Risk factors for cardiovascular diseases (CVDs) patients in Bhutan

Yeshi Dorji and Montip Tiensuwan

Department of Mathematics, Faculty of Science, Mahidol University, Bangkok, Thailand

Email: yeepee89@gmail.com

Abstract. This paper aims to identify a survival curve involving risk factors for cardiovascular diseases patients in Bhutan. There are 5,206 cardiovascular diseases patients recorded in Jigme Dorji Wangchuck National Referral Hospital (JDWNRH) medical record department as of January 2006 to December 2016. In addition, it examines a relationship between survival time and some risk factors by using univariate and multivariate Cox’s proportional hazard models. Application of Log-rank test and Cox’s proportional hazard model compares the survival curves and find the covariates that affects the survival time. The results showed that, during the follow up a total of 844 (16.2%) CVD patients died which 216 (13.3%) was young-aged, 201 (13.7%) a middle-aged and 427 (20.2%) an old-aged. The survivorship curves are plotted using Kaplan Meier method and identified some risk factors using univariate and multivariate Cox’s model. The risk factors were identified as, namely, age, sex, frequency of hospital visits, CVD subtype, number of diagnosis, ward, hypertension, diabetes, myocardial infarction, and cerebral infarction. Further, risk factors for each age group are illustrated at the significant p-value <0.05.

1. Introduction

Cardiovascular Diseases (CVDs) are the number one cause of death globally and to the less economically developed country like Bhutan. Every year, more number of people die of cardiovascular disease than from any other cause. Cardiovascular Disease (CVD) is the condition that affects the functioning of heart organ and flow of blood in the circulatory system [1]. The competitive risk factors like diabetes, obesity, and the metabolic syndrome contributes to the ascending trend of CVD patients and its mortality. Thus, in the face of public health and individual’s economy, it is a huge burden requiring a renewed approach towards prevention of risk factors [2-4].

Survival models in finding risk factors for cardiovascular patients would be effective in lengthening mortality, treatment and lifestyle modification for CVD patients [5-7]. Various studies conducted to frame a risk factors model for short and long terms risk [8]. The association between CVD sub-types and age groups are significant [9-11]. Kaplan Meier and Cox’s models have widely used to study the relation between CVD’s patients and associated risk factors [12-14].

The aim of this study is to compare the survival function for each age group within covariates and examine the relationship between survival time and risk factors by Cox’s proportional hazard models.
2. Methods

2.1. Study Participants
The JDWNRH is the medical center of Bhutan keeping the update of the statistical data available regarding patients suffering from any disease. Country is yet to cope with technology and trend regarding the system of automatic (computerization) like other developed nations. It’s low financial and economic background keeps Bhutan in peace with neither hunger of it nor craving under the philosophy of GNH (Gross National Happiness). Thus, country’s data is statistically in-qualitative as most of the variables/factors are lacking in the package. However, in this registry, 1,618 were registered as young-aged (<45 years) patients, 1,471 as middle-aged (45-65 years) and 2,117 as an old-aged (>65 years) from January 2006 to December 2016. None of the patients were excluded from the study.

2.2. Risk factors
To identify the risk factors after onset of CVD which further affects the survival time, we considered some conventional variables available in the data, namely, age, sex, frequency of hospital visits, CVD subtypes, number of diagnosis, ward, hypertension, diabetes mellitus, myocardial infarction and cerebral infarction. Some other risk factors like, HDL and LDL cholesterol, BMI, current smoking, and physical activity were absent in the data. Its absence has weighted the data low. However, available variables were used to find its significance to the survival time/time to event. Age of patients were categorized into <45 years as young aged/less risk patient, 45-65 years as middle aged/moderate risk patients, and >65 years as an old aged/high risk patient. The cardiovascular disease is generally considered as an old age disease. The CVD subtype, hypertension, diabetes mellitus, myocardial infarction and cerebral infarction were extracted from the diagnosis 1, 2, 3, and 4 of the data based on ICD-10 coding for CVD.

2.3 Statistical Analysis
All the participants (n = 5,206) were used throughout the data analysis and interpretation. Kaplan-Meier (KM) method was used to estimate the cumulative distribution function of each risk factor, and also compared and plotted the survivorship within each level of all risk factors. The KM method remained helpful in finding risk factors which affect the survival time of CVD patients.

The KM estimator of the survivorship function at time t is obtained by the equation below

\[
\hat{S}(t_j) = \frac{\text{number of surviving past } t_j}{\text{number of subject in the study}},
\]

where \( t_j \) are ordered failure times for \( j=0, 1, 2, \ldots, k \); where \( k \) denotes the number of distinct times at which some subject failed.

