Journal of Statistics and Its Application, Likelihood Vol 1 No 2, December 2025

The Cox Proportional Hazards Approach in Survival Analysis of an Ovarian Cancer Dataset with Stratification and an Extended Model

Luqyana Zakiya Almas [1], Rizki Adisetya Tahwin [2], Ifayanti Rohmatul H [3], Anisya Safira [4], Sri Sulastri [5]

[1] Institut Sepuluh Nopember, Indonesia, email: 6003241031@student.its.ac.id
[2] Institut Sepuluh Nopember, Indonesia, email: 6003241020@student.its.ac.id
[3] Institut Sepuluh Nopember, Indonesia, email: 6003241022@student.its.ac.id
[4] Institut Sepuluh Nopember, Indonesia, email: anisyafira29@gmail.com
[5] Institut Sepuluh Nopember, Indonesia, email: 6003241039@student.its.ac.id

Abstract

This analysis uses the Ovarian dataset available on the R Pubs platform to evaluate the factors influencing the risk of an event occurring. This dataset is one of the most commonly used survival datasets for analyzing treatment effectiveness or risk factors in medical research. The aim of this analysis is to evaluate the Cox Proportional Hazards model in analyzing the factors influencing the risk of an event occurring, while accounting for interactions between variables. The two model approaches tested are the Stratified Cox and Extended Cox models, both with and without interactions. The evaluation was conducted using four primary methods: the Concordance Index, Dynamic-AUC, Brier Score, and Integrated Square Error. The likelihood ratio test results indicate that the inclusion of interactions does not significantly improve model fit, as evidenced by p-values > 0.05 for both model approaches. Therefore, the Stratified Cox model without interactions was selected as the best model. This model is considered simpler yet still capable of providing an accurate interpretation of the relationship between the studied variables and the risk of an event. The selection of this model is expected to make a significant contribution to the development of survival analysis, particularly in applications involving complex data with stratification factors. The analysis results indicate that patient age is significantly associated with the risk of death; each additional year of age increases the risk of death by approximately 2.4%. Treatment type and the presence of residual disease are not significantly associated with the risk of death; however, the analysis results show that patients receiving experimental treatment have a 28.4% lower risk of death, and patients with residual disease have a 38.9% higher risk of death compared to patients without residual disease.

Keywords: *Survival Analysis, Cox Proportional Hazards Model, Stratified Cox Model, Variable Interaction, Ovarian Cancer Dataset, Mortality Risk*

Introduction

Ovarian cancer is the seventh most common cancer among women, with genetic factors being one of the most significant contributors; therefore, preventive measures and early detection are strongly recommended (Momenimovahed et al., 2019). This cancer also has the poorest prognosis and the highest mortality rate (Coburn et al., 2017). It is estimated that by 2040, the mortality rate for this cancer will increase significantly (Bray, 2018). The high mortality rate of ovarian cancer is attributed to asymptomatic tumor growth that goes unnoticed, delayed symptom onset, and a lack of appropriate screening, leading to diagnosis only at an advanced stage. One approach that can be used to analyze the factors influencing the survival of ovarian cancer patients is through survival analysis. Survival analysis provides insights into the time required for a specific event to occur, such as death or disease recurrence. Survival analysis has significantly impacted clinical trials through the Kaplan-Meier method, log-rank statistics, and the Cox proportional hazards model, with future developments

https://doi.org/10.30605/likelihood.1040

expected across various fields (Fleming & Lin, 2000). Previous survival analysis research on ovarian cancer datasets was conducted using the Weibull distribution approach. That study concluded that the variables influencing patient outcomes were treatment history and patient age. Patients with a history of treatment had a fourfold higher likelihood of better recovery, and increasing age reduced the likelihood of recovery (Mukaromah, 2020). This study applies the Cox Proportional Hazards model to the ovarian cancer dataset, considering stratification based on factors that may influence survival, and develops a more complex model to provide more accurate estimates regarding patient survival predictions. The model's ability to provide estimates is evaluated using various comprehensive evaluation methods, with the aim of ensuring the accuracy and reliability of the prediction results generated by this model in the context of ovarian cancer.

