World Applied Sciences Journal 19 (5): 608-614, 2012
ISSN 1818-4952;

DOI: 10.5829/idosi.wasj.2012.19.05.3054

Elimination of Causes in Competing Risks: A Hazards Model Approach

[1]M. Ataharul Islam and [2]Rafiqul I. Chowdhury

[1]Department of Statistics and OR, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
[2]Department of Epidemiology and Biostatistics, University of Western Ontario, Canada

Abstract: The problem concerning elimination of causes and its impact has been a longstanding problem in the field of survival analysis. In this paper, a covariate dependent model based on proportional hazards has been proposed. In addition, a test procedure is suggested for testing the independence of competing risks based on Islam [1, 2]. The estimation and test procedures are applied to analyze the competing causes in maternal morbidity data from Bangladesh. The results confirm that although there are slight changes in the estimates for before and after elimination of causes, the competing causes satisfy independence hypothesis.

Key words: Cause elimination · Competing risks · Multiple failure types · Test for independence

INTRODUCTION

It is commonly observed in the fields of survival analysis and reliability that we need to model for multiple failure types. Some examples are: (i) we may consider various causes of deaths or failures, (ii) different types of diseases or failures, (iii) failure of different components of machines, etc. Some of these events are transient and others are absorbing. In both the situations, we need to know the behaviour of survival functions in the presence and absence of certain type of cause.

The first work in the field of competing risks can be traced back to Bernoulli in 1766. He tried to measure the impact of removing smallpox as a cause of death. Seal [3] provided a nice review of competing risks theory where an attempt was made to summarize the findings from the literature to find an answer to the question that given two states A and B such that individuals in state A have mutually exclusive probabilities, possibly dependent on the time spent in state A, of having that state because of death or passage to state B, what is the probability of of an individual passing to state B and dying there within a given period? Karn [4] discussed about the role of increase in survival by eliminating the causes of death. The elimination of causes and its impact in the presence of covariates has been studied by Pocock *et al*. [5] and Hakulinen and Tenkanen [6]. The independence assumption of failure types has been addressed by number of researchers [7-10]. Oakes [11] observed that if latent failure times for U causes assumed continuous for each type of failure then still the observable data are single variable for failure time, T. The joint distribution of failure time for each cause is not identifiable.

In this paper, we have examined the role of covariates before and after elimination of a certain cause. This was addressed by Prentice *et al*. [12] and since then many others made attempts to employ the competing cause duration data [13, 14]. However, Kay [15] and Islam [1] provided the outline for dealing with competing risk as a special case of much broader multistate analysis.

In this paper, we propose a model based on proportional hazards approach to show the impact of elimination of causes on the hazard functions or transition probabilities. A simple method is suggested in this paper for testing the independence of competing causes. The estimation and test procedures are illustrated on the basis of data on maternal morbidity.

The Cause Specific Hazard Functions: Let us define the hazard function as a function of covariate vector Z:

$$\lambda(t;Z) = \lim_{\Delta t \to 0} P\{t \le T \le t + \Delta t; Z(t)\} / \Delta t$$

Where Z(t) denotes the value of the regression vector of representing p covariates ($Z(t) = [Z_1(t),Z_2(t),...,Z_p(t)]$) at time t. The above hazard function is used for single transient and single absorbing states.

Corresponding Author: M. Ataharul Islam, Department of Statistics and OR, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia. Tel: +966-597779806.

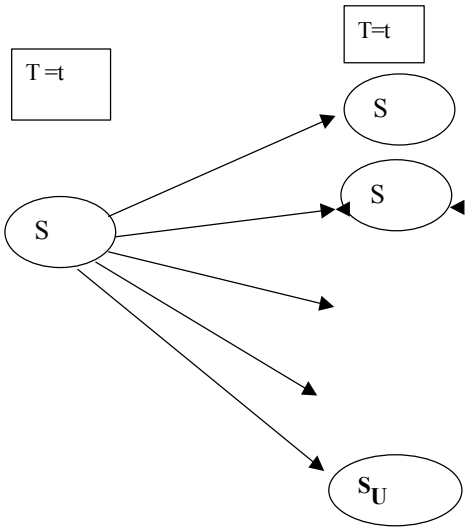

Fig. 2.1: States and Transitions for Failure Time Analysis with Competing Risks

Prentice *et al*. [12] and Farewell [16] extended this for competing causes. Cause specific hazard functions [17-21] are defined by

$$\lambda_u(t;Z(t)) = \lim_{\Delta t \to 0} P\{t \le T \le t+\Delta t \mid T \ge t, S = u \mid T \ge t; Z(t)\}/\Delta t$$

for u=1,2,……,U. The function λ_u (*t; Z*(*t*)) gives the instantaneous failure rate from cause u at time t, given the regression vector Z(t), in the presence of other failure types.