A non-parametric test (Log-rank test) is used for comparing the survival time between different groups of variables and risk factors. It is very much powerful for non-overlapping survival curves. When survival curves are compared, we use the following testing hypothesis:

\[ H_0 : \text{No difference between survival curves.} \]
\[ H_1 : \text{Difference between survival curves.} \]

The most popular testing method is called the Log-rank test. The Log-rank statistic for two groups are given by

\[
\frac{(O_i - E_i)^2}{\text{Var}(O_i - E_i)}; i = 1,2,\ldots,k,
\]

where

\[ (O_i - E_i) = \sum_j (m_y - e_y), \]
and $m_{ij}$ is the observed counts for the $i^{th}$ group at the $j^{th}$ ordered failure time, $e_{ij}$ is the expected counts for the $i^{th}$ group at the $j^{th}$ ordered failure time.

The univariate and multivariate Cox regression allowed us to analyse the effects on survival of several prognostic variables in a survival model. It assumed the effects of different prognostic variables on survival as constant over time. The Cox’s PH model is a special type of regression model, which can be used to analyse data that contains censored observations. The hazard function formula for the Cox’s PH model is written as

$$h(t, X) = h_0(t) \exp(\sum_{i=1}^{p} \beta_i x_i),$$

where $X$ is a set of covariates $(x_1, x_2, ..., x_p)$,

$h_0(t)$ is the baseline hazard function that is a function of $t$, but does not involve the $x$’s

$\beta_i$ is the coefficient of $x_i$.

All statistical analysis was performed with SPSS Version 18.

3. Results

The cumulative survival functions of all CVD patients are considered and is shown in Figure 1. The cumulative survival of all CVD patients were more than 80%. It indicates that patients are going to survive some more years if the death is not caused due to other factors. The CVD patient’s data were then used to plot cumulative survival functions and compare survival functions for each age group covariates. Referring to Figure 2, it shows that patients less than or equal to 45 years have better survival prognosis than other two middle and an old aged groups. Though middle aged groups survival function is comparatively less significant with young aged group, the curves do not overlap. The log-rank statistics and corresponding p-value on Table 1, indicate that the three categories of patients differ to each other in survival function.

| Test statistic for equality of survival distribution. |
|-------------------------------------------------------|
| Covariate                                             | Log-rank statistic | d.f. | p-value   |
| Age group for CVD patients                            | 40.003**           | 2    | 0.000     |

**significant at p-value <0.01
Some patients had multiple diagnosis, up to four diagnosis. Some patients had diabetes mellitus and hypertension, where we considered hypertension and diabetes mellitus as another separate covariate. Their associated diagnostics was supposed to affect the mortality rate of patients and evaluated in Tables 3, 4 and 5 for all covariates within age groups.

Table 2 Baseline characteristics of participants grouped in three age categories: number (%)

| Gender       | Less than 45 yrs (<45) | 45-65 yrs | More than 65 yrs (>65) |
|--------------|------------------------|-----------|------------------------|
| Female       | 797 (49.3)             | 625 (42.5) | 958 (45.3)             |
| Frequency Visit |
| 1 Time       | 1429 (88.3)            | 1339 (91.0) | 1907 (90.1)            |
| 2 Times      | 153 (9.5)              | 104 (7.1)  | 168 (7.9)              |
| 3 Times      | 36 (2.2)               | 28 (1.9)   | 42 (1.9)               |
| CVD Subtype  |
| ARF          | 73 (4.5)               | 21 (1.4)   | 8 (0.4)                |
| CRHD         | 188 (11.6)             | 69 (4.7)   | 31 (1.5)               |
| HpD          | 177 (10.9)             | 258 (17.5) | 447 (21.1)             |
| IHD          | 63 (3.9)               | 85 (5.8)   | 173 (8.2)              |
| PHD & DPC    | 73 (4.5)               | 36 (2.4)   | 55 (2.6)               |
| Other HD     | 241 (14.9)             | 229 (15.6) | 616 (29.1)             |
| CrD          | 208 (12.8)             | 337 (22.9) | 579 (27.4)             |
| DAAC         | 22 (1.4)               | 16 (1.1)   | 22 (1.0)               |
| DVLL         | 548 (33.9)             | 415 (28.2) | 172 (8.1)              |
| Others       | 25 (1.5)               | 5 (0.3)    | 14 (0.7)               |
| No. of Diagnostics |
| 1 Disease    | 1128 (69.7)            | 849 (57.7) | 994 (46.9)             |
| 2 Diseases   | 382 (23.6)             | 451 (30.7) | 721 (34.1)             |
| 3 Disease    | 90 (5.6)               | 134 (9.1)  | 295 (13.9)             |
| 4 Disease    | 18 (1.1)               | 37 (2.5)   | 107 (5.1)              |
| Ward         |
| Medical      | 494 (30.5)             | 640 (43.5) | 1141 (53.9)            |
| Surgical     | 478 (29.5)             | 411 (27.9) | 257 (12.1)             |
| Pvt Cabin    | 79 (4.9)               | 147 (9.9)  | 355 (16.8)             |
| Others       | 567 (35.0)             | 273 (18.6) | 364 (17.2)             |
| Hypertension | 44 (2.7)               | 120 (8.2)  | 233 (11.0)             |
| Diabetes     | 18 (1.1)               | 47 (3.2)   | 92 (4.3)               |
| Myocardial Infarction | 26 (1.6) | 47 (3.2) | 77 (3.6) |
| Cerebral Infarction | 54 (3.3) | 97 (6.6) | 210 (9.9) |

