OPEN ACCESS
Journal of Multidisciplinary Applied Natural Science Vol. xx No. xx (xxxx) https://doi.org/10.47352/jmans.2774-3047.448
Check for updates
Research Article

Bivariate Survival Copula Analysis of Cataract Patients During Blindness: Cataract Cases at Undaan Eye Hospital, Surabaya

Jerry Dwi Trijoyo Purnomo*, Purhadi Purhadi, Shofi Andari, Wahyu Dwi Rahmawati, and Sri Harini

Received : December 15, 2025	Revised : April 3, 2026	Accepted : June 1, 2026	Online : June 28, 2026

Abstract
About 51% of all occurrences of blindness worldwide are caused by cataracts, which are the most common cause of blindness. According to the Rapid Assessment of Avoidable Blindness (RAAB) survey conducted in Indonesia between 2014 and 2016, East Java had the highest rate of blindness (4.4%), with cataracts accounting for 81.1% of cases. Approximately 10,000 cataract cases were reported at the Undaan Eye Hospital in Surabaya, one of the major eye hospitals in East Java, in 2023. This study aims to analyze the recovery time of patients undergoing cataract surgery on both eyes, considering the influencing factors. Data from bilateral cataract patients at the Undaan Eye Hospital in Surabaya between January 2023 and December 2024 were utilized. Gender, age, history of hypertension, heart disease, diabetes, gastric illness, stroke and cholesterol were among the characteristics that were examined. To determine the relationship between the recovery durations of both eyes following surgery, a bivariate survival analysis with Clayton and Gumbel copula functions was used for the analysis. The Gumbel copula was determined to be the better model for describing the recovery of bilateral cataract patients, with a QIC value of 1529.078 and a Kendall's Tau of 0.6636, indicating a moderate and positive dependency between the recovery periods of the right and left eye. The model estimate results showed a substantial correlation between recovery time and a few covariates, including age (hazard ratio of 0.516), history of hypertension (hazard ratio of 0.357), and history of diabetes (hazard ratio of 0.615). Compared to younger individuals, older patients recover more slowly. Furthermore, people with a history of diabetes and hypertension typically recover at a lower rate than those without these conditions.

Keywords: bilateral cataracts, bivariate survival analysis, Clayton's Copula, Gumbel's Copula

1. INTRODUCTION

One of the most common causes of blindness in the world, cataracts, are becoming more common as the world's population ages. The World Health Organization (WHO) estimates that 2.2 billion people worldwide suffer from blindness or visual impairment. At least 1 billion of these 2.2 billion people suffer from avoidable or treatable blindness or vision impairment. About 51% blindness globally is caused by cataracts, which have a major negative influence on the quality of life for people who are affected [1]. The results of the 2014–2016 Rapid Assessment of Avoidable Blindness (RAAB) survey, which was carried out in 15 provinces—including Java-Bali, Sumatera, Kalimantan, Sulawesi, Maluku, and Papua zones—represent 75% of Indonesia's total population. According to the RAAB data analyzed by the Health Research and Development Agency (Litbangkes), untreated cataracts accounted for over 70% of the country's 3% blindness prevalence. Based on the survey's findings, East Java was identified as the region that contributed the most to the blindness rate, with 4.4% of cases coming from people over 50.0 and 81.1% from cataract cases. In this sense, the capital of East Java Province, the City of Surabaya, offers medical facilities that can aid in cataract management and healing. One of these is the Undaan Eye Hospital in Surabaya, which treats and examines patients with eye conditions, such as cataracts, both independently and through referrals from clinics or community health centers.

Undaan Eye Hospital will have 10,000 cataract patients by 2023. As a result, they work hard to give patients the best care possible to aid in their recuperation. Surgery is currently the final treatment for cataracts. Vitamins C and E have been demonstrated in several studies to reduce the growth of cataracts, although they are not yet effective enough to treat them. The best treatment for cataracts at the moment is surgery. Only one eye is initially affected by surgical operations at the

Table 1. Research variables in this study.

Variable	Name	Description
T_1	Survival time	Time until the right eye was declared cured of cataracts by the doctor since surgery (in days)
T_2	Survival time	Time until the left eye was declared cured of cataracts by the doctor since surgery (in days)
d	Event status	0: other (censored); 1: *Event*
X_1	Gender	0: Male; 1: Female
X_2	Age	0: < 60 years; 1: ≥ 60 years
X_3	history of hypertension	0: No; 1: Yes
X_4	history of heart disease	0: No; 1: Yes
X_5	history of diabetes mellitus	0: No; 1: Yes
X_6	history of gastric illness	0: No; 1: Yes
X_7	history of stroke	0: No; 1: Yes
X_8	history of cholesterol	0: No; 1: Yes

Undaan Eye Hospital in Surabaya. To determine which eye will be operated on, observations are made prior to surgery. A thorough analytical approach is required to comprehend the pattern of disease progression and the factors influencing it, given the rising number of cataract cases. For an object of observation, survival analysis typically involves only one primary event [1]. However, in certain situations, such as in cases of chronic and bilateral disease, observations entail many events. In situations such as these, the timing of the occurrence of events in each organ associated with the same research subject can be examined using the survival analysis approach with bivariate survival regression. The survival time of an event that occurs twice from the same research subject but from distinct objects is one of the two dependent variables used in bivariate survival regression. This method connects the two dependent variables that may mutually impact the risk of the event occurring. According to Sun and Ding [2], there are three general methods for analyzing bivariate survival data: the copula model, the frailty model, and the marginal model method.

In order to connect two dependent variables in bivariate survival models, the copula function combines two marginal distribution functions into one while taking their dependencies into account. A copula model for bivariate survival data can be created by applying the copula function to a semiparametric survival model. This model is distinguished by a parametric model with a copula connection and a nonparametric model of two marginal survival functions [3]. One of the earliest distribution families with many copula functions is the Archimedean copula family. The Clayton, Gumbel, Frank, and Joe copula functions are the four commonly utilized copula functions [2].

Several previous studies have demonstrated the success of applying a copula in bivariate survival analysis in various medical cases. Zhang et al. [4] used the copula approach to analyze the time to recurrence in bilateral breast cancer patients. Suresh et al. [5] applied this method to analyze the survival time of kidney transplant patients. Huang, Zhou, and Ibrahimi [6] applied this method to the diabetic retinopathy study (DRS). Petti et. al. [7] used copula on age-related macular degeneration (AMD) patient data. He et al. [8] used the copula approach on tumor disease. Filho nand Demarqui [9] applied the survival copula method to ovarian cancer patient data. However, the application of this method in the context of eye diseases, especially cataracts, is still limited. Hong et. al. [10] use simple linear regression and correlation to compare the outcomes and complications of immediate sequential bilateral cataract surgery and unilateral cataract surgery in a tertiary hospital in South Korea. Although bilateral surgery was generally superior to unilateral surgery, this approach is

unable to capture the multivariate dependency between response variables. In this study, the bivariate survival model with copula functions was applied on data from bilateral cataract patients at the Undaan Eye Hospital, Surabaya. The main problem in this study was how to model the recovery time of patients after undergoing surgery by considering risk factors and the dependency between healing of the right and left eyes. The copula functions used included the Clayton and Gumbel copula functions.