Assuming distinct failure types, the overall hazard function can be expressed in terms of cause-specific hazard function as

$$\lambda(t;Z(t)) = \sum_{u=1}^{U} \lambda_u(t;Z(t))$$

The overall survivor function is

$$S(t;Z(t)) = e^{-\int_0^t \lambda(\tau;Z(\tau))d\tau}$$

and the probability function for time to failure and cause of failure

$$f_u(t;Z(t)) = \lambda_u(t;Z(t)).S(t;Z(t))$$

Let us consider that there are U causes of failure at any stage. If we denote u=1,2,…,U for the causes of failure, then let us assume that each cause at the elimination will be considered as the last one, U. We are assuming only one cause of elimination at a time. The cause specific hazard function can be expressed as follows:

$$\lambda_{iu}(t;z) = \lambda_{0iu}(t)e^{Z_{iu}(t)\beta_u}$$

where β_u is the vector of regression coefficients corresponding to covariate vector Z_{iu} for the failure of type u (u=1,2,…,U).

Age Specific Hazard Functions and Elimination of Cause: Following Pocock *et al*. [5] Hakulinen and Tenkanen [6] proposed a model for cause elimination on the basis of grouped data for the follow-up interval. They considered G groups of subintervals denoted by [t_k, t_{k+1}, k=1,2,…,G. Then the combined hazard function is expressed as

$$\lambda_k(t;z) = \lambda_{0k}(t)e^{Z_{uk}(t)\beta_k} + \lambda_k^{-U}(t;z)$$

where λ^{-U} (t;z) is the hazard function in a population free of the particular disease U. They also defined

$$\lambda_{kU}(t;z) = \lambda_k(t;z) - \lambda_k^{-U}(t;z)$$

as the hazard function due to disease U. Pocock *et al*. [5] defined the expected survival function for the subinterval as

$$S_k^{-U}(t) = e^{-\int_{t_k}^{t_{k+1}} \lambda_k^{-U}(\tau)d\tau}$$

One of the weaknesses of this approach is that Pocock *et al*. [5] assumed λ^{-U} (*t; z*) as constant because it is based on a section of the large population. In reality, the hazard functions may vary widely over time and particularly when we consider the proportional hazards regression model this can be misleading. Hence we propose an alternative approach of solving this problem.

Proposed Hazard Model for Elimination of Cause: We propose the following cause specific hazard function after elimination of cause U:

$$\lambda_{iu}^{-U}(t;z) = \lambda_{0u}^{-U}(t)e^{Z_{iu}(t)\beta_u^{-U}}$$

Where i=1,2,…,n and β_u^{-U} is the vector of regression coefficients corresponding to covariate vector Z_{iu}, where

$$Z(t) = [Z_1(t), Z_2(t), ..., Z_p(t)]$$

Estimation and Test of Hypothesis: Kay [15] and Islam [1] showed that the shape functions can be cancelled for numerator and denominator for the partial likelihood contribution for transitions as well as for repeated transitions. The model for competing risk without cause elimination can be estimated from the following likelihood function:

$$\text{Model I:} L(\beta\ ) = \prod_{u=1=1}^{U}\prod_{i=1}^{n}\frac{e^{\beta_u Z_{iu}(t_{iu})}}{\sum_{l\in R(t_{iu})}\beta_u Z_l(t_l)} \qquad (4.1)$$

If the cause U is eliminated then we may consider all the individuals making transition to U as censored and the likelihood function is:

$$\text{Model II:} L(\beta^{-U}) = \prod_{u=1}^{U-1}\prod_{i=1}^{n}\frac{e^{\beta_u^{-U} Z_{iu}(t_{iu})}}{\sum_{l\in R(t_{iu})}\beta_u^{-U} Z_l(t_l)} \qquad (4.2)$$