Ovarian cancer is a type of gynecological cancer with a high mortality rate, particularly among patients with advanced-stage disease. Survival analysis is frequently used in cancer studies to evaluate factors that influence patient survival time. In the context of ovarian cancer, survival time refers to the duration from diagnosis or the start of treatment until a specific event occurs, such as death from cancer. Patient age is often associated with prognosis in ovarian cancer. Patient age is a significant prognostic factor, with patients under 50 years of age having a higher five-year survival rate compared to patients aged 75 and older (Sait et al., 2022). The patient's performance status, as measured by the Eastern Cooperative Oncology Group (ECOG) scale, is also an important predictor. This score reflects the patient's physical condition, with patients in better performance status typically having a higher likelihood of surviving longer (Oken et al., 2009). The type of treatment received by ovarian cancer patients—whether standard or experimental—plays a crucial role in influencing survival time. Clinical trials are often designed to compare the effectiveness of these two types of treatment in improving patient survival duration (Peto, et al, 1977).

Survival analysis is a method of data analysis that examines the time to the occurrence of a specific event, such as death, disease recurrence, or other relevant events, and serves to indicate the duration from the start of observation to the point at which the event in question occurs (Kartisonaki, 2016). The survival function serves as the primary foundation in survival analysis, as it provides important information regarding the distribution of survival times (Mondall, 2017). Suppose T is a continuous random variable representing survival time with probability density function f(t) and cumulative distribution function $F(t) = P(T \leq t)$. The survival function S(t) is then defined as the probability that an event has not yet occurred up to time t, namely:

$$S(t) = P(T \geq t) = 1 - F(t)$$

The hazard function is used to measure the probability that an individual will experience an event within a small time interval Δt, given that the individual has survived until time t. This function is defined as follows:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(\, t \leq T < t + \Delta t \mid T \geq t\,)}{\Delta_t}$$

This function can also be related to the survival function and the probability density function via the following relationship:

$$h(t) = \frac{S(t)}{f(t)}$$

Survival analysis requires a specialized statistical approach, particularly for handling censored data that is, data that provide only partial information regarding the time of an event (Kartisonaki, 2016).

Data distribution testing is a crucial step in statistical analysis to determine whether a dataset follows a specific theoretical distribution, such as the normal, exponential, or other distributions. One method frequently used for this purpose is the Anderson-Darling test (AD Test). This test is an extension of the Kolmogorov-Smirnov method, designed to be more sensitive to deviations in the tails of the distribution. The Anderson-Darling test compares the empirical cumulative

distribution function (ECDF) with the theoretical cumulative distribution function (CDF), placing greater weight on the tails of the distribution. The Anderson-Darling test statistic is defined as follows:

$$A^2 = -n - \frac{n}{1} \sum_{i=1}^{n} [(2i-1)(\boldsymbol{ln}F(X_i) + \boldsymbol{ln}(1 - F(X_{n+1-i})))]$$

where n is the number of data points and $F(X_i)$ is the theoretical cumulative distribution function at the value X_i.

The Kaplan-Meier method is a nonparametric statistical technique used to analyze survival or event-time data, such as time to death, system failure, or patient recovery. This method is frequently used in various fields, including medicine, engineering, and the social sciences, to evaluate survival time distributions and compare survival curves across groups. The Kaplan-Meier calculation formula is shown in the following equation:

$$S(t) = \prod_{ti \le t} (1 - \frac{d_i}{n_i})$$

where d_i is the number of events at a specific time t_i and n_i is the number of individuals still at risk immediately before t_i. This method estimates the survival function that is, the probability that a subject will survive until time using the formula . Here, t_i is the time of event, n_i is the number of events at time t_i, and n_i is the number of individuals still at risk before time t_i. The advantage of the Kaplan-Meier method lies in its ability to handle censored data, that is, data where the exact time of event is unknown. The Kaplan-Meier method also facilitates the visualization of survival probabilities using survival curves. The analysis is typically supplemented with a log-rank test to compare survival curves across groups. However, this method has limitations, such as the assumption of no confounding factors and the inability to analyze the effects of covariates. The Kaplan-Meier method has been widely applied, particularly in the medical field to evaluate treatment effectiveness or risk factors. For example, research served as the basis for the development of this method. The Kaplan-Meier method remains the primary tool in survival data analysis (Meier, 1958).