2. MATERIALS AND METHODS

2.1. Materials

The data used in this study were secondary data from bilateral cataract patients who underwent eye surgery on their right and left cataracts and were declared cured. The data were obtained through the patients' medical record data management information system from January 2024 to December 2024 at the Undaan Eye Hospital, Surabaya. The raw data of these bilateral cataract patients who underwent surgery on both eyes at the Undaan Eye Hospital in Surabaya from January 2024 to December 2024 were 1,790 patients. Dependent and independent factors make up the research variables employed in this study. Survival time (T_1, T_2) and censored status (d) are the response variables employed. Patient status (d) shows whether or not an incident happened throughout the trial, whereas survival time (T_1, T_2) represents the amount of time between the cataract patient's right and left eye surgeries and their declaration of cure. Gender, age, history of hypertension, history of heart disease, history of diabetes mellitus, history of gastric illness, history of stroke, and history of cholesterol are among the independent variables utilized in this study that may have an impact on the recovery duration of cataracts. A description of the variables used in the study is explained in Table 1, and R studio was used to analyze this dataset.

2.2. Methods

2.2.1. Bivariate Survival Regression

Univariate survival analysis assumes that the time-to-event observations are independent each other. However, this assumption doesn't hold when the research unit involves two correlated events, such as twin pairs, married couples, bilateral diseases, or other paired cases [11]. Several notations are defined for bivariate survival data. Let (T_{1i}, T_{2i}) and (C_{1i}, C_{2i}) denote the time to the bivariate event and the right censoring time for the i -th subject, respectively, where i = 1, 2, ..., n. The censoring indicator is denoted by $\delta_i = (\delta_{1i}, \delta_{2i})$, and x_i represents the covariate vector for the i-th subject. To simplify the discussion, the observed data are denoted by D as follows:

$$D = \{(y_{1i}, y_{2i}, \delta_{1i}, \delta_{2i}, \boldsymbol{x}_i): i = 1, 2, \ldots, n\},$$

where $y_{1i} = \min(T_{1i}, C_{1i})$, $y_{2i} = \min(T_{2i}, C_{2i})$, $\delta_{1i} = I(y_{1i} \leq C_{1i})$, $\delta_{2i} = I(y_{2i} \leq C_{2i})$ and $x_i = (x^T_{1i}, x^T_{2i})^T$, also the bivariate survival time (T_1, T_2) and the censoring time (C_1, C_2) are assumed to be independent. The marginal survival function for (T_1, T_2) given x_i is denoted by $S_k(t_k; x_{ki})$, k = 1, 2 under the Cox regression model. The hazard function with known x_i is given by Equation 1 [12];

$$h(t; \boldsymbol{x}_i) = h_0(t) \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}) \quad (1)$$

where β is a p-dimensional regression vector and $h_0(t) = \int_0^t \lambda_0(s)ds$ represents the cumulative baseline hazard. To define the joint survival function, the copula function, as a commonly used function to link the two marginal functions, will be employed.

2.2.2. Bivariate Cox Survival Model with Copula

The copula survival model is a statistical analysis approach that incorporates the copula function in analyzing survival data. According to Sklar's theorem, there exists a copula function C

Table 2. Description of cataract patient data.

	Right Eye (OD) (Proportion)	Left Eye (OS) (Proportion)	Total (Proportion)
Censored (δ = 0)	24 (17.39%)	21 (15.22%)	45 (16.30%)
Recovery (δ = 1)	114 (82.61%)	117 (84.78%)	231 (83.70%)
Total	138 (100.00%)	138 (100.00%)	276 (100.00%)

that combines two continuous marginal survival functions S into a joint survival function to capture the dependence structure that characterizes the association between the two marginal distributions [13]. In the bivariate case, where T_1 and T_2 are non-negative random variables representing the survival times of two observed subjects with marginal survival functions $S_1(t_1) = P(T_1 > t_1)$ and $S_2(t_2) = P(T_2 > t_2)$, the joint survival function $S(t_1,t_2) = P(T_1 > t_1, T_2 > t_2)$ under the copula model is defined as follows Equation 2 [14]:

$$S(t_1,t_2) = C\big(S_1(t_1),S_2(t_2)\big); t_1,t_2 \geq 0 \quad (2)$$

with the density function of (T_1, T_2) given by Equation (3):

$$f(t_1,t_2) = c\big(S_1(t_1),S_2(t_2)\big); t_1,t_2 \geq 0 \quad (3)$$

where $c(u, v) = \partial^2 C(u, v)/\partial u\partial v$ for $u, v \in (0,1)$.

The Archimedean copula family is one of the most popular copula models due to its flexibility and simplicity in modeling events with bivariate endpoints [15]. Another idea derived from Sklar's theorem is to model the Archimedean copula function and the marginal distributions separately in order to construct the joint distribution.

2.2.3. Copula Function

One of the oldest families of distributions commonly used to model bivariate correlation is the Archimedean copula family. Its copula function is defined by the following Equation (4):

$$C(u,v) = \varphi^{-1}\big(\varphi(u) + \varphi(v)\big) \quad (4)$$

where φ is a continuous generator function of the copula, φ^{-1} is the generator's inverse, (U, V) is a pair of random variables, and C contains information about the dependence structure between the two variables (U, V), where $C(u, v) = P(U \leq u, V \leq v)$.

Clayton, Frank, Gumbel, and Joe are the four Archimedean copula families that are frequently utilized. The dependence parameter of the copula function will be derived from the implementation results of two copula functions within the bivariate Cox survival model because there is no particular method for determining the value of Kendall's τ in this study. The Gumbel copula, sometimes referred to as the Gumbel–Hougaard copula, is an asymmetric copula with greater upper tail dependence than lower tail dependence. The Gumbel copula can be defined by the following Equation 5 [16]:

-

$$C_\theta(u,v) = \exp\left(-\{(-\log u)^\theta + (-\log v)^\theta\}^{\frac{1}{\theta}}\right) \quad (5)$$

with $\theta > 0$ and the generator function $\varphi(t) = (-\log t)^{1/\theta}$. The Clayton copula is also asymmetric, however, in contrast to the Gumbel copula, it exhibits stronger lower tail dependence than upper tail dependence. The Clayton copula can be defined by the following Equation 6 [17]:

$$C_\theta(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}} \quad (6)$$

with $\theta > 0$ and the generator function $\varphi(t) = \frac{t^{-\theta}-1}{\theta}$.

2.2.4. Dependence Measurement

Kendall's τ and other dependence indicators are directly correlated with the copula parameter θ. Equation (7) defines the value of Kendall's τ, which may be obtained from the function θ in the copula model. It is evident that Kendall's τ is independent of the marginal distributions and only depends on the bivariate copula function [17]:

$$\tau = 4\int_0^1\int_0^1 C_\theta(u,v)c_\theta(u,v)dudv - 1 \quad (7)$$

\

$c(u, v) = \partial^2 C(u, v)/\partial u\partial v$. The formula for computing Kendall's τ for the Gumbel copula is given by Equation (8), and for the Clayton copula, it is presented in Equation (9).

$$\tau = 1 - \frac{1}{\theta} \quad (8)$$

$$\tau = \frac{\theta}{\theta + 2} \quad (9)$$

2.2.5. Marginal Cox Proportional Hazard Model

In certain cases, parametric marginal distributions may lead to biased estimation results due to a mismatch between the assumed distribution and the actual data of the study. On the other hand, the estimation of the dependence parameter in the copula model is not robust to misspecification of

Table 3. Description of cataract patient data based on independent variables.