We can estimate the pth parameter of u$^{-\text{th}}$ cause as follows:

$$\text{Model I:}\frac{\partial \ln L}{\partial \beta_{pu}} = 0, p = 1,2,....,P; u = 0,1,2,.....,U.$$

$$\text{Model II:}\frac{\partial \ln L}{\partial \beta_{pu}^{-U}} = 0, p = 1,2,....,P; u = 0,12,....,U-1.$$

The (p,q)th element of the observed information matrix can be obtained from equations 3.6 and covariance matrix $I^{-1}(\beta)$ is obtained from the following:

$$I_{pq}(\beta) = -\frac{\partial^2 \ln L}{\partial\, \beta_p^{-U} \partial\, \beta_q^{-U}}$$

For attaining convergence in, we have used the Newton-Raphson method as shown below:

$$\hat{\beta}_j = \hat{\beta}_{j-1} + \Gamma^{-1}\left(\hat{\beta}_{j-1}\right) \mathrm{U}\left(\hat{\beta}_{j-1}\right)$$

Where $\hat{\beta}_j$ denotes the estimate at j-th iteration. Here β is a column vector comprising of all the parameters for the models for transitions. For testing the hypotheses H_o: $\beta = 0$ we can use the asymptotic chi-square for large samples as follows:

$$\chi^2 = U'(\beta) I^{-1}(\beta) U(\beta)$$

Test for Independence: For testing of independence of causes, Farewell [16] suggested that the occurrence of one type of cause can be represented in the Z(t) vector in the hazard function for the other type as binary variable coded 1 after the occurrence of cause and 0 otherwise. According to Farewell, a test for a non-zero regression co-efficient corresponding to this variable is then a test of independence. On the other hand, Oakes [11] suggested that the time to failure T and the type of failure J may be analyzed separately, perhaps using T as an explanatory variable in the analysis of J. The hazard function for cause J=j are proportional as functions of t if and only if T and J are independent. Oakes [11] referred to Koziol and Green [22] for this test. Both the test results can be misleading due to the selection of variables or censoring pattern, respectively. Hence, it is evident that still there is no direct test of independence on the basis of estimates obtained before and after elimination. In this section, we demonstrate that the test proposed by Islam [1] and Islam *et al*. [2] can be employed conveniently for testing independence of causes in analyzing before and after elimination of a cause using proportional hazards model. The original test was proposed for equality of parameters for transitions and repeated transitions. In this case, we can employ the same procedure for estimates of parameters before and after elimination. If there is no statistically significant change after elimination of a disease then it is indicative of the fact that there is no violation of independence of causes.

For equality of parameters before and after elimination of cause U, we can employ the test developed by Islam [1] and Islam *et al*. [2]. To test the equality of two sets of parameters, the null hypothesis is $H_0 : \beta_1 = \beta_2$ and the asymptotic chi-square for large samples is

$$\chi^2 = (\hat{\beta}_1 - \hat{\beta}_2)' I(\hat{\beta})(\hat{\beta}_1 - \hat{\beta}_2) \qquad (4.3)$$

where $I^{-1}(\hat{\beta})$ is the pooled variance-covariance matrix for $\hat{\beta}_1$ and $\hat{\beta}_2$. For more details see Islam [1] and Islam *et al*. [2].

We can apply the same test procedure on the basis of inclusion of extended data for the individuals under consideration. Instead of assuming only the first cause of failure, if we consider a second cause by extending the time after the first cause, then for some individuals incidence of the second cause can also be observed. If we denote the first and second causes, both, as U (the cause of failure) then we can employ expression 4.2 for the combined causes. The possibility of such test is indicated

in Kalbfleisch and Prentice [23, pp262]. Then using the test procedure 4.3, we can test for independence for the combined causes U. This will reveal the pattern of independence more specifically. If the null hypothesis of independence is rejected then the p-value will demonstrate value less than 0.05 or 0.01.