The log-rank test is a statistical method used to compare survival curves between two or more groups. This test is often used in conjunction with the Kaplan-Meier method in survival data analysis, particularly to determine whether differences in survival functions between groups are statistically significant (Altman et al., 2020). The log-rank test compares event-time distributions by testing the null hypothesis that there is no difference between groups in survival functions. The log-rank test statistic is calculated based on the difference between the observed and expected number of events at each time point (Harrington, 2021). The formula for the log-rank test statistic is provided, where the number of observed events in the first group is the number of expected events in the first group based on the pooled data. This test statistic follows a chi-square distribution with degrees of freedom equal to the number of groups minus one (Altman et al., 2020). The log-rank test formula is based on the chi-square test and is formulated as follows:

$$\chi 2 = \frac{\left(\sum(O_i - Ei)\right)^2}{\sum V_i}$$

The log-rank test has several advantages. It is simple and efficient, easy to calculate, and provides reliable results for survival data without requiring specific distribution assumptions. Additionally, the log-rank test is suitable for censored data because it is designed to handle survival data containing censoring. However, this method has limitations, including the assumption of proportional hazards, meaning that the hazard ratio between groups is assumed to remain constant over the observation period. Furthermore, this test is sensitive to sample size; with small sample

sizes, the log-rank test may be less sensitive in detecting significant differences (Mitchell et al., 2019). The log-rank test is frequently used in medical research to compare the effectiveness of treatments or therapies between patient groups. For example, clinical trials often use this test to evaluate differences in survival between patients receiving a new treatment and those receiving standard treatment (Altman, 2020). The Cox proportional hazards model is a regression method used to analyze the relationship between the time to a specific event and several independent variables or covariates.

$$h(t \mid X) = h_0(t) \cdot exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p)$$

This model estimates the event hazard $h(t \mid X)$ as the product of the baseline hazard $h_0(t)$ and the exponential of a linear combination of covariates. One of the key assumptions of this model is that the hazard ratio between individuals remains constant over time (proportional hazards). If this assumption is not met, the analysis can be adapted using the stratified Cox approach, which groups data based on specific covariates to eliminate their effects on the hazard, or using the extended Cox model, which introduces time-dependent covariates to capture time-dependent changes in risk.

The Stratified Cox model is used when the proportional hazards (PH) assumption is not met for one or more covariates. The data are divided into groups (strata) based on specific covariates that violate the PH assumption, so that the baseline hazard function $(h_0(t))$ differs for each stratum. In the Stratified Cox model, interactions can be included to evaluate the combined effects of variables on the hazard, allowing for more specific analysis within each stratum, although this requires a large sample size due to the complexity of the parameters. Without interactions, the model only performs stratification based on specific variables to address violations of the PH assumption without adding complexity, making it simpler and reducing the risk of overfitting, but it is unable to capture the combined effects of variables.

$$h(t \mid X) = h_{0k}(t) \cdot exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p)$$

where $h_{0k}(t)$ is a baseline hazard function that differs for each stratum k.

The Extended Cox model is used to address violations of the PH assumption by adding time-dependent covariates $(X(t))$. In the Extended Cox model, interactions can be included to evaluate the combined effects of the fixed covariates (X) and the time-dependent covariates $(X(t))$, providing deeper insights into changes in risk over time, although the model becomes more complex and difficult to interpret. Without interactions, the model simply adds time-dependent covariates to address the PH assumption; it remains capable of capturing changes in risk over time without increasing complexity, but does not assess the combined effects of the variables.

$$h(t \mid X, X(t)) = h_0(t) \cdot exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \gamma_1 X_1(t) + \cdots + \gamma_q X_q(t))$$

where $X(t)$ is a time-dependent covariate, and γ is the coefficient for the time-dependent covariate.

