Survival Analysis: Impact of Marital Bereavement on Mortality among the Elderly

Yimin Chen (yc4195), Qingyue Zhuo (qz2493), Jiarui Yu (jy3360)

2023-12-11

Contents

Abstract

This study addresses the critical issue of understanding how marital bereavement affects mortality risks in the elderly. Previous studies have suggested that losing a spouse can significantly impact an individual's health, but quantifying this effect remains challenging. The primary objective of this project is to assess the impact of marital bereavement on the risk of death among the elderly, with a focus on how this risk is modified by demographic and health-related factors. We applied non-parametric and semi-parametric methods in the study. Our result shows that the change of bereavement status doesn't have a significant effect on mortality rate.

Background and Objectives

The aging process is accompanied by various life events that can significantly impact health outcomes. The loss of a spouse stands out as a particularly traumatic experience with potential long-term effects on the survivor's well-being. Marital bereavement has been associated with increased mortality rates, commonly referred to as the "widowhood effect." Understanding the dynamics of how marital bereavement influences mortality is critical for public health and for developing interventions aimed at supporting the bereaved elderly population.

The primary objective of this research is to quantify the impact of marital bereavement on all-cause mortality among the elderly. While previous studies have established a link between spousal loss and increased mortality risk, many questions remain regarding the mechanisms at play and the moderating role of demographic and health-related factors. This study seeks to examine the extent of these influences and captures relevant variables. Given the aging global population and the associated increase in the incidence of bereavement among older adults, this research holds significant value for health policy and geriatric care.

Methods

Study data

The study utilizes data from the Bereavement dataset (BRV) from the "biostat3" package in R. It includes 399 elderly individuals who were part of a larger longitudinal study. This dataset focuses on the elderly whose spouses were alive at the onset of the study, allowing for an in-depth analysis of the effects of marital bereavement on mortality.

The dataset includes a wide range of variables, including individual ID (id), couple identifiers (couple), dates of birth (dob), entry into the study (doe), exit from the study (dox), death of the spouse (dosp), and the mortality of the study subjects (fail). The primary outcome of the study is the time to event (death), with the event status indicated by the 'fail' variable, where 0 represents subjects who were alive at the end of the study (censored), and 1 indicates subjects who died during the study period (uncensored).

The data features right-censoring where the event of interest (death) has not occurred for all subjects by the end of the study period. The 'dosp' variable provides information about the date of the spouse's death; it is coded with a specific date if the spouse died during the study period or a uniform date (January 1, 2000) if the spouse was still alive at the study's conclusion.

Table 1: Summary Statistics for Date Variables

Variable	Description	Minimum Date	Maximum Date
dob	Date of Birth	1888-02-22	1906-03-12
doe	Date of Entry into Study	1981-01-15	1981-10-23
dox	Date of Exit from Study	1981-03-13	1991-01-01
dosp	Date of Death of Spouse	1981-05-22	2000-01-01

The study population primarily consists of elderly couples categorized into three distinct groups based on their age and living arrangements at the time of the initial interview. Group 1 comprises 226 index cases, all of whom were 75 years or older and lived in a two-person household. Group 2 includes 118 index cases with spouses younger than 75 years, also living in two-person households. Lastly, Group 3 contains 55 index cases who resided with a spouse and another person. The baseline characteristics of the sample include demographic information (age and sex), health status, and level of disability.

Table 2: Summary Statistics for Factor Variables

Variable	Description	Factor Top Counts	Meaning
group	Group	1: 597, 2: 302, 3: 137	1 = group 1, 2 = group 2, 3 = group 3
disab	Disability Level	0: 758, 1: 131, 3: 81, 2: 66	0 = none, 1 = mild, 2 = mod, 3 = severe
health	Health Status	2: 634, 1: 349, 0: 53	0 = poor, 1 = mod, 2 = good
sex	Sex	1: 554, 2: 482	1 = male, 2 = female
brv	Bereavement Status	0: 735, 1: 301	1 = bereaved, 0 = not bereaved
fail	Status at Exit	cen: 758, 1: 278	cen = censored, 1 = death
age	Age Band	(80: 408, (75: 326, (85: 258, (90: 38	(70,75], (75,80], ..., (95,100]

##Data Processing

We performed a series of data transformations on the BRV dataset. Firstly, we created time variables `y_before_sp_dth` and `y_after_sp_dth` relative to the spouse's death. These represent the number of years before and after the spouse's death, calculated from the date of entry into the study (`doe`) and the date of the spouse's death (`dosp`), as well as the date of exit from the study (`dox`). Then, we split each subject's follow-up time into two records: one for the time before the spouse's death and another for the time after (if the spouse died during the study). The dataset is further modified to calculate the time each subject was at risk of the event (death) after the spouse's death. This is represented by the `t_sp_at_risk` variable. We created a new binary variable, `brv`, to indicate whether the subject was bereaved (1) or not (0) based on whether `y_after_sp_dth` is greater than 0. We also created additional time variables: `age_sp_dth` (age at spouse's death), `age_start` (age at the start of the observation period), and `age_end` (age at the end of the observation period). New variables `t_at_risk` (time at risk within each age band) and `age` (categorical age band) are created to reflect the duration and age category of each interval.

