Comparative studies for Cox hazards model based on the population-based cohort study of Japan

Michikazu Nakai[1], Y.L. Lio[2], Ding-Geng (Din) Chen[3]
Kunihiro Nishimura[1], Makoto Watanabe[1], Yoshihiro Miyamoto[1]

[1]National Cerebral and Cardiovascular Center, 5-7-1 Fujishir-dai Suita-city, Osaka, 565-8565, Japan
[2]University of South Dakota, 414 East Clark, Vermillion, SD 57069, USA
[3]Center for Research, School of Nursing & Department of Biostatistics and Computational Biology, University of Rochester Medical Center, 255 Crittenden Blvd, Box SON Rochester, NY 14642, USA

Abstract
In cohort study, there is always a time-gap between the date of occurrence and the date of diagnosis when the endpoint is a lifestyle-related disease such as hypertension (HT). However, in survival analysis, the date of diagnosis is typically used. In this paper, we investigate whether these time-gaps could cause a significant bias in the analysis result of the Suita study, the population-based prospective cohort study of Japan. Participants who had no HT at baseline (1,591 men and 1,973 women) aged 30-84 years were included. During the mean follow-up of 7.2 years, 1,325 participants (640 men and 685 women) developed HT. Missing data were created at the date of HT occurrence. Then, we compared the efficiency of median imputation and multiple imputation (MI) along with left and right-censoring approaches. The Cox proportional hazards model was used to estimate hazard ratios and C-index of waist-circumference and body mass index by sex adjusted for age, cigarette smoking and alcohol drinking. Right-censoring and MI estimated equally well compared to others regardless of indicators and sex.

Key Words: imputation, interval censor, Cox hazard model, cohort study, epidemiology

1. Introduction

In epidemiology, follow-up study is one of very informative resources to evaluate some hypotheses about human health and acknowledge the outcome. Most endpoints of the cohort study are to study the diseases like stroke or coronary heart disease. Recently, life-related diseases such as hypertension (HT) are featured people's interest, and habitual health check-up is important to lead to early detection. It is better to be able to detect the illness before being affected by disease.

In research aspect, when the life-related diseases are developed for one of participants, we may only know that the participant was affected between the two diagnosis point, the previous date of diagnosis and the date of diagnosis. Even though the exact time of being affected can be unknown, date of diagnosis is commonly used for statistical analysis.

This paper investigates whether time-gaps between the date of HT occurrence and the date of diagnosis have caused a significant bias in analysis result. If it is true, how much time gaps may be a limitation for a regular health checkup.

2. Subjects and Methods

2.1 Subjects

Suita cohort study which is a prospective population-based cohort study of an urban area in Japan was established in 1989. These participants were randomly selected from the municipality population based on a random sampling from 12,200 Japanese residents living in Suita-city in Japan. The participants have continued to visit the National Cerebral and Cardiovascular Center every two years for a regular health checkup.

Participants without HT at baseline (1989-1994) who visited at least twice were chosen in this study. The age range was between 30 and 84 years old. The total participants in this analysis were 3,564 (1,591 men and 1,973 women). During the mean follow-up of 7.2 years, 1,325 (640 men and 685 women) developed HT. The Cox proportional hazards model was used to estimate hazard ratios (HRs) and C-index of waist-circumference (WC) and Body Mass Index (BMI) by sex adjusted for age, cigarette smoking and alcohol drinking.

In this paper, HT was defined as SBP≧140 mmHg ,DBP≧90 mmHg or current use of antihypertensive medication. For an analysis purpose, missing value was created in the date of diagnosis when the participant developed HT. If the participant didn't develop HT during observed period, the last observed time point was left as non-missing. Since the interval of a regular health checkup has two years, the analysis included the case of two-year as well as four-year. The four-year case has obviously decreased the sample size 2,049 (924 men and 1,125 women).

SAS9.2 (SAS Institute Inc., Cary, NJ, USA) and STATA (College Station, TX, USA) were used for the analysis.

2.2 Imputation Methods

In this paper, we compare four techniques, left-censoring, median imputation, multiple imputation and right-censoring. Let the X_i be the time of HT developed for the participant $i, i = 1, \cdots, n$. Assume X_i is interval censored where $X_i \in [L_i, R_i]$. Left-censoring approach (Left) can be applied L_i point in place of missing value, which is the previous date of diagnosis. Since left-censoring can be placed before HT developed, this approach may be including some bias regard of time-gap. Median imputation (Median) is used a middle point between two diagnosis date where $\frac{L_i+R_i}{2}$. This approach is the representative approach for a single imputation method in censoring problem to compare [3,8]. Also, the algorithm of Multiple imputation (MI) is designed to calculate the posterior distribution of the parameters of interest and iterates between imputation and posterior steps [2,7]. The detail of this method is described elsewhere [1,5]. The idea behind this method is exactly that of Rubin's MI [6]. In this analysis, since MI is required to have at least two points to impute, when the participants developed HT at second diagnosis, it is ignored and counted as non-missing value. At last, right-censoring can be applied R_i point in place of missing value, which is the date of diagnosis. Since the exact date X_i is unknown, most analysis use this point as the time of HT developed. This approach could be referenced for other approaches.

3. Results

Table 1 shows that data characteristics in Suita Study. T-test and chi-squared test are used for continuous and categorical variables, respectively. All variables are significantly different between men and women. Table 2 explains the percentage of missingness for two-year and four-year. Since we ignore the HT occurrence at the second diagnose, in two-year, both sex includes about 20% missingness whereas, in four-year, missingness is about 30% which is maximum amount of missingness to use as effective imputation [4]. Tables 3 and 4 indicate the result in two-year and four-year for WC and BMI with four imputation approaches by sex. Examining Tables 3a and 4a, there are no huge differences in HR among four imputation approaches. However, C-index in Tables 3b and 4b shows that Left and Median have lower value than Right and MI, which explains that two imputations include some biases. In term of medical concept, HT developed more close to right-censoring point than middle point. Also, since HR of risk indicators (WC and BMI) have similar value in all approaches, the bias could be considered other variable besides adjusted factors. Furthermore, as time-interval gets longer, Left and Median were lower value than others (Table 4b).