The hazard ratio (HR) is a statistical measure used in survival analysis, including the Cox Proportional Hazards model, to compare the risk of a specific event between two groups. The hazard ratio describes the magnitude of the relative risk associated with a one-unit change in a covariate or between two different groups. The hazard ratio can be calculated using the following formula:

$$HR = exp(\beta)$$

where β is the regression coefficient from the Cox model, representing the effect of the covariate on the log-hazard, and $\exp(\beta)$ is the hazard ratio (HR), which indicates the relative risk.

Interpretation:

- HR=1: There is no difference in risk between the compared groups.
- HR>1: The group with higher covariate values has a higher risk.
- HR<1: The group with lower covariate values has a lower risk.

The Likelihood Ratio (LR) test is a statistical method used to compare two models, typically a simpler model (without interactions or without certain additional parameters) with a more complex model (with interactions or additional parameters). This test aims to determine whether adding parameters to a model significantly improves the fit to the data. The LR is frequently used in regression analysis, including in the Cox Proportional Hazards model, to evaluate the effects of additional variables or interactions. The LR formula is as follows:

$$\Lambda = -2 \cdot ln\left(\frac{L_{complex}}{L_{simple}}\right)$$

The Λ statistic follows a chi-square distribution with degrees of freedom equal to the difference in the number of parameters.

Test Hypotheses:

- H0: The simple model is adequate.
- H1: The complex model is better.

If the p-value is greater than 0.05, the simple model is selected. Conversely, if the p-value is less than or equal to 0.05, the complex model is preferred. The likelihood ratio (LR) test is often used to evaluate the effects of interactions or violations of the Pareto-Hirschman (PH) assumptions. Adjustments can be made using the Stratified Cox model to stratify the data or the Extended Cox model to include time-dependent covariates.

Model evaluation in survival analysis aims to assess the extent to which a model can predict event times with good accuracy. The following is an explanation of the four evaluation methods used: the Concordance Index. The Concordance Index is a measure of discrimination that indicates how well a model can distinguish between two individuals with different event times. Its values range from 0.5 (random prediction) to 1.0 (perfect prediction). The formula for the CI is as follows:

$$C = \frac{Number\ of\ concordant\ pairs}{Number\ of\ comparable\ pairs}$$

Concordant pairs are those in which a higher risk prediction corresponds to an earlier time of event, and comparable pairs are all pairs of individuals except those with censoring, which cannot be compared. Dynamic-AUC is an Area Under the Curve (AUC)-based method used to evaluate a model's predictive ability at various time points. This method accounts for changes in risk over time and data censoring. The higher the Dynamic-AUC value, the better the model is at distinguishing individuals at time t. The formula for Dynamic-AUC is as follows:

$$AUC_t = \frac{\Sigma_{i \neq j} I(h_i > h_j) I(T_i > t) I(T_j \leq t)}{\Sigma_{i \neq j} I(T_i > t) I(T_j \leq t}$$

where h_i is the predicted hazard rate for individual t, T_i is the time of event, and I is an indicator variable (1 if the condition is met, 0 if not). The Brier Score quantitatively measures prediction error by calculating the mean squared difference between the predicted probability and the actual outcome. Its values range from 0 (perfect prediction) to 1 (poor prediction).

$$Brier\ Score(t) = \frac{1}{n}\sum_{i=1}^{n}[\delta_i \cdot \left(1 - \hat{S}_i(t)\right)^2 + (1 - \delta_i) \cdot \hat{S}_i(t)^2]$$

where $\hat{S}_i(t)$ is the predicted survival probability at time t for individual i, and δ_i is the event indicator (1 if the event occurs, 0 otherwise). The Integrated Square Error evaluates the quality of the prediction by calculating the average of the squared differences between the predicted survival function and the actual survival function, integrated over time. The ISE formula is as follows:

$$ISE = \int_0^{\tau} \left(S(t) - \hat{S}(t)\right)^2 dt$$

where τ is the maximum time (end of observation), $S(t)$ is the actual survival function, and $\hat{S}(t)$ is the predicted survival function.