Table 3: Summary Statistics for Numeric Variables

Variable	Description	Mean	Median	Sd
y_before_sp_dth	Years Before Spouse's Death	-8.07	-5.58	7.89
y_after_sp_dth	Years After Spouse's Death	-2.72	0.00	7.13
t_sp_at_risk	Time at Risk After Spouse's Death	5.35	5.25	3.03
age_sp_dth	Age At Spouse's Death	87.71	85.80	7.26
age_start	Start Age	81.33	80.00	3.80
age_end	End Age	83.68	84.81	3.69
t_at_risk	Time at risk	2.35	2.06	1.61

By the end of this process, the dataset is transformed to facilitate survival analysis by accommodating the change in risk status due to bereavement and by considering the age of the subjects as they progress through different age bands. This detailed partitioning allows for a nuanced analysis of how the risk of death changes with age and bereavement status over time.

Statistical methods

Life Table

To investigate our research questions, we first applied the Life-table, which was constructed using standard life table methodology.

% Probability function at t_middle

$$t_{mi} = \frac{t_i + t_{i-1}}{2}$$

$$f(t_{mi}) = \frac{\hat{S}_L(t_{i-1}) - \hat{S}_L(t_i)}{t_i - t_{i-1}}$$

$$\hat{S}_L(t_{mi}) = \frac{\hat{S}_L(t_{i-1}) + \hat{S}_L(t_i)}{2}$$

% Number of events per person-time-units

$$\hat{h}(t_{mi}) = \frac{d_i}{I(t_i - t_{i-1})(n_i' - \frac{d_i}{2})}$$

% Based on the definition

$$\hat{h}(t_{mi}) = \frac{f(t_{mi})}{S(t_{mi})} = \frac{2f(t_{mi})}{S(t_i) + S(t_{i-1})}$$

% Variance

$$\text{var}\{\hat{h}(t_{mi})\} = \frac{(\hat{h}(t_{mi}))^2}{n_i' q_i}\left(1 - \frac{(\hat{h}(t_{mi})(t_i - t_{i-1}))^2}{2}\right)$$

Kaplan-Meier and Fleming-Harrington model

For nonparametric estimator, Kaplan-Meier(KM) model and Fleming-Harrington(FH) model were used to measure the fraction of subjects living for a certain amount of time after the change in bereavement status.

The Kaplan-Meier estimator

$$\hat{S}_K(t) = \begin{cases} 1 & \text{if } t < t_1 \\ \prod_{t_i \le t}[1 - \frac{d_i}{n_i}] & \text{if } t \ge t_1 \end{cases}$$

note: $d_i = \#\ of\ failure\ at\ time\ t_i$, $n_i = \#\ at\ risk\ at\ t_i^-$, $c_i = \#\ censored\ during\ the\ interval$ $[t_i, t_{i+1}]$

The Fleming-Harrington estimator

$$\hat{S}_F(t) = \begin{cases} 1 & \text{if } t < t_1 \\ \prod_{t_i \le t} exp[-\frac{d_i}{n_i}] & \text{if } t \ge t_1 \end{cases}$$

Log-rank Test

We then use a non-parametric log-rank test, which makes no distributional assumptions about our data. The log-rank test compares differences between expected and observed events at each event time point, k, to derive the following test statistics:

The likelihood function is defined as:

$$L = \sum_{i=1}^{k} (d_{0i} - e_{0i});$$

And the variance is given by:

$$\text{var}(L) = \sum_{i=1}^{k} \frac{n_{0i} n_{1i} d_i (n_i - d_i)}{n_i^2 (n_i - 1)};$$

A significant result from the log-rank test indicates a statistically significant difference between the two groups. Notably, the log-rank test does not allow us to adjust for covariates or prognostic variables.

Cox Proportional Hazard Model

As a further step in this analysis, we consider the Cox Proportional Hazard (Cox PH) model, which allows us to model the hazard ratio based on covariates to understand their impact on the survival function.

The Cox PH typically takes the form:

$$h(t \mid Z = z) = h_0(t) e^{\beta' z}$$

Results

Discussion

Limitation and Future Work

The primary limitation of our study pertains to the modest size of our dataset, comprising 399 observations. This smaller sample size introduces potential challenges, including Reduced Statistical Power, Increased Type II Error Risk, Unreliable Estimates, Limited Generalizability, and Instability in Model Fitting. These limitations may impact the robustness and generalizability of our findings. Additionally, another noteworthy limitation stems from the absence of a "treatment" variable in our dataset. Consequently, our analysis lacks insight into potential health interventions or treatments related to the observed outcomes. Instead, our focus centers on investigating the impact of bereavement on the life expectancy of the elderly. To address this limitation and enhance the scope of our study, we propose the inclusion of additional datasets capturing information on interventions relevant to bereavement, such as counseling, psychotherapy, or physical exercise provided to older individuals. This augmentation would enable us to assess the influence of these interventions on survival outcomes, thereby enriching the depth and applicability of our analysis.

Conlusion

References

[1] López, P., Rodríguez, A. C., Escapa, S. (2022). "Psychosocial effects of gentrification on elderly people in Barcelona from the perspective of bereavement." Emotion, Space and Society, 43, 100880.

[2] Matthys, O., Dierickx, S., Deliens, L., Lapeire, L., Hudson, P., Van Audenhove, C., De Vleminck, A., Cohen, J. (2023). "Is pre-bereavement collaboration between family caregivers and healthcare professionals associated with post-bereavement emotional well-being? A population-based survey." Patient Education and Counseling, 110, 107654.

[3] Denckla, C. A., Hahn, J., Cowden, R. G., Ho, S., Gao, K., Espinosa Dice, A. L., Jha, S. C., Kang, J. H., Shear, M. K. (2023). "Bereavement, Memorial Attendance, and Mental Health During the COVID-19 Pandemic: Longitudinal Results from the Nurses' Health Study." The American Journal of Geriatric Psychiatry, 31(12), 1045-1057.

Appendix