Variable	Right Eye (OD) (Proportion)			Left Eye (OS) (Proportion)		
	Censored ($\delta = 0$)	Recovery ($\delta = 1$)	Total	Censored ($\delta = 0$)	Recovery ($\delta = 1$)	Total
Gender						
Man	11 (19.64%)	45 (80.36%)	56 (100.00%)	8 (14.29%)	48 (85.71%)	56 (100.00%)
Woman	13 (15.85%)	69 (84.15%)	82 (100.00%)	13 (15.85%)	69 (84.15%)	82 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)
Age						
< 60 y.o.	11 (18.03%)	50 (81.97%)	61 (100.00%)	10 (16.39%)	51 (83.61%)	61 (100.00%)
≥ 60 y.o.	13 (16.88%)	64 (83.12%)	77 (100.00%)	11 (14.29%)	66 (85.71%)	77 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)
Hypertension History						
No	22 (18.49%)	97 (81.51%)	119 (100.00%)	19 (15.97%)	100 (84.03%)	119 (100.00%)
Yes	2 (10.53%)	17 (89.47%)	19 (100.00%)	2 (10.53%)	17 (89.47%)	19 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)
Heart Disease History						
No	24 (17.52%)	113 (82.48%)	137 (100.00%)	21 (15.33%)	116 (84.67%)	137 (100.00%)
Yes	0 (0.00%)	1 (100.00%)	1 (100.00%)	0 (0.00%)	1 (100.00%)	1 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)
Diabetic History						
No	17 (16.35%)	87 (83.65%)	104 (100.00%)	14 (13.46%)	90 (86.54%)	104 (100.00%)
Yes	7 (20.59%)	27 (79.41%)	34 (100.00%)	7 (20.59%)	27 (79.41%)	34 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)
Stomach Disease History						
No.	24 (17.8%)	111 (82.22%)	135 (100.00%)	20 (14.81%)	115 (85.19%)	135 (100.00%)
Yes	0 (0.00%)	3 (100.00%)	3 (100.00%)	1 (33.33%)	2 (66.67%)	3 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)
Stroke History						
No	24 (17.52%)	113 (82.48%)	137 (100.00%)	21 (15.33%)	116 (84.67%)	137 (100.00%)
Yes	0 (0.00%)	1 (100.00%)	1 (100.00%)	0 (0.00%)	1 (100.00%)	1 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)
Stroke History						
No	21 (17.07%)	102 (82.93%)	123 (100.00%)	19 (15.45%)	104 (84.55%)	123 (100.00%)
Yes	3 (20.00%)	12 (80.00%)	15 (100.00%)	2 (13.33%)	13 (84.67%)	15 (100.00%)
Total	24 (17.39%)	114 (82.61%)	138 (100.00%)	21 (15.22%)	117 (84.78%)	138 (100.00%)

the marginal distributions. Therefore, the marginal distributions used in this study are semiparametric models, in which the baseline hazard function doesn't follow any specific parametric distribution and is treated as a piecewise constant across all uncensored event times [2]. The right-censored marginal Cox proportional survival function for T_{ki} with covariate x_i can be expressed as follows Equation 10:

$$S_k(t_{ki}|\boldsymbol{x}_{ki}) = P(T_{ki} > t_{ki}|\boldsymbol{x}_{ki}) = \exp(-S_{0k}(t_{ki})\exp(\boldsymbol{x}_{ki}^T\boldsymbol{\beta}_k)); \quad k = 1, 2;\ i = 1, 2, .., n, \tag{10}$$

where $S_{0k}(t_{ki}) = \sum_{i=1}^{n} \frac{I(y_{ki} \le t)\delta_{ki}}{\sum_{j \in R_{ik}} \exp(x_{ki}^T \beta_k)}$ is the Breslow baseline cumulative hazard function for the k-th marginal, $\beta_k = [\beta_1^T, \beta_2^T]^T$ and $R_{ik} = \{j: y_{kj} \ge y_{ki}\}$ denotes the at-risk set at time y_{ki}.

2.2.6. Pseudo-Likelihood Function for Right-Censored Bivariate Data

The pseudo-likelihood method is one way to get a parameter estimator for a model built from the joint distribution of a set of random variables [18]. This method is employed in this study because a joint distribution model is represented by the bivariate Cox survival model with a copula function. Previously, the survival function $S(t_1, t_2)$ and the joint density function $f(t_1, t_2)$ for (T_1, T_2) were defined. Consider the data $D = \{D_i\}^n_{i=1}$, there are four types of possible observations, including [19]:

1. Both T_{1i} and T_{2i} are observed with no censoring, meaning $\delta_{1i} = \delta_{2i} = 1$. The likelihood component for this case is $\frac{\partial^2 S(y_{1i}, y_{2i}|x_i)}{\partial y_{1i} \partial y_{2i}}$.
2. If only T_{1i} is censored and T_{2i} is observed, meaning $\delta_{1i} = 0$ and $\delta_{2i} = 1$. The likelihood component for this case is $\frac{-\partial S(y_{1i}, y_{2i}|x_i)}{\partial y_{2i}}$.
3. If only T_{2i} is censored and T_{1i} is observed, meaning $\delta_{1i} = 1$ and $\delta_{2i} = 0$. The likelihood component for this case is $\frac{-\partial S(y_{1i}, y_{2i}|x_i)}{\partial y_{1i}}$.
4. Both T_{1i} and T_{2i} are censored, meaning $\delta_{1i} = \delta_{2i} = 0$. The likelihood component for this case is $S(y_{1i}, y_{2i} \mid x_i)$.

The unknown parameters in the function $S(t_1, t_2)$ are the function S_{0k} and $\theta = (\beta_1^T, \beta_2^T, \eta)^T$, where S_{0k} and β_k represent the baseline survival function for the k-th margin and the regression coefficient vector, respectively, and η denotes the dependence parameter. The likelihood function can be written as follows Equation 11 [2]:

$$L(\boldsymbol{\theta}|\hat{S}_{01}, \hat{S}_{02}) = \prod_{i=1}^{n} [f(y_{1i}, y_{2i}|\boldsymbol{x}_i)]^{\delta_{1i}\delta_{2i}} \times \left[-\frac{\partial S(y_{1i}, y_{2i}|\boldsymbol{x}_i)}{\partial y_{1i}}\right]^{\delta_{1i}(1-\delta_{2i})} \times \left[-\frac{\partial S(y_{1i}, y_{2i}|\boldsymbol{x}_i)}{\partial y_{2i}}\right]^{(1-\delta_{1i})\delta_{2i}} \times S(y_{1i}, y_{2i}|\boldsymbol{x}_i)^{(1-\delta_{1i})(1-\delta_{2i})} \tag{11}$$

or can be expressed as Equation 12:

$$L(\boldsymbol{\theta}|\hat{S}_{01}, \hat{S}_{02}) = \prod_{i=1}^{n} [c\{S_1(y_{1i}|\boldsymbol{x}_{1i}), S_2(y_{2i}|\boldsymbol{x}_{2i})\}]^{\delta_{1i}\delta_{2i}} \times \left[-\frac{\partial C\{S_1(y_{1i}|\boldsymbol{x}_{1i}), S_2(y_{2i}|\boldsymbol{x}_{2i})\}}{\partial y_{1i}}\right]^{\delta_{1i}(1-\delta_{2i})} \times \left[-\frac{\partial C\{S_1(y_{1i}|\boldsymbol{x}_{1i}), S_2(y_{2i}|\boldsymbol{x}_{2i})\}}{\partial y_{2i}}\right]^{(1-\delta_{1i})\delta_{2i}} \times [C\{S_1(y_{1i}|\boldsymbol{x}_{1i}), S_2(y_{2i}|\boldsymbol{x}_{2i})\}]^{(1-\delta_{1i})(1-\delta_{2i})} \tag{12}$$

where $(\delta_{1i}, \delta_{2i}) \in \{(0, 0),(0, 1),(1, 0),(1, 1)\}$ and the log-likelihood function can be written as $\ell(\theta|\hat{S}_{01},\hat{S}_{02}) = \log L(\theta|\hat{S}_{01},\hat{S}_{02},D_i)$. Due to the complex form of the likelihood, a two-stage estimation method using the Brooyden–Fletcher–Goldfarb–Shanno (BFGS) numerical iteration approach will be implemented in estimating the parameter θ. This method has been proven to be both computationally stable and efficient [20].

2.2.7. Two-Step Estimation Procedure

In 1970, Broyden, Fletcher, Goldfarb, and Shanno created the BFGS method, a numerical iteration methodology that partially modified the Newton-Raphson algorithm. The adjustment entails substituting an approximation matrix that only needs gradient information for the Hessian matrix. This approximation matrix's positive definiteness guarantees a quicker rate of convergence and improves robustness by transferring knowledge from earlier iterations. Let θ denote the vector of unknown parameters in the model, and the following presents the BFGS iterative algorithm for estimating θ:

1. Estimate β_k using Cox PH and S_{0k} using Breslow estimator then save as $\beta_k^{(0)}$ and $\hat{S}_{0k}$.
2. Estimate η using $\ell(\theta|\beta_k^{(0)}, \hat{S}_{0k.})$; $k = 1, 2$, then save as $\eta^{(0)}$.
3. Determine the gradient vector $\boldsymbol{g}(\boldsymbol{\theta}) = (\frac{\partial \ell(\boldsymbol{\theta}|S_{01}, S_{02})}{\partial \boldsymbol{\beta}_1}, \frac{\partial \ell(\boldsymbol{\theta}|S_{01}, S_{02})}{\partial \boldsymbol{\beta}_2}, \frac{\partial \ell(\boldsymbol{\theta}|S_{01}, S_{02})}{\partial \eta})$.
4. Set initial value $\theta^{(0)} = (\beta_k^{(0)}, \eta^{(0)})$ for joint estimation.

Table 4. Results of the model analysis with the Clayton Copula.

Independent Variable	$\hat{\beta}_j$	$SE(\hat{\beta}_j)$	Stat(Z)	p-value
X_1 : Gender	0.168	0.177	0.899	0.342
X_2 : Age	-0.747	0.212	12.413	0.000*
X_3 : Hypertension History	-0.425	0.333	1.631	0.202
X_4 : Heart Disease History	-0.396	1.201	0.108	0.742
X_5 : Diabetic History	-0.506	0.289	3.042	0.081*
X_6 : Stomach Disease History	0.042	0.466	0.008	0.928
X_7 : Stroke History	1.416	2.099	0.455	0.499
X_8 : Cholesterol History	0.256	0.237	1.168	0.280
η : Clayton Copula Coefficient	3.362	0.607	30.660	3.060×10^{-8}*

5. Estimate θ iteratively using the following Equation 13 :

$$\hat{\boldsymbol{\theta}}^{(m+1)} = \hat{\boldsymbol{\theta}}^{(m)} - \left(\boldsymbol{A}^{(m)}\right)^{-1} \boldsymbol{g}\left(\hat{\boldsymbol{\theta}}^{(m)}\right), \quad \textbf{(13)}$$

where $\boldsymbol{A}^{(0)} = \boldsymbol{I}$ Using the iterative BFGS method, the following formula is used to update the matrix $\boldsymbol{A}$ Equation 14.

$$\boldsymbol{A}^{(m+1)} = \boldsymbol{A}^{(m)} + \frac{\boldsymbol{q}_m \boldsymbol{q}_m^T}{\boldsymbol{q}_m^T \boldsymbol{s}_m} - \frac{\boldsymbol{A}^{(m)} \boldsymbol{s}_k \boldsymbol{s}_k^T \boldsymbol{A}^{(m)}}{\boldsymbol{s}_m^T \boldsymbol{B}^{(m)} \boldsymbol{s}_m} \quad \textbf{(14)}$$

where $\boldsymbol{s}_m = \hat{\boldsymbol{\theta}}^{(m+1)} - \hat{\boldsymbol{\theta}}^{(m)}$ and $\boldsymbol{q}_m = \boldsymbol{g}\left(\hat{\boldsymbol{\theta}}^{(m+1)}\right) - \boldsymbol{g}\left(\hat{\boldsymbol{\theta}}^{(m)}\right)$

6. The iteration is terminated if $||\hat{\theta}^{(m+1)} - \hat{\theta}^{(m)}|| < \epsilon$ with $\epsilon > 0$.

7. Store the optimal estimated parameters as $\hat{\theta}$.

2.2.8. Significance Testing of Model Parameters

2.2.8.1. Simultaneous Test

The simultaneous significance test of parameters is carried out using the likelihood ratio test (LRT) with the following hypotheses [21]:

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_p = 0$ (no variables have a significant effect on survival time)
H_1: at least one $\beta_j \neq 0$ (there exists at least one variable that significantly affects survival time).

The likelihood ratio test statistic is given by the following Equation 15:

$$G^2 = 2\left(\ell(\hat{\boldsymbol{\theta}}) - \ell(\boldsymbol{\theta}^{\sim})\right) \approx^d \chi_p^2 \quad \textbf{(15)}$$

where $\ell(\hat{\theta})$ is the maximum log-likelihood of the model under $H_0 \cup H_1$, and $\ell(\theta^{\sim})$ represents the maximum log-likelihood of the model under H_0. n_f denotes the number of individuals at risk of experiencing the event at time $t_{(f)}$. H_0 is rejected if $G_2 > \chi^2_{\alpha,p}$ or if the p-value < α. When the simultaneous test leads to a rejection of H_0, indicating that at least one variable significantly affects survival time, a partial significance test is then performed to identify which variables have significant effects.

2.2.8.2. Partial Test

The partial significance test of parameters is carried out using the Wald Test with the following hypotheses [21]:

$H_0 : \beta_j = 0$ (the j-th variable does not have a significant effect on the patient's survival time);
$H_1 : \beta_j \neq 0$, $j = 1, 2, \ldots, p$ (the j-th variable significantly affects the patient's survival time).