Note: ARF- Acute rheumatic fever, CRHD-chronic rheumatic heart disease, HpD-hypertensive disease, IHD-Ischemic heart disease, PHD & DPC-pulmonary heart disease and disease of pulmonary circulation, Other HD- other forms of heart disease, CrD-cerebrovascular disease, DAAC- disease of arteries, arterioles and capillaries, DVLL-disease of veins, lymphatic vessels and lymph nodes, others- other unspecified disorders of circulatory system.

Therefore, we divided the data into three groups as follows: young aged CVD patients, middle aged CVD patients and old aged CVD patients. In addition, we used type and risk factor in clinical data, and all covariates in personal data to plot and compare survival functions for each group. Each covariates association and comparison of Log-rank test with its p-value are given in Tables 3, 4 and 5.

Table 2 shows the baseline characteristics of participants grouped into three age categories. 49.3%, 42.5% and 45.3% were female patients of young aged, middle aged and old aged group respectively. Patients visiting hospital ranged to the maximum of nine times and for the convenience, it was grouped into one time, two times and more than three times. Some patients experienced the event during nine times of visit, some during one time only and some still survived even after nine times visit. The visit frequency to the hospital for young patients found to be 88.3%, 9.5% and 2.2%, respectively. For middle aged patients, 91%, 7.1% and 1.9%, respectively. An old-aged patient’s visit was 90.1%, 7.9% and 1.9%, respectively. Above figures depicts that the frequency of CVD patients visiting hospital one time is more and concludes that patients have less complications. Thus, more the frequency of visit, more the complications of the disease.

CVD is not just one disease; it is anything related to the defect of circulatory system. Therefore, the disease is grouped into 10 according to the ICD-10-CM diagnosis codes. After onset of disease, some patients carry multiple diagnosis, up to four diagnosis. Some patients had diabetes mellitus and hypertension, where we considered hypertension and diabetes mellitus as another separate covariate.
Mean age of total participants was 57 years and 41% were an old age. Some prognostic variables were associated to the survival time of all three categories of age. However, gender remained statistically insignificant for all groups of age. The follow up duration of the study was 11 years from January 2006 to December 2016. During that period, 844 events occurred, of which 216 were young-aged, 201 middle aged and 427 an old-aged participant. Tables 3, 4 and 5 shows the log-rank test of three groups of participants respectively.

**Table 3** Test statistics for equality of survival distributions for young-aged CVD patients.

| Covariate                  | Log-rank statistics | d.f. | p-value |
|----------------------------|---------------------|------|---------|
| Gender                     | 0.002               | 1    | 0.963   |
| Frequency of visit         | 4.013               | 2    | 0.134   |
| CVD sub-types              | 284.927**           | 9    | 0.000   |
| No. of diagnostics         | 118.162**           | 3    | 0.000   |
| Ward                       | 192.510**           | 3    | 0.000   |
| Hypertension               | 2.281               | 1    | 0.131   |
| Diabetes                   | 0.075               | 1    | 0.784   |
| Myocardial infarction      | 12.726**            | 1    | 0.000   |
| Cerebral infarction        | 4.442*              | 1    | 0.035   |

*significant at p-value <0.05, **significant at p-value <0.01.