Method

The data used in this study were obtained from the Ovarian Dataset, which is a compilation of data from four double-blind randomized clinical trials involving patients with advanced ovarian cancer. This dataset was developed as part of the Ovarian Cancer Meta-Analysis Project in 1991. Overall, this dataset includes 299 original records, but the number of valid records for the study is 297. This dataset is available from the R source, which is a widely used platform for statistical analysis and data science.

This study uses several variables from the Ovarian Dataset, where the dependent variable (Y) is Time, representing survival time or censoring time in days. The independent variables (X) include fustat (censoring status, with a value of 1 for deceased and 0 for alive/censored), age (patient age in years), resid. ds (presence of postoperative residual disease, with a value of 1 for absent and 2 for present), rx (type of treatment received, with a value of 1 for the standard treatment group and 2 for the experimental treatment group), and ecog.ps (ECOG performance score, which describes the patient's level of disability, with a value of 1 for better performance and 2 for worse performance). These variables are used to analyze the factors influencing the survival duration of patients with advanced-stage ovarian cancer.

The research steps include: a. testing data distribution, b. using non-parametric methods because the data do not follow a specific distribution, c. applying the Cox proportional hazards (PH) model as the approach, d. testing the assumptions of the Cox PH model, e. conducting stratified Cox PH and extended Cox PH analyses, f. testing for the presence or absence of interactions between variables, g. selecting the best model based on the likelihood ratio test, h. evaluating the model using the Concordance Index, Brier Score, and others, i. interpreting the HR of the best model.

Results and Discussion

Distribution Test

The statistical test used was the Anderson-Darling test on the survival time variable against several distributions, namely the Weibull, Exponential, Log-normal, Gaussian, and Log-logistic distributions. The null hypothesis (H_0) is that the data follow a specific distribution, and the alternative hypothesis (H_1) is that the data do not follow a specific distribution. Using the Easyfit software, all distributions yielded a p-value of 2.5018 at $\alpha = 0.05$; therefore, H_0 cannot be rejected, and the data do not follow a specific distribution.

Kaplan-Meier and Log-Rank Tests

Table 1. This is the title of your table

Column 1	Column 2	Column 3
Data 1	Data 4	Data 7
Data 2	Data 5	Data 8
Data 3	Data 6	Data 9
Total	Sum Column 2	Sum Column 3

The Kaplan-Meier method was used to estimate survival curves for all variables in the dataset. These estimates provide a visual representation of differences in survival rates between groups based on specific variables. To test the significance of these differences in survival rates between groups, the log-rank test was used, which statistically compares the survival curves.

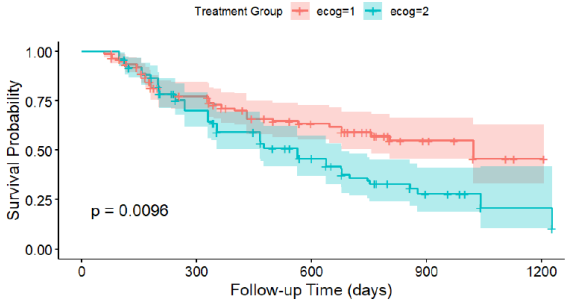

Figure 1. Kaplan-Maeier Plot Variable Ecog

The Kaplan-Meier plot in Figure 1 illustrates patient survival probabilities based on two Eastern Cooperative Oncology Group (ECOG) categories: ECOG=1 (red line) and ECOG=2 (blue line). The X-axis represents follow-up time (in days), while the Y-axis represents survival probability. The graph shows that patients with ECOG=1 have a higher survival probability compared to patients with ECOG=2 throughout the follow-up period. This indicates that patients with ECOG=2 have a poorer prognosis. The listed p-value (p=0.0096) indicates that the difference in survival between the two groups is statistically significant ($p<0.05$). The shaded area around the survival curve reflects the confidence interval, providing an indication of the uncertainty in the estimate. Overall, ECOG status was found to have a significant effect on patient survival outcomes.

