X8
	No depression	X3	X9
No diabetes	Major depression	X4	X10
	Minor depression	X5	X11
	No depression	X6	X12

X_i denotes a vector of covariates (explanatory variables) for group i ($1 \leq i \leq 12$). For example, diabetes = 1, major = 1, minor = 0, and intervention = 1 for X1 and diabetes = 1, major = 1, minor = 0, and intervention = 0 for X7. The adjusted hazard ratio of 0.47 in **Table 2** came from the ratio of the hazard rates for these two groups. That is,

$$\begin{aligned}
 HR &= \hat{h}(t, X1) / \hat{h}(t, X7) \\
 &= \hat{h}_0(t) \exp(\sum_{j=1}^p \hat{\beta}_j X1_j) / \hat{h}_0(t) \exp(\sum_{j=1}^p \hat{\beta}_j X7_j) \\
 &= \exp(\sum_{j=1}^p \hat{\beta}_j (X1_j - X7_j)) \\
 &= \exp(\hat{\beta}_1 \text{intervention} + \hat{\beta}_2 \text{major} * \text{intervention} + \hat{\beta}_3 \text{intervention} * \text{diabetes} + \hat{\beta}_4 \text{major} * \text{intervention} * \text{diabetes})
 \end{aligned}$$

where $\hat{h}_0(t)$ is the baseline hazard function. The fifth TEST statement (labeled as X1_vs_X7) computed the estimated coefficient ($\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \hat{\beta}_4$) and the corresponding variance estimate. The DATA step was used to change the variable names and calculate a hazard ratio with 95% CIs. The PROC PRINT statement was used to print the final results with two decimal places. The same analyses were applied to the remaining five TEST statements.

DISCUSSION

Recent decades have seen tremendous applications of survival analysis in various academic fields. However, few have reported detailed information on examining two-way and three-way interaction terms into the Cox proportional hazards model using SAS. In this paper, I examined whether primary care patients with diabetes were more or less likely to benefit from the depression care intervention using the test statement within the PROC PHREG procedure in SAS. Other specific medical conditions (e.g., heart disease, hypertension, peripheral vascular disease, stroke, cancer, chronic pulmonary disease) were addressed in the JAGS paper (1).

This paper provides the example SAS code to explain the use of interaction terms in the Cox proportional hazards model, along with step-by-step instructions to perform each technique in SAS. The emphasis is on statistical tools and model interpretations which are useful in medical follow-up studies and in general time-to-event studies. Moreover, this paper could be used as a valuable reference for biostatisticians, epidemiologists, and other healthcare professionals who are working in settings involving various lifetime events. Investigators interested in subgroup analysis in the randomized controlled trials should find it a useful guide for the incorporation of two-way and three-way interactions into their Cox models.

REFERENCES

1. Bogner HR, Joo JH, Hwang S, et al: Does a Depression Management Program Decrease Mortality in Older Adults with Specific Medical Conditions in Primary Care? An Exploratory Analysis. Journal of American Geriatrics Society 2015
2. Bruce ML, Ten Have TR, Reynolds CF, 3rd, et al: Reducing suicidal ideation and depressive symptoms in depressed older primary care patients: a randomized controlled trial. JAMA : the journal of the American Medical Association 2004; 291:1081-1091

3. Gallo JJ, Morales KH, Bogner HR, et al: Long term effect of depression care management on mortality in older adults: follow-up of cluster randomized clinical trial in primary care. *BMJ* 2013; 346:f2570
4. Allison PD: *Survival analysis using SAS: A practical guide.*, 2nd. SAS Institute, 2010
5. *SAS/STAT 9.3 User's Guide*, Cary, NC, SAS Institute Inc., 2011

ACKNOWLEDGMENTS

I took the opportunity to use data from the Prevention of Suicide in Primary Care Elderly: Collaborative Trial (PROSPECT) study. The PROSPECT is supported by a grant from the National Institute of Mental Health, United States (PI: Joseph J. Gallo, M.D., M.P.H.; R01 MH065539 and K24 MH070407).

This work is dedicated to my parents, Jungja Han and Okgil Hwang. All I have and will accomplish are only possible due to their love and sacrifices. As always, most special thanks to Yun Kyoung Ryu for her advice and encouragement in getting me to finish this paper. She has been incredibly generous with her time in reviewing the draft and making many helpful suggestions. All remaining errors, omissions, and weaknesses are my sole responsibility.

CONTACT INFORMATION

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

Seungyoung Hwang, MS, MSE, GStat®
 Biostatistician, Department of Mental Health,
 DrPH Student, Department of Health Policy and Management
 Bloomberg School of Public Health
 Johns Hopkins University
 624 North Broadway, Hampton House 880
 Baltimore, MD 21205
 Phone: 410-440-2040
 Email: shwang25@jhu.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.