Young aged CVD patients of Bhutan have a strong association with CVD sub-types, number of diagnostics, ward, myocardial infarction and cerebral infarction covariates (p-value <0.05). The gender is not significant with p-value 0.963, which does not have association at all. It means, it has a probability error of 96.3%, where male and female does not have different survival time. In addition, frequency of visit, hypertension and diabetes does not have association with young-aged CVD patients as shown in Table 3 as their p-value is greater than 0.05.

**Table 4** Test statistics for equality of survival distributions for middle-aged CVD patients.

| Covariate                  | Log-rank statistics | d.f. | p-value |
|----------------------------|---------------------|------|---------|
| Gender                     | 1.724               | 1    | 0.189   |
| Frequency of visit         | 10.244**            | 2    | 0.006   |
| CVD sub-types              | 161.223**           | 9    | 0.000   |
| No. of diagnostics         | 41.879**            | 3    | 0.000   |
| Ward                       | 152.686**           | 3    | 0.000   |
| Hypertension               | 9.805**             | 1    | 0.002   |
| Diabetes                   | 0.450               | 1    | 0.502   |
| Myocardial infarction      | 0.558               | 1    | 0.455   |
| Cerebral infarction        | 7.375**             | 1    | 0.007   |

**significant at p-value <0.01.

Regarding the middle-aged patients (45-65 years), covariates like frequency of visit, CVD sub-types, number of diagnostics, ward, hypertension and cerebral infarction have p-value less than 0.05. It indicates that the middle-aged patient’s survival time will depend on those covariates. Some variables like, gender, diabetes and myocardial infarction does not make a difference in survival time as their significant level is more than 0.05 as shown in Table 4.
Table 5 Test statistics for equality of survival distributions for old-aged CVD patients.

| Covariate                | Log-rank statistics | d.f. | p-value |
|--------------------------|---------------------|------|---------|
| Gender                   | 2.280               | 1    | 0.131   |
| Frequency of visit       | 0.002               | 2    | 0.999   |
| CVD sub-types            | 126.554**           | 9    | 0.000   |
| No. of diagnostics       | 18.964**            | 3    | 0.000   |
| Ward                     | 60.921**            | 3    | 0.000   |
| Hypertension             | 0.779               | 1    | 0.377   |
| Diabetes                 | 0.538               | 1    | 0.463   |
| Myocardial infarction    | 4.522*              | 1    | 0.033   |
| Cerebral infarction      | 3.473               | 1    | 0.062   |

*significant at p-value <0.05, **significant at p-value <0.01.

Table 5 shows the test statistics for equality for survival distributions for old-aged CVD patients. Where CVD sub-types, number of diagnostics, ward and myocardial infarction associates and depicts that, an old-age CVD patient's survival time would depend on those factors. Some covariates like gender, frequency of visit, hypertension, diabetes, and cerebral infarction does not have relation with the old-age CVD patients. It’s survival time wont depend on those variables.

Thus, some of the considered risk factors have shown an association with test statistics at significance level <0.05. The risk of mortality increases significantly with age. CVD sub-types, number of diagnostics and ward were statistically significant for all groups of age and they are identified to be the significant risk factors for CVD patients. When survivor functions are plotted using KM method, every plot showed no overlapping and each groups are different in all categorical variables. Moreover, the overall hazard ratio of the all-categorical variables grouped into three age groups are summarized in Table 6.

All the covariates were entered into the Cox’s proportional hazard model. The table shows the general hazard ratio of the all covariates in three groups. The covariates which are not related to survival time at significance level 0.05 are gender, number of diagnostics and cerebral infarction. Hazard ratio explains the comparison of effect of each covariate on survival time of patients.

The subjects who have visited 2 times and more than 3 times have hazard ratio 1.02 and 1.04 for young aged patients, 0.49 and 0.47 for middle aged patients with 1.03 and 1.09 for old aged patients respectively. Hazard ratio for young aged and old aged patients is greater than subjects who visited hospital one time. So chances of death for 2 times visit and 3 times visit for young aged patients increases at a rate of 1.02 and 1.04 times respectively and old aged patients follows the same. In the similar manner, some CVD sub types, number of diagnostics, ward, hypertension, diabetes, myocardial infarction and cerebral infarction had hazard ratio greater than reference (value 1). A young aged patient who has diabetes are associated with increased death rate of 2.28.