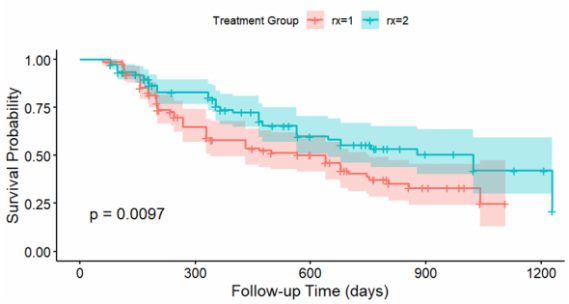

Figure 2. Kaplan-Maeier Plot Variablel Rx

This Kaplan-Meier plot shows patient survival probabilities based on two treatment groups: rx=1 (red line) and rx=2 (blue line). The X-axis represents follow-up time (in days), while the Y-axis shows survival probability. The graph shows that the probability of survival for patients in the rx=2 group tends to be higher than that of the rx=1 group throughout the follow-up period. The p-value ($p = 0.0097$) indicates that the difference in survival between the two groups is statistically significant ($p<0.05$). The shaded area around the survival curves reflects the confidence intervals, providing information about the uncertainty of the estimates. These results indicate that the type of treatment (rx) significantly affects patient survival outcomes.

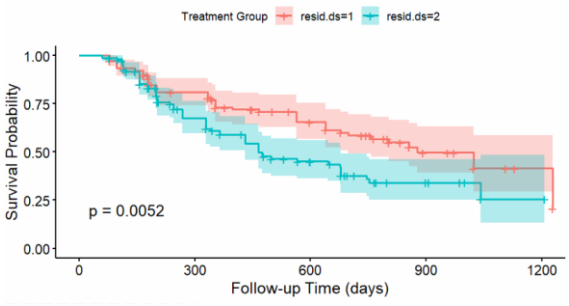

Figure 3. Kaplan-Maeier Plot Variable Resid.Ds

This Kaplan-Meier plot shows patient survival probabilities based on two residual disease categories (resid.ds): resid.ds=1 (red line) and resid.ds=2 (blue line). The X-axis represents follow-up time (in days), while the Y-axis represents survival probability. The graph shows that patients with resid.ds=1 have a higher survival probability compared to patients with resid.ds=2 throughout the follow-up period. The p-value (p=0.0052) indicates that this difference is statistically significant (p<0.05). The shaded area around the survival curve reflects the confidence interval, providing an indication of the uncertainty in the estimate. These results indicate that residual disease status significantly influences patient survival.

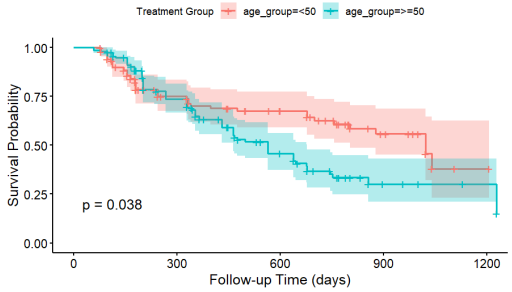

Figure 4. Kaplan-Maeier Plot Variable Age

This Kaplan-Meier plot shows patient survival probabilities based on two age groups: age <50 years (red line) and age ≥50 years (blue line). The X-axis represents follow-up time (in days), while the Y-axis depicts survival probability. The graph shows that patients aged <50 years have a higher survival probability compared to patients aged ≥50 years throughout the follow-up period. The p-value (p=0.038) indicates that this difference is statistically significant (p<0.05). The shaded area around the survival curve represents the confidence interval, which indicates the uncertainty of the estimate. These results suggest that age significantly influences patient survival prognosis.

Cox Propotional Hazards

The next analysis involves modeling using the Cox proportional hazards model. The parameter estimates for the ovarian cancer patient data are shown in the following table.