Estimation of the Hazard Function: For estimating the cause specific hazard function

$$\lambda_{iu}^{-U}(t;z) = \lambda_{0u}^{-U}(t)e^{Z_{iu}(t)\beta_u^{-U}}$$

We need to estimate the underlying hazard, $\lambda_{0u}^{-U}(t)$, after estimates of the hazard function parameters are obtained. Kalbfleisch and Prentice [23], Kay [15] and Islam [1] showed that the underlying hazard function can be approximated by the step function

$$\lambda_{0u}^{-U}(t) = \lambda_{ul}^{-U}, \quad t \in (b_{m-1}, b_m] = I_l$$

where timescale is divided into pre-specified points denoted by b_l, l=1,2,…,r. Then if the total number of transitions to cause u is d_{ul} during the time interval I_l, then the estimate of the underlying hazard function is

$$\hat{\lambda}_{ul} = \frac{d_{ul}}{\sum_{u \in R(b_l)} e^{Z_{iu}(t)\beta_u^{-U}}}$$

Application: This study uses the data from the Maternal Morbidity in Bangladesh conducted during November 1992 to December 1993. The prospective data from the study is employed in the study. A total of 1020 pregnant women were followed up (993 had at least one follow-up) to delivery and postnatal stage. Several complications were recorded during the antenatal period such as haemorrhage, cough/fever for more than 3 days and fits/convulsion. These three complications are considered as potential causes in this application as the causes of the competing risks. The time of failure is considered since the entrance in the study. It is noteworthy that we have considered the time of the first reported occurrence in this analysis for these three causes for the results displayed in Tables 1-3. The joint occurrences are being considered in Table 4. The covariates are Age at Marriage (<=15=0, 16+=1), Education (No education=0, Some schooling=1), Wanted Pregnancy (No=0, Yes=1).

Table 1 shows that 46.4 percent of the respondents were free from the three complications, 12.5 percent suffered from haemorrhage, 36.9 percent mentioned about cough/fever for more than 3 days and the incidence of fits/convulsion recorded for 4.2 percent of the respondents in the study as the first cause.

Table 1: Distribution of three Pregnancy Complications for the First Time Occurrences

Causes	N	%
No Complications	461	46.4
Haemorrhege	124	12.5
Cough/Fever >3 days	366	36.9
Fits/Convulsion	42	4.2
Total	993	100.0

Table 2: Competing Risk proportional Hazards Regression Models without Elimination of Causes for Pregnancy Complications

Causes and Variables	Coefficients	Standard error	t-Value	p-Value
Haemorrhege				
Age at Marriage	-0.523	0.211	-2.479	0.013
Education	-0.337	0.187	-1.804	0.071
Wanted Pregnancy	-0.334	0.190	-1.755	0.079
Cough/Fever >3 days				
Age at Marriage	-0.106	0.112	-0.953	0.340
Education	-0.058	0.106	-0.550	0.582
Wanted Pregnancy	-0.417	0.108	-3.855	0.000
Fits/Convulsion				
Age at Marriage	0.206	0.316	0.650	0.515
Education	-0.159	0.313	-0.510	0.610
Wanted Pregnancy	-0.968	0.311	-3.115	0.002
Model Chi-square	43.302 (p-value= 0.000002)			

Table 3: Competing Risk proportional Hazards Regression Models with Elimination of a Cause for Pregnancy Complications

Models and Variables	Coefficients	Standard error	t-Value	p-Value
Model I		After Elimination of Haemorrhege		
Cough/Fever >3 days				
Age at Marriage	-0.059	0.113	-0.522	0.602
Education	-0.027	0.072	-0.382	0.702
Wanted Pregnancy	-0.403	0.109	-3.716	0.000
Fits/Convulsion				
Age at Marriage	0.294	0.320	0.917	0.359
Education	0.016	0.203	0.081	0.936
Wanted Pregnancy	-0.942	0.312	-3.020	0.003
Model Chi-square		25.402 (p-value=0.00023)		
Model II		After Elimination of Cough/Fever >3 days		
Haemorrhege				
Age at Marriage	-0.466	0.212	-2.197	0.028
Education	-0.240	0.134	-1.790	0.073
Wanted Pregnancy	-0.194	0.190	-1.024	0.306
Fits/Convulsion				
Age at Marriage	0.233	0.320	0.728	0.467
Education	0.030	0.205	0.146	0.884
Wanted Pregnancy	-0.812	0.311	-2.614	0.009
Model Chi-square		18.736 (p-value=0.004632)		
Model III		After Elimination of Fits/Convulsion		
Haemorrhege				
Age at Marriage	-0.536	0.211	-2.547	0.011
Education	-0.327	0.187	-1.755	0.079
Wanted Pregnancy	-0.258	0.190	-1.356	0.175
Cough/Fever >3 days				
Age at Marriage	-0.104	0.111	-0.937	0.349
Education	-0.053	0.106	-0.503	0.615
Wanted Pregnancy	-0.357	0.108	-3.296	0.001
Model Chi-square		26.044 (p-value=0.00022)		