4. Discussion

In this paper, we investigated whether a time-gap between the date of occurrence and the date of diagnosis has caused significant biases in the analysis when the disease progresses with time such as a lifestyle-related disease. We used the Suita cohort study, which is the population –based prospective cohort study in Japan.

From the result, there was no significant difference between right-censoring and MI, indicating that MI could be replaced if right-censoring would include more missing values. And, both approaches may be applied to a medical examination in every two year (four year) reasonably to predict incident HT in statistical aspect.

HR in Median imputation was similar with right-censoring. However, C-index showed less accuracy than right-censoring or MI, which was due to some bias in Median imputation. Simulation studies in [8] showed that bias could be resulted when the time-gaps are more than two years even though it doesn’t consider the endpoint of its disease. In our paper, bias included even for two-year interval of developing HT.

HR in left-censoring was similar with other imputation. However, C-index was lowest among four approaches regardless of indicators/sex. Thus, it is risky to impute left-censoring in place of missing value. However, even in left-censoring, more than 60% holds to predict HT. This result in statistical aspect explains that it may be possible to prevent HT in medical examination ahead of time when the participant is about to be pre-hypertension (SBP is between 120 mmHg and 139 mmHg or DBP is between 80 mmHg and 89 mmHg).

References

1. Bebchuk,J.D. and Betensky, R.A., 2000. Multiple imputation for simple estimation of the hazard function based on interval censored data, Statistics in medicine, 19,405-419

2. Chen, DG., Sun,J. and Peace, K.E, 2012, Interval-censored time to event data- methods and application, Chapman & Hall/CRC Biostatistics Series

3. Dorey, F.J., Little, R.J.A. and Schenker N., 1993. Multiple imputation for threshold-crossing data with interval censoring, Statistics in medicine, 12, 1589-1603

4. Little,.R.J.A. and Rubin,.D.B., 2002, Statistical analysis with missing data (2nd Edition), New York, Wiley

5. Pan,W., 2001. A multiple imputation approach to regression analysis for doubly censored data with application to AIDS studies, Biometric, 57,1245-1250

6. Rubin, D.B.,1987. Multiple imputation for nonresponse in surveys. New York, Wiley

7. Sun,J.,2006. The statistical analysis of interval censored failure time data, Springer

8. Zhang, W., Zhang, Y., Chaloner, K., and Stapleton, J.T., 2009. Imputation methods for doubly censored HIV data, J Stat Comput Simul, 79(10), 1245-1257

Table

Table1: Baseline Data Characteristics

Variable	Men	Women	P-value
Sample size	1,591	1,973	
WC (cm)	82±8	76±10	<0.01
BMI (kg/m^2)	22.6±2.8	21.7±3.0	<0.01
SBP (mmHg)	118±12	114±13	<0.01
DBP (mmHg)	74±9	71±9	<0.01
Current smoker (%)	52.5	13.3	<0.01
Current drinker (%)	73.8	34.5	<0.01

Table 2: Missing values in two-year and four-year dataset

	Variable	Men	Women
Two year	Sample size	1,591	1,973
	HT occurrence n (%)	640 (40.23)	685 (34.72)
	HT occurrence at second diagnose	256 (16.09)	263 (13.33)
Four year	Sample size	924	1,125
	HT occurrence n (%)	452 (48.92)	484 (43.02)
	HT occurrence at second diagnose	137 (14.83)	149(13.24)

Table 3a: HR in two years

		Left	Median	MI	Right
Men	Waist	1.02	1.03	1.02	1.03
	BMI	1.08	1.08	1.08	1.08
Women	Waist	1.02	1.02	1.03	1.03
	BMI	1.09	1.09	1.10	1.10

Table 3b: C-index (95% Confidence Interval) in two-year

		Left	Median	MI	Right
Men	Waist	0.75(0.65,0.85)	0.74(0.63,0.84)	0.84(0.75,0.92)	0.84(0.75,0.92)
	BMI	0.75(0.64,0.85)	0.74(0.63,0.83)	0.84(0.74,0.92)	0.84(0.74,0.92)
Women	Waist	0.65(0.55,0.74)	0.66(0.57,0.76)	0.74(0.64,0.82)	0.73(0.64,0.82)
	BMI	0.65(0.55,0.74)	0.66(0.56,0.75)	0.73(0.64,0.82)	0.73(0.64,0.82)

Table 4a: HR in four-year

		Left	Median	MI	Right
Men	Waist	1.02	1.02	1.02	1.02
	BMI	1.05	1.05	1.05	1.05
Women	Waist	1.02	1.02	1.02	1.02
	BMI	1.07	1.05	1.07	1.07

Table 4b: C-index (95% Confidence Interval) in four-year

		Left	Median	MI	Right
Men	Waist	0.70(0.56,0.82)	0.73(0.60,0.84)	0084(0.73,0.93)	0.84(0.73,0.93)
	BMI	0.70(0.56,0.82)	0.73(0.60,0.84)	0.84(0.74,0.93)	0.84(0.73,0.93)
Women	Waist	0.65(0.53,0.76)	0.69(0.57,0.80)	0.76(0.65,0.85)	0.76(0.65,0.86)
	BMI	0.65(0.53,0.76)	0.69(0.58,0.80)	0.76(0.65,0.86)	0.76(0.65,0.86)