The Wald test statistic is given by the following Equation 16:

$$W_j = \frac{\hat{\beta}_j}{\widehat{SE}(\hat{\beta}_j)} \approx^d N(0,1) \text{ under } H_0 \quad \textbf{(16)}$$

with $\left(\widehat{SE}(\hat{\beta}_j)\right)^2 = \widehat{var}(\hat{\beta}_j)$. H_0 is rejected if or if $|W_j| > z_{\frac{\alpha}{2}}$ the p-value < α, which means that the patient's survival time is significantly affected by the j-th variable.

2.2.9. Model Selection

Model selection aims to decide the best model close to the true one among several possible models. Model selection is a very important step, as the resulting dependence structure will vary depending on the copula function applied. The Akaike's information criterion (AIC) is probably the most well-known model selection criterion. Another criterion, the Bayesian information

criterion (BIC) serves as an approximation to a transformation of the Bayesian posterior probability of a candidate model. Since these two conventional criteria are likelihood-based and full multivariate likelihoods, they cannot be directly applied in pseudo-likelihood-based approaches because pseudo-likelihood-based methods do not represent these kinds of likelihood. Here, we introduce the modification of AIC, called quasi-information criterion (QIC), for bivariate survival model. The QIC was constructed by replacing the likelihood in the Kullback-Leibler information with the quasi-likelihood under the working independence assumption [22]. To understand the basic concept of QIC, we first briefly review the derivation of AIC.

Suppose we have a candidate model M_{Can} and the true model M_{True} with log-likelihood function $\ell(\theta; D_{True})$ and log-likelihood $\ell(\theta_{True}; D_{True})$, respectively. A well-known measure of separation between two models is given by Kullback-Leibler information [22]. The Kullback-Leibler information of M_{Can} from M_{True} is Equation 17.

$$\Delta_0(\boldsymbol{\theta}_{True}, \boldsymbol{\theta}) = E_{\boldsymbol{\theta}_{True}}(\ell(\boldsymbol{\theta}_{True}; \mathbf{D}_{True}) - \ell(\boldsymbol{\theta}; \mathbf{D}_{True}) \quad (17)$$

where the expectation θ is taken with respect to the true distribution of D_{True}. From a set of candidate models M_{All}, in which each can be indexed by θ, we would like to choose the smallest $\ell(\theta_{True}; \theta)$. However, both θ and θ_{True} are unknown. The AIC is defined as Equation 18:

$$AIC = -2\ell(\hat{\boldsymbol{\theta}}; \mathbf{D}_{True}) + 2p \quad (18)$$

where p is the dimension of θ and $\hat{\theta}$ is the maximum likelihood estimator (MLE) under any candidate model in M_{All}. Model selection is achieved by selecting from M_{All} the one that minimizes AIC. Since in pseudo-likelihood methods we do not have a true likelihood function, we replace ℓ in (17) by the quasi-log-likelihood Q under the working independence model and define a new discrepancy as Equations 19 - 21:

$$\Delta(\boldsymbol{\theta}_{True}, \boldsymbol{\theta}; \ell) = E_{\boldsymbol{\theta}_{True}}(\ell(\boldsymbol{\theta}_{True}; \ell, \mathbf{D}_{True}) - Q(\boldsymbol{\theta}; \mathbf{D}_{True}) \quad (19)$$

Where:

$$E_{\boldsymbol{\theta}_{True}}\left(-\frac{\partial Q(\boldsymbol{\theta}; \ell, \mathbf{D}_{True})}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{True}}\right) = 0 \quad (20)$$

and

$$\Omega = E_{\boldsymbol{\theta}_{True}}\left(-\frac{\partial^2 Q(\boldsymbol{\theta}; \ell, \mathbf{D}_{True})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{True}}\right) = \sum_{i=1}^{n} \boldsymbol{D}_i^T \boldsymbol{V}_i \boldsymbol{D}_i \quad (21)$$

Ω is the model-based covariance estimator under independence working correlation structure, and $\hat{V}_{san}$ is the sandwich covariance estimator under the working correlation structure **R** in Equation (21).

$$QIC(R) = -2Q(\hat{\boldsymbol{\theta}}(R); \boldsymbol{I}, \boldsymbol{D}_{True}) + 2tr(\boldsymbol{\Omega}\hat{\boldsymbol{V}}_{san}) \quad (22)$$

The selection of the survival model with a copula function is based on the model that produces the smallest QIC value (Equation 22).

3. RESULTS AND DISCUSSIONS

3.1. Characteristics of Cataract Patient Data

The data used in this study consist of 276 observations obtained from 138 patients who

Table 5. Results of the model analysis with the Gumbel Copula.

Independent Variable	$\hat{\beta}_j$	$SE(\hat{\beta}_j)$	Stat(Z)	p-value
X_1 : Gender	0.149	0.209	0.507	0.476
X_2 : Age	-0.664	0.257	6.639	0.009*
X_3 : Hypertension History	-1.048	0.445	5.548	0.018*
X_4 : Heart Disease History	-0.458	1.408	0.105	0.745
X_5 : Diabetic History	-0.474	0.277	2.909	0.088*
X_6 : Stomach Disease History	-0.328	0.630	0.272	0.602
X_7 : Stroke History	1.021	2.357	0.187	0.665
X_8 : Cholesterol History	0.099	0.223	0.198	0.656
η : Clayton Copula Coefficient	2.979	0.305	94.936	2.200×10^{-16} *

underwent cataract surgery on both the right eye (*oculus dextra*) and the left eye (*oculus sinistra*) and were subsequently declared recovered.

The number and percentage of censoring status for the 276 observations (138 for each eye) are shown in Table 2. The majority of patients were found to have recovered, with 114 right eyes, or 82% of all patients, and 117 left eyes, or 84% of all patients. The remaining cases, which represented circumstances in which details regarding the patients' recuperation times were either lacking or imprecise, were censored.

Table 3 presents a description of cataract patient data using contingency tables (crosstabs) based on the independent variables. Eleven male patients (19.64%) and thirteen female patients (15.85%) were censored for the right eye. There were more female patients (69, or 84.15%) than male patients (45, or 80.36%) who either recovered or had the incident. The percentage of censored data for the left eye was 13 females (15.85%) and 8 males (14.29%). There were 48 male patients (85.71%) and 69 female patients (84.15%) who had the incident. Although female patients showed a little higher recovery rate than male patients in both eyes, overall, there was no discernible difference between the sexes in terms of recovery status.

The gender disparities in survival curves were then statistically evaluated using a log-rank test. The alternative hypothesis (H_1) asserts that there is a difference between the survival curves of male and female patients, whereas the null hypothesis (H_0) asserts that there is no difference. The p-values obtained from the log-rank test were 0.54 for the left eye and 0.92 for the right. Since both p-values are significantly greater than the significance level of α=0.1, H_0 cannot be rejected. Consequently, it can be said that the survival curves for the right and left eyes of male and female patients do not differ statistically significantly.