Table 2. Estimation of Cox Proportional Hazards Model Parameters

Parameter	Coef	SE	Z	p-value
age	0.023798	0.008835	2.694	0.00707
resid.ds	0.312512	0.201289	1.553	0.12053
rx	-0.316901	0.205241	-1.544	0.12258
ecog	0.206089	0.198733	1.037	0.29973

Based on Table 2, a Cox proportional hazards model can be formulated using the following equation:

$$h(t \mid X) = h_0(t) \cdot exp(0.023\ age + 0.312 resid.ds + (-0.137 rx + 0.206\ ecog)$$

Before conducting the parameter significance test, the Cox PH assumption test was first performed. The results of the parameter assumption test are shown in Table 2 with $\alpha = 0.05$.

Table 3. Estimation of Cox Proportional Hazards Model Parameters

Variable	Chi-square	Df	p-value
residual	0.595	1	0.08
age	3.062	1	0.44
rx	0.047	1	0.83
ecog	4.713	1	0.03
GLOBAL	6.928	1	0.14

Based on Table 3, the proportional hazards assumption is met for the residuals, the age variable, and the rx variable. The ecog variable violates the proportional hazards assumption with a p-value < 0.05; this is addressed using the stratified Cox and extended Cox approaches.

Stratified Cox Proportional Hazard

A stratified model analysis was performed on the ecog variable because it violated the proportional hazards assumption. The parameter estimates for the stratified model are shown in Table 4.

Table 4. Parameter Estimates for the Stratified Cox Proportional Hazards Model

Parameter	Coef	SE	Z	p-value
age	0.023772	0.008827	2.693	**0.00708**
resid.ds	0.328712	0.201119	1.634	0.10217
rx	-0.334400	0.205149	-1.630	0.10309

Based on Table 4, a stratified Cox proportional hazards model can be formulated using the following equation:

$$h(t \mid X) = h_0(t) \cdot exp(0.023\ age + 0.328 resid.ds + (-0.137 rx + 0.206\ ecog)$$

Extended Cox Proportional Hazard

An extended model analysis was performed on the ecog variable because it violated the proportional hazards assumption. The parameter estimates for the extended model are shown in Table 5.

Table 5. Parameter Estimates for the Extended Cox Proportional Hazards Model

Parameter	Coef	SE	Z	p-value
age	0.023525	0.008761	2.467	0.0136
resid.ds	0.301971	0.200046	1.605	0.1083
rx	-0.331149	0.205042	-1.719	0.0857
ecog	-0.311234	1.542346	-2.168	0.0301
logtecog	0.632610	0.275516	2.335	0.0196

Based on Table 3, the extended Cox proportional hazards model can be formulated as follows:

h(t|X)=h_0 (t)·exp(0.023 age+0.301971resid.ds-0.137rx-0.311234ecog+0.632610logtecog)

Likelihood Ratio Test

In this section, the models were evaluated using the Likelihood Ratio (LR) test to determine whether the models with interactions provided a significant improvement over the models without interactions. This test was applied to the Stratified Cox and Extended Cox models to ensure the selection of the optimal final model.

Table 6. Parameter Estimates for the Extended Cox Proportional Hazards Model

Model	p-value	Conclusion
Stratified Cox	0.8201	H_0 accepted
Extended Cox	0.7833	H_0 accepted

The results of the Stratified Cox test showed a p-value of 0.8201 (P > 0.05), so the null hypothesis which states that the model without interaction is just as good as the model with interaction was accepted. There was no significant improvement in model fit when interaction was added. Thus, the model without interaction was selected as the final model. Furthermore, in the Extended Cox model, the resulting p-value was 0.7833 (P > 0.05). Similar to the results in the Stratified Cox model, the null hypothesis was accepted. The model without interaction was again selected because there was no significant improvement in data fit when interaction was included. The conclusion from this LR test is that in both the Stratified Cox and Extended Cox models, the model without interaction provides simpler yet still optimal results. Therefore, the model without interaction is used as the final model in this analysis.

Model Evaluation

An evaluation of the performance of survival prediction models using the Stratified Cox model without interactions and the Extended Cox model without interactions. The evaluation was conducted using four methods, as shown in Table 7.