Table 2 shows the fit of the competing risk model with or without elimination of pregnancy complications. t is evident from Table 1 that before elimination, haemorrhage is negatively associated with age at marriage ($p<0.05$), education ($p<0.10$) and wanted pregnancy ($p<0.10$). For analyzing cough/fever for more than 3 days, we observe that only wanted pregnancy is negatively associated (p-value<0.01). Similar finding is demonstrated for fits/convulsion, negative association with wanted pregnancy (p-value<0.01). The chi square test for independence shows that there is no strong evidence of rejecting the null hypothesis of independence. The results are summarized in Table 3. After elimination of causes, we observe that there are some changes in the estimates after elimination of causes cough/fever for more than 3 days and fits/convulsion but the test for independence presented in Table 5 from the panel for elimination of one cause, the changes are not statistically significant after employing the test proposed by Islam [1]. The test results suggest independence of causes in the competing risk framework. Table 4 displays the results based on elimination of two causes occurred to respondents. In this case also, although we observed some changes in the estimates after elimination of two causes, we observe from the second panel of Table 5 that the hypothesis of independence of causes is still valid.

Table 4: Competing Risk proportional Hazards Regression Models with Elimination of Two Causes for Pregnancy Complications

Causes and Variables	Coefficients	Standard error	t-Value	p-Value
Model I: After Elimination of Cough/Fever >3 days and Fits/Convulsion				
Haemorrhege				
Age at Marriage	-0.469	0.212	-2.214	0.027
Education	-0.230	0.134	-1.715	0.086
Wanted Pregnancy	-0.053	0.190	-0.277	0.782
Model Chi-square		9.509 (p-value= 0.02323)		
Model II: After Elimination of Haemorrhege and Fits/Convulsion				
Cough/Fever >3 days				
Age at Marriage	-0.058	0.112	-0.519	0.603
Education	-0.029	0.072	-0.400	0.689
Wanted Pregnancy	-0.346	0.108	-3.195	0.001
Model Chi-square		11.464 (p-value= 0.00946)		
Model III: After Elimination of Haemorrhege and Cough/Fever >3 days				
Fits/Convulsion				
Age at Marriage	0.287	0.287	0.799	0.371
Education	0.054	0.054	0.070	0.792
Wanted Pregnancy	-0.806	-0.806	6.704	0.010
Model Chi-square		7.528 (p-value= 0.0570)		

Table 5: Tests for Independence

Causes Before and After Elimination	Chi-square	p-value
Elimination of one cause		
Before and after elimination of Haemorrhege	0.96	0.987
Before and after elimination of Cough/Fever >3 days	1.96	0.923
Before and after elimination of Fits/Convulsion	0.49	0.998
Elimination of two causes		
Before and After Elimination of Cough/Fever >3 days and Fits/Convulsion	2.99	0.393
Before and After Elimination of Haemorrhege and Fits/Convulsion	0.84	0.841
Before and After Elimination of Haemorrhege and Cough/Fever >3 days	1.12	0.772

Concluding Remarks: The problem of cause elimination under competing risk framework has been a longstanding issue of concern among the researchers and potential users for formulating policies. This paper highlights a procedure based on the proportional hazards model to estimate the covariates associated with competing risks before and after elimination of causes. In addition, a test for independence of competing risks is also employed. It has been shown that the test procedure proposed by Islam [1] initially developed for testing equality of parameters in multistate transition and repeated transition models using hazard functions can be used for testing independence of causes conveniently. The applications to maternal morbidity complications indicate that the proposed cause elimination in proportional hazards model under competing risks provides the estimates of the factors before and after cause elimination along with the independence of competing causes without difficulty.