Using 60 as the age criterion, the distribution of recovery status among cataract patients was categorized. In the age group under 60, there were 11 censored patients (18.03%) for the right eye; in the group over 60, there were 13 censored patients (16.88%). Patients 60 years of age and beyond had a greater recovery rate—64 patients, or 83.12%—than the under-60 group, which included 50 patients, or 81.97%. The percentage of censored data for the left eye was 14.29% (11 individuals) in the over-60 group and 16.39% (10 people) in the under-60 group. In a similar vein, patients 60 years of age and above had a slightly greater recovery rate for the left eye (85.71%, 66 individuals) than those under 60 (83.61%, 51 individuals). Patients 60 years of age and older tended to have a somewhat higher recovery rate in both eyes, but overall there was no significant difference between the two age groups.

The differences in survival between the two age groups were statistically evaluated using a log-rank test. The alternative hypothesis (H_1) asserts that there is a difference between the survival curves of patients under 60 and those over 60. The null hypothesis (H_0) asserts that there is no difference. For the right and left eyes, the log-rank test findings yield p-values of 0.0012 and 0.0002, respectively. The rejection of H_0 results from both p-values being less than the significance level of α=0.1. Thus, it can be said that the survival curves for both the right and left eyes differ statistically significantly between individuals under 60 and those above 60.

According to the contingency table (Table 3), which describes cataract patients according to their history of hypertension, 22 individuals (18.49%) of patients without a history of hypertension were censored for the right eye, whereas 97 individuals (81.51%) experienced the event (recovery). Of the patients having a history of hypertension, 17 (89.47%) had the incident, and 2 (10.53%) were censored. Two patients (10.53%) with a history of hypertension and 19 patients (15.97%) without one experienced censoring in the left eye. Patients without a history of hypertension had a slightly lower recovery rate of 84.03% (100 individuals) compared to 89.47% (17 persons) among those with a history of hypertension, notwithstanding the small differences. Patients without a history of hypertension tended to recover more quickly for the right eye. Patients with a history of hypertension, however, showed a somewhat greater recovery rate for the left eye than those without.

A log-rank test was also conducted to examine whether there is a significant difference between the survival curves of patients with and without hypertension. The null hypothesis (H_0) states that there is no difference between the survival curves of the two groups, while the alternative hypothesis

Table 6. BIC score of Copula Model.

Copula Function	QIC	Kendall Tau (τ)
Clayton	1533.551	0.627
Gumbel	1529.078	0.664

(H_1) states that such a difference exists. The results of the log-rank test yielded p-values of 0.0003 for the right eye and 0.0001 for the left eye. Both p-values are smaller than the significance level of α = 0.1, leading to the rejection of H_0. This indicates that there is a statistically significant difference between the survival curves of patients with and without hypertension, for both the right and left eyes.

Based on Table 3, it can be observed that for the right eye, 24 patients (17.52%) without a history of heart disease were censored, while 113 patients (82.48%) experienced the event (recovery). While in the group with a history of heart disease, no censored data were found, and only one patient (100%) experienced the event. For the left eye, the results were nearly identical to those for the right eye: 21 patients (15.33%) without a history of heart disease were censored, while 116 patients (84.67%) experienced the event. Similarly, among patients with a history of heart disease, there were no censored cases, and only one patient (100%) experienced the event.

A log-rank test was performed with the null hypothesis (H_0) stating that there is no difference between the survival curves of patients with and without a history of heart disease, and the alternative hypothesis (H_1) stating that there is a difference between the two groups. The log-rank test results p-values of 0.2 for the right eye and 0.13 for the left eye. Both p-values are greater than the significance level of α = 0.1, leading to fail to reject H_0. This indicates that there is no significant difference between the survival curves of patients with and without a history of heart disease for either eye.

Patients without a history of diabetes had an event proportion of 83.65% (87 people) and a censoring proportion of 16.35% (17 individuals) for the right eye. In contrast, patients with a history of diabetes had an event proportion of 79.41% (27 individuals) and a censoring proportion of 20.59% (7 individuals). Patients without a history of diabetes had a censoring proportion of 13.46% (14 individuals) and an event proportion of 86.54% (90 individuals) for the left eye, while patients with a history of diabetes had a censoring proportion of 20.59% (7 individuals) and an event proportion of 79.41% (27 individuals). Following bilateral cataract surgery, individuals with a history of diabetes often showed a lower recovery % and a higher censoring rate in both eyes, indicating that a history of diabetes may contribute to a slower postoperative healing process.

To infer the difference between the categories, a log-rank test was conducted with the null hypothesis (H_0) stating that there is no difference between the survival curves of patients with and without a history of diabetes mellitus, and the alternative hypothesis (H_1) stating that such a difference exists. The log-rank test results yielded p -values of 0.011 for the right eye and 0.025 for the left eye. Both p-values are smaller than the significance level of α = 0.1, leading to the rejection of H_0. This indicates that there is a statistically significant difference between the survival curves of patients with and without a history of diabetes mellitus for both the right and left eyes.

Table 3 presents the features of cataract patients according to their history of gastrointestinal disease. For the right eye, 111 patients (82.22%) who did not have a history of stomach disease experienced the incident; 24 patients (17.78%) were excluded. With no censored data, all three patients (100%) who had a history of stomach disease experienced the incident. Patients having a history of stomach disease indicated that 66.67% (2 individuals) had the event for the left eye, whereas 33.33% (1 individual) were excluded. A log-rank test was also conducted, with the alternative hypothesis (H_1) asserting that there is a difference between the survival curves of patients with and without a history of stomach disease, and the null hypothesis (H_0) asserting that there is no difference. The test yields p-values of 0.83 for the left eye and 0.41 for the right. Since both p-values are significantly

greater than the significance level of $\alpha = 0.1$, H_0 cannot be rejected. This suggests that the survival curves for both the right and left eyes of people with and without a history of gastrointestinal disease do not differ significantly. Table 3 reveals that among patients without a history of stroke, 113 (82.48%) experienced the occurrence in their right eye, and 24 (17.52%) were censored. On the other hand, no censored data were gathered for people with a history of stroke, and just one person (100%) experienced the occurrence. Similar results are shown for the left eye, where the event occurred in only one person (100%) and no suppressed data was found in those with a history of stroke. 116 (84.67%) of the patients without a history of stroke experienced the occurrence, whereas 21 (15.33%) were excluded.

The alternative hypothesis (H_1) for the log-rank test states that patients with and without a history of stroke have different survival curves, whereas the null hypothesis (H_0) states that no such difference exists. The log-rank test yields p-values of 0.39 and 0.44 for the right and left eyes, respectively. Given that both p-values exceed the significance threshold of $\alpha = 0.1$, H_0 cannot be ruled out. This implies that there is no statistically significant difference in the survival curves for the right and left eyes of patients with and without a history of stroke.

Lastly, Table 3 shows the recovery distribution based on cholesterol disease. For the right eye, 17.07% (21 individuals) patients without a history of cholesterol disease were censored, and the remaining 82.93% (102 individuals) experienced an event. Meanwhile, patients with a history of cholesterol disease showed a slightly higher proportion of censored data at 20.00% (3 individuals), and the event proportion is 80.00% (12 individuals). For the left eye, 19 individuals (15.45%) without a history of cholesterol disease were censored, and 104 individuals (84.55%) experienced the event. Among patients with a history of cholesterol disease, 2 individuals (13.33%) were censored, and 13 individuals (86.67%) experienced the event.