4. Discussion
We use Kaplan-Meier model to plot the cumulative survival curves of each covariate and test the difference of cumulative survival curves within each covariate by using log-rank statistics. We found that younger CVD patients have better survival prognosis than other two middle and old aged patients. For each group of patients, the gender covariates do not have difference in survival time.

For univariate Cox’s proportional hazard model, we use one covariate for one model to find the association. Next, we use all covariates which we get from the univariate Cox’s proportional hazard model in multivariate Cox’s proportional hazard model.
A participant who lays in detection of the disease. Thus, the survival time due to inaccuracy or determinants used variables like marital status, BMI, hypertension, age, myocardial infarction and cerebral infarction showed differences in survival commonly to young patients.

Further, study could be done with more detail with more varied variables like cholesterol, marital status, BMI, current smoking status.

The risk factors that affect the survival time are the frequency of hospital visits, CVD subtypes, number of diagnosis, ward, hypertension, myocardial infarction and cerebral infarction. A participant who visited hospital not so frequently has better survival prognosis than those who visited more than one time. It depicts the severity of the disease or need of better treatment. Other unidentified cardiovascular disease has less survival time. It depicts the severity of the disease or need of better treatment. Other unidentified cardiovascular disease has less survival time due to inaccuracy or delay in detection of the disease. Thus, further delaying treatment.

A patient with more number of disease has poor prognosis. Some participants who were treated in the surgical ward showed a better prognosis in survival as immediate surgery saves lives. Participants treated in other wards like ears and nose (ENT) ward, psychiatric ward, orthopaedic ward etc., showed poor survival time. Hypertension affected the survival time of only middle-aged patients. Myocardial infarction and cerebral infarction showed difference in survival time commonly to young patients.

Further, study could be done with more exclusively with more detailed variables like cholesterol, marital status, BMI, current smoking status.

### References

[1] Steinbaum RS, (2017, July 25). *Heart Disease Home*. Retrieved from Heart and Cardiovascular

Diseases: http://www.webmd.com/heart-disease/guide/diseases-cardiovascular#1

[2] Hata J, et al. (2016). Risk prediction models for mortality in patients with cardiovascular disease: The biobank japan project. Journal of epidemiology, s71-s76.

[3] Pearson A T, et al. (2002). Consensus Panel Guide to Comprehensive Risk Reduction for Adult Patients Without Coronary or Other Atherosclerotic Vascular Diseases. AHA Scientific Statement, 388-399.

[4] Yusuf S, et al. (2001). Part II: Variations in Cardiovascular Disease by Specific Ethnic Groups and Geographic Regions and Prevention Strategies. Clinical Cardiology: New Frontiers. 2855-2864.

[5] Freedman S D, et al. (1998). The Relation of Overweight to Cardiovascular Risk Factors Among Children and Adolescents. The Bogalusa Heart Study. 1175-1182.

[6] Berenson S G, et al. (1998). Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The New England Journal of Medicine. 1650-1656.

[7] Kannel B W, et al. (1979). Diabetes and Glucose Tolerance as Risk Factors for Cardiovascular Disease. The Framingham Study. 120-126.

[8] Turner CR et al. (1998). Risk factors for coronary artery disease in non-insulin dependent diabetes mellitus: United Kingdom Prospective diabetes study (UKPDS:23). British Medical Journal (BMJ), 823-828.

[9] Conroy MR et al. (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. European Heart Journal, 987-1003.

[10] Lloyd-Jones MD, et al. (2006). Prediction of Lifetime Risk for Cardiovascular Disease by Risk factor Burden at 50 Years of Age. Epidemiology, 791-798.

[11] Soares PG, et al. (2016). Progression of Mortality due to Diseases of the Circulatory System and Human Development Index in Rio de Janeiro Municipalities. Original Article, Epidemiology, 314-322.

[12] Wilson WFP et al. (1998). Prediction of coronary heart disease using risk factor categories. Circulation: Special report, 1837-1847.

[13] Blacher J et al. (2001). Artrial Calcifications, Arterial Stiffness, and Cardiovascular Risk in End-Stage Renal Disease. Hypertension, 938-942.

[14] Assmann G et al. (2002). Simple Scoring Scheme for Calculating the Risk of Acute Coronary Events Based on the 10_year Follow-Up of the Prospective Cardiovascular Munster (PROCAM) Study. Circulation, 310-315.

Acknowledgements

We would like to thank Thailand International Cooperation Agency (TICA) for the scholarship and Record Section of Jigme Dorji Wangchuck National Referral Hospital (JDWNRH) in Bhutan for providing and supporting with possible required CVDs patient’s data.