ACKNOWLEDGEMENTS

We are grateful to Dr. Halida Hanum Akhter (ex-Director, BIRPERHT) for providing us the data employed in this study. The authors would like to thank Mahbub E Elahi K. Chowdhury and Arindom Sen, for their assistance during different phases of this work. The authors are greatly indebted to the Ford Foundation for funding the data collection of the maternal morbidity study. The project was supported by the King Saud University, Deanship of scientific research, Collage of science research centre.

REFERENCES

1. Islam, M.A., 1994. Multistate Survival Models for Transitions and Reverse Transitions: An Application to Contraceptive Use Data. Journal of the Royal Statistical Society A, 157: 441-455.
2. Islam, M.A., R.I. Chowdhury, N. Chakraborty and W. Bari, 2004. A Multistage Model for Maternal Morbidity During Antenatal, Delivery and Postpartum Periods. Statistics in Medicine, 23: 137-158.
3. Seal, H.L., 1977. Studies in the History of Probability and Statistics.XXXV Multiple Decrements or Competing Risks. Biometrika, 64: 429-39.
4. Karn, M.N., 1933. A Further Study of Methods of Constructing Life Tables when certain Causes of Death are Eliminated. Biometrika, 25: 91-101.
5. Pocock, S.J., S.M. Gore and G.R. Kerr, 1982. Long Term Survival Analysis: the Curability of Breast Cancer. Statistics in Medicine, 1: 93-104.
6. Hakulinen, T. and L. Tenkanen, 1987. Regression Analysis of Relative Survival Rates. Applied Statistics, 36: 309-317.
7. Williams, J.S. and S.W. Lagakos, 1977. Models for Censored Survival Analysis: Constant-Sum and Variable-sum Models. Biometrika, 64: 215-24.
8. Kalbfleisch, J.D. and R.J. McKay, 1979. On Constant-Sum Models for Censored Survival Data. Biometrika, 66: 87-90.
9. Slud, E.V. and L.V. Rubinstein, 1983. Dependent Competing Risks and Summary Survival Curves. Biometrika, 70: 643-9.
10. Zheng, M. and J.P. Klein, 1995. Estimates of Marginal Survival for Dependent Competing Risks Based on an Assumed Copula. Biometrika, 82: 127-38.
11. Oakes, D., 2001. Biometrika Centenary: Survival Analysis. Biometrika, 88: 99-142.
12. Prentice, R.L., J.D. Kalbfleisch, A.V. Peterson, N. Flournoy, V.T. Farewell and N.E. Breslow, 1978. The Analysis of Failure Times in the Presence of Competing Risks. Biometrics, 34: 541-54.
13. Alvarez, R.M. and J. Nagler, 1995. Economics, Issues and the Perot Candidacy: Voter Choice in the 1992 Presidential Election. American Journal of Political Science, 39: 714-744.
14. Gordon, S.C., 2002. Stochastic Dependence in Competing Risks. American Journal of Political Science, 46: 200-217.
15. Kay, R., 1982. The Analysis of Transition Times in Multistate Stochastic Processes Using Proportional Hazard Models. Communications in Statistics Theory and Methods, 11: 1743-56.
16. Farewell, V.T., 1979. An Application of Cox's Proportional Hazards Model to Multiple Infection Data. Applied Statistics, 28: 136-143.
17. Chiang, C.L., 1968. Introduction to Stochastic Processes in Biostatistics. Wiley, New York..
18. Altshuler, B., 1970. Theory for the Measurement of Competing Risks in Animal Experiments. Mathematical Bioscience, pp: 6-11.
19. Holt, J.D., 1978. Competing Risk Analyses with Special Reference to Matched Pair Experiments. Biometrika, 65: 159-165.
20. Beck, G.J., 1979. Stochastic Survival Models with Competing Risks and Covariates. Biometrics, 35: 427-438.
21. Prentice, R.L. and N.E. Breslow, 1978. Retrospective Studies and Failure Time Models. Biometrika, 65: 153-158.
22. Koziol, J.A. and S.B. Green, A Cramer-Von Mises Statistics for Randomly Censored Data. *Biometrika* 1976; 63:465-474.
23. Kalbfleisch, J.D. and R.L. Prentice, 2002. The Statistical Analysis of Failure Time Data. 2nd Edition. Wiley, New York.