By a log-rank test, the null hypothesis (H_0) states that there is no difference in the survival curves between patients with and without a history of cholesterol disease, and the alternative hypothesis (H_1) states that a difference exists. The log-rank test results give p-values of 0.62 for the right eye and 0.61 for the left eye. Both p-values are well above the significance level of $\alpha = 0.1$, leading to the failure to reject H_0. This indicates that there is no significant difference between the survival curves of patients with and without a history of cholesterol disease, for both the right and left eyes.

3.2. Bivariate Cox Survival Model with Clayton Copula

The estimated regression coefficients from the model with the Clayton copula are presented in Table 4. The simultaneous test of the model yielded a p-value=0.001 and $G = 1.940 > \chi^2_{0.1;8} = 13.360$. Using a significance level $\alpha = 0.1$, we conclude to reject H_0; that is, there is evidence that at least one independent variable has a significant effect on the recovery of cataract patients after surgery.

Partially, at $\alpha = 0.1$ and with a $z_{\frac{0.1}{2}} = 2.710$, it is concluded that the independent variables age and history of diabetes have a significant effect on the recovery time of cataract patients who underwent surgery on both eyes. The same conclusion also applies to the dependence parameter, $\hat{\eta}$, so it can be inferred that recovery times of the right and left eyes are significantly dependent, with Kendall's $\tau = 0.627$.

3.3. Bivariate Cox Survival Model with Gumbel Copula

A similar conclusion was also obtained from this model. The results of the simultaneous significance test show a p-value=0.001 and $G = 15.740 > \chi^2_{0.1;8}$. At a significance level of $\alpha = 0.1$, this leads to the rejection of H_0, indicating that at least one independent variable significantly influences the recovery of cataract patients following surgery.

Table 5 presents the estimated regression coefficients from the model using the Gumbel copula function. Partially, at $\alpha = 0.1$ and with a $z_{\frac{0.1}{2}} = 2.710$, it is concluded that the independent variables age, history of hypertension, and history of diabetes have a significant effect on the recovery time of cataract patients who underwent surgery on both eyes. The dependence parameter $\hat{\eta}$ also has a very small p-value with a big statistic W, so it can be inferred that recovery times of the right and left eyes are significantly dependent, with Kendall's $\tau = 0.664$.

3.4. Model Selection with the Best Copula Function

Finding the best-fitting model is the goal of model evaluation based on each model's QIC values. Table 6 displays the QIC values and Kendall's Tau for both models, demonstrating that the model using the Gumbel copula function produces a lower QIC value than the model using the Clayton copula function. Thus, it can be said that the recovery events of cataract patients who had surgery on both eyes are best explained by the bivariate Cox survival model with the Gumbel copula.

According to the model, the combined survival probability falls as the surviving duration for both eyes grows. This suggests that the probability that both eyes have not yet recovered decreases with time, which is consistent with the survival function's general feature. Additionally, the patient data indicates that the recovery times for the right and left eyes are comparatively close, indicating that recovery in one eye is probably going to be followed by recovery in the other eye in a short amount of time. The Kendall's Tau value of 0.664 in the estimation findings shows a moderate and positive relationship between the recovery times of the left and right eyes. Therefore, it can be inferred that patients who recover more quickly in their right eye also typically recover more quickly in their left eye, and vice versa. This result is in line with the Gumbel copula's ability to capture upper-tail reliance, or the correlation between greater recovery times and simultaneous recovery over longer durations.

3.5. Hazard Ratio Analysis

The hazard ratio (HR) values were obtained from the exponentiated coefficients exp $\hat{\beta}_j$ of the independent variables identified as risk factors influencing the recovery of cataract patients who underwent surgery on both eyes. Table 7 summarizes the estimated HR values, focusing only on variables that showed a statistically significant effect in the best model, as determined by the previous BIC evaluation.

The estimated HR of 0.517 for the age variable, where x^*=1 denotes patients 60 years and older and x=0 denotes patients under 60, shows that patients 60 years and older have a recovery rate that is 0.52 times higher than that of younger patients. This indicates that compared to people under 60, the recovery rate for older patients is about 48% slower. This result is consistent with earlier research indicating that physiological changes, such as reduced tissue flexibility, slower cell regeneration, and impaired immune response, tend to cause older patients' healing processes to be slower. Even though cataract surgery is increasingly common in older people, the body's ability to adapt to healing processes still has a significant impact on recovery time.

Patients with hypertension have a recovery rate roughly 0.36 times higher than that of patients without hypertension, according to the estimated HR of 0.357 for the history of hypertension variable, where x^*=1 denotes patients with a history of hypertension and x = 0 denotes patients without such a history. This suggests that people with hypertension recover about 64% more slowly. Increased intraocular pressure, poor microcirculation, and an increased risk of postoperative hemorrhage are some of the ways that hypertension can hinder the healing process following surgery. Because uncontrolled hypertension can result in systemic problems that impede wound healing, hypertensive patients are frequently classified as high-risk during surgical planning and postoperative treatment in ophthalmic practice.

Patients with diabetes have a recovery rate that is 0.61 times greater than that of patients without diabetes, according to the estimated HR of 0.615 for the history of diabetes variable, where x^*=1

Table 7. Hazard ratio.

Variable	$\hat{\beta}_j$	exp ($\hat{\beta}_j$)
Age	-0.660	0.517
Hypertension History	-1.029	0.357
Diabetic History	-0.486	0.615

indicates patients with diabetes and $x = 0$ represents those without. This suggests that those with diabetes heal around 39% more slowly than those without the condition. Diabetes is known to hinder wound healing and increase the risk of complications after surgery, especially ophthalmologic surgery. This is explained by circulatory and metabolic issues associated with diabetes, which hinder ocular tissue regeneration and generally slow down the healing process.

4. CONCLUSIONS

A bivariate copula model was applied to analyze the dependence between the recovery times of the right and left eyes in patients with bilateral cataracts. There were several important findings coming from this study. Both two copula functions tested, the Clayton and Gumbel copula, the evaluation based on the QIC indicated that the Gumbel copula was the best-fitting model. According to the estimated dependence parameter and Kendall's Tau revealed a moderate positive association between the recovery times of both eyes, indicating that the recovery processes of the right and left eyes were not independent. Further analysis showed that among several examined variables—including gender, age, and medical histories such as hypertension, heart disease, diabetes mellitus, gastric disorders, stroke, and cholesterol—three variables significantly affected recovery time at the 10% significance level: age ($p = 0.009$), hypertension history ($p = 0.018$), and diabetes mellitus history ($p = 0.088$). The estimated HR indicated that patients aged 60 years and older had a 48% slower recovery rate compared to younger patients, those with a history of hypertension had a 64% slower recovery rate than non-hypertensive patients, and those with a history of diabetes mellitus had a 39% slower recovery rate compared to patients without diabetes. For future research, it is recommended to consider incorporating additional clinical and behavioral variables that may influence the patient's recovery process, such as adherence to medication, nutritional intake, and physical activity levels. Also, to reduce selection bias and enhance the generalizability of the model, it is recommended to use data with a larger sample size and covering a longer time period. Furthermore, it is also recommended to try generalized estimating equation (GEE) method to allow the modelling of correlated data. Instead of assuming a multivariate distribution of these data (responses), GEE specifies a variance function and a working covariance guess for the covariance structure among responses.

AUTHOR INFORMATION

Corresponding Author

Jerry Dwi Trijoyo Purnomo — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);
orcid.org/0000-0001-5150-5066
Email: jerry@its.ac.id

Authors

Purhadi Purhadi — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);
orcid.org/0000-0002-1592-6645

Shofi Andari — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);
orcid.org/0000-0001-5794-2646

Wahyu Dwi Rahmawati — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);
orcid.org/0009-0001-4350-6509

Sri Harini — Department of Mathematics, Universitas Islam Negeri Maulana Malik Ibrahim, Malang-65141 (Indonesia);
orcid.org/0000-0001-9664-027X

Author Contributions

J. D. T. P.: Conceptualization, Methodology, and Software. P. P.: supervision. S. A.: visualization and investigation. W. D. R.: Data curation and writing original draft preparation. S. H.: Validation, reviewing and editing. All authors discussed the results and contributed to the final manuscript.

Conflicts of Interest

The authors affirm that the work described in this publication was not influenced by any known competing financial interests or personal relationships.

ACKNOWLEDGEMENT

The authors gratefully acknowledge financial support from the Ministry of Research, Technology, and Higher Education for this work, under the project scheme of Regular Fundamental Research (BIMA) 2025 with contract number 1275/PKS/ITS/2025.

DECLARATION OF GENERATIVE AI

Grammarly was used by the authors to improve their grammar and style while preparing this work. The authors examined and revised all AI-generated content for accuracy, bias, and completeness and assume full responsibility for the final manuscript's content, integrity, and originality.

REFERENCES

[1] L. Lin, Y. Liang, G. Jiang, Q. Gan, T. Yang, P. Liao, and H. Liang. (2025). "Global, regional, and national burden of cataract: A comprehensive analysis and projections from 1990 to 2021". *PLOS ONE.* **20** (6): e0326263. 10.1371/journal.pone.0326263.

[2] T. Sun and Y. Ding. (2020). "CopulaCenR: Copula Based Regression Models for Bivariate Censored Data in R". *The R Journal.* **12** (1): 266-282. 10.32614/RJ-2020-025.

[3] G. Cheng, L. Zhou, X. Chen, and J. Z. Huang. (2014). "Efficient Estimation of Semiparametric Copula Models for Bivariate Survival Data". *Journal of Multivariate Analysis.* **123** : 330-344. 10.1016/j.jmva.2013.10.008.

[4] Y. Zhang, Y. Chen, and H. Li. (2020). "Copula-Based Analysis of Bilateral Breast Cancer Data". *Statistical Methods in Medical Research.* **29** (4): 1023-1037.

[5] K. Suresh, J. M. G. Taylor, and A. Tsodikov. (2021). "A Gaussian Copula Approach for Dynamic Prediction of Survival with a Longitudinal Biomarker". *Biostatistics.* **22** (3): 504-521. 10.1093/biostatistics/kxz049.

[6] J. Huang, H. Zhou, and N. Ebrahimi. (2022). "Bayesian Bivariate Cure Rate Models Using Copula Functions". *International Journal of Statistics and Probability.* **11** (3): 9-21. 10.5539/ijsp.v11n3p9.

[7] D. Petti, A. Eletti, G. Marra, and R. Radice. (2022). "Copula Link-Based Additive Models for Bivariate Time-to-Event Outcomes with General Censoring Scheme". *Computational Statistics & Data Analysis.* **175** : 107550. 10.1016/j.csda.2022.107550.

[8] W. He, G. Y. Yi, and A. Yuan. (2024). "Analysis of Multivariate Survival Data under Semiparametric Copula Models". *Canadian Journal of Statistics.* **52** (2): 380-413. 10.1002/cjs.11776.

[9] W. dos Reis Miranda Filho and F. N. Demarqui. (2025). "A Class of Semiparametric Models for Bivariate Survival Data". *Lifetime Data Analysis.* **31** : 102-125. 10.1007/s10985-024-09642-x.

[10] S. Hong, W. Park, Y. Eom, H. M. Kim, and J. S. Song. (2022). "Comparisons of Outcomes and Complications of Immediate Sequential Bilateral Cataract Surgery and Unilateral Cataract Surgery in a Tertiary Hospital in South Korea". *Scientific Reports.* **12** : 22382. 10.1038/s41598-022-26851-2.

[11] D. R. Cox and D. O. Oakes. (1984). "Analysis of Survival Data". Chapman & Hall/CRC, New York.

[12] T. H. Scheike, K. K. Holst, and J. B. Hjelmborg. (2015). "Measuring Early or Late Dependence for Bivariate Lifetimes of Twins". *Lifetime Data Analysis.* **21** (2): 280-299. 10.1007/s10985-014-9309-5.

[13] A. Sklar. (1959). "Fonctions de Répartition à n Dimensions et Leurs Marges". *Publications de l'Institut de Statistique de l'Université de Paris.* **8** (3): 229-231.

[14] T. Emura and Y.-H. Chen. (2018). "Analysis of Survival Data with Dependent Censoring: Copula-Based Approaches". Springer, Singapore. 10.1007/978-981-10-7164-5.

[15] P. A. Shaw and M. P. Fay. (2016). "A Rank Test for Bivariate Time-to-Event Outcomes When One Event Is a Surrogate". *Statistics in Medicine.* **35** (19): 3413-3423. 10.1002/sim.6950.

[16] E. J. Gumbel. (1960). "Bivariate Exponential Distributions". *Journal of the American Statistical Association.* **55** (292): 698-707.

10.1080/01621459.1960.10483368.

[17] D. G. Clayton. (1978). "A Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence". *Biometrika.* **65** (1): 141-151. 10.1093/biomet/65.1.141.

[18] P. B. Mpofu, G. Bakoyannis, C. T. Yiannoutsos, A. W. Mwangi, and M. Mburu. (2020). "A Pseudo-Likelihood Method for Estimating Misclassification Probabilities in Competing-Risks Settings When True-Event Data Are Partially Observed". *Biometrical Journal.* **62** (7): 1747-1768. 10.1002/bimj.201900198.

[19] M. Mei. (2016). "A Goodness-of-Fit Test for Semi-Parametric Copula Models of Right-Censored Bivariate Survival Times". Simon Fraser University.

[20] T. Sun and Y. Ding. (2021). "Copula-Based Semiparametric Regression Method for Bivariate Data under General Interval Censoring". *Biostatistics.* **22** (2): 315-330. 10.1093/biostatistics/kxz032.

[21] D. W. Hosmer, S. Lemeshow, and S. May. (2008). "Applied Survival Analysis: Regression Modeling of Time-to-Event Data". John Wiley & Sons, New Jersey. 10.1002/9780470258019.

[22] L. Y. Hin and T.-G. Wang. (2009). "Working -Correlation-Structure Identification in Generalized Estimating Equations". *Statistics in Medicine.* **28** (4): 642-658. 10.1002/sim.3489.