Table 7. Model Evaluation

Method	*Stratified Cox*	*Extended Cox*	Interpretation
CI	**0.5631**	0.3126	The Stratified model has a higher CI of 56.31%, making it superior to the Extended Cox model.
AUC(t)	**0.612**	0.248	The AUC(t) value for the Stratified model is 61.2%, making it significantly better than the Extended model, which has a value of 24.8%.
BSC	**0.2896**	0.515	The Stratified model has a BSC value closer to 0, specifically 0.2896, making the Stratified model superior.
ISE	14.4754	**6.0667**	The Extended model has a smaller ISE value, specifically 6.0067, making the Extended model superior.

Based on Table 7, which presents model evaluations using CI, Dynamic AUC, BSC, and ISE, the Stratified Cox model demonstrated better evaluation scores across all three metrics, indicating that the Stratified Cox model outperforms the Extended Cox model.

Hazard Ratio

As the final step in this analysis, the results from the best-performing model are presented. The following table shows the HR estimates for the variables included in the best model.

Table 8. Hazard Ratio from the Stratified Cox Model.

Parameter	*Hazard Ratio*	p-value	Significance
age	1.024	**0.007**	Significant
resid.ds	1.389	0.102	Not Significant
rx	0.716	0.103	Not Significant

Parameter Explanation, age (0.024): For every 1-year increase in a patient's age, the risk of an event increases by 2.4%. Resid.ds (0.329): Patients with residual disease have a 38.9% higher risk of an event compared to patients without residual disease. rx (-0.335): Experimental treatment reduces the risk of event by 28.4% compared to standard treatment.

Based on the analysis results, only the age variable has a significant effect on the risk of events. The resid.ds and rx variables are not statistically significant, so they have no real effect on the risk of events in this Stratified Cox model. The Hazard Ratios in TABLE 8 can be modeled in the following equation.

$$h(t \mid X) = h_0(t) \cdot exp(0.024 \cdot age + 0.329 \cdot resid.ds2 - 0.335 \cdot rx)$$

Conclusion

Based on the results of the analysis and evaluation of the model using various methods, the Stratified Cox model without interactions was selected as the best model. This selection was supported by the results of the Likelihood Ratio test, which showed that adding interactions to the model did not significantly improve the fit of the data. With a p-value > 0.05 for the Stratified Cox model, the null hypothesis—stating that the model without interactions is as good as the model with interactions—was accepted. The Stratified Cox model without interactions provides simpler yet still accurate results in explaining the relationship between the study variables and the risk of events, and was therefore selected as the final model. The analysis results indicate that patient age is significant for the risk of death in patients with ovarian cancer; each additional year of age increases the risk of death by approximately 2.4%. Although the variables of treatment type and the presence of residual disease post-surgery were not significant for the risk of death, the analysis results showed that patients receiving experimental treatment had a 28.4% reduction in the risk of death, and patients with residual disease had a 38.9% higher risk of death compared to patients without residual disease.

Acknowledgment

-

Bibliography

Bradburn, M. J., Clark, T. G., Love, S. B., & Altman, D. G. (2003). Survival analysis part II: multivariate data analysis–an introduction to concepts and methods. *British journal of cancer*, *89*(3), 431-436.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 68(6), 394-424. https://doi.org/10.3322/caac.21492

Coburn, S. B., Bray, F., Sherman, M. E., & Trabert, B. (2017). International patterns and trends in ovarian cancer incidence, overall and by histologic subtype. International journal of cancer, 140(11), 2451-2460. https://doi.org/10.1002/ijc.30676

D. W. Hosmer, S. Lemeshow och S. (2011), Applied Survival Analysis: Regression Modeling of Time-to-Event Data, NJ: John Wiley & Sons.

Du Bois, A., Reuss, A., Pujade-Lauraine, E., Harter, P., Ray-Coquard, I., & Pfisterer, J. (2009). Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers de l'Ovaire (GINECO). Cancer, 115(6), 1234-1244. https://doi.org/10.1002/cncr.24149

Fleming, T. R., & Lin, D. Y. (2000). Survival analysis in clinical trials: past developments and future directions. Biometrics, 56(4), 971-983.

Harrington, D. P., & Fleming, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika*, *69*(3), 553-566.. https://doi.org/10.1093/biomet/69.3.553

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. Journal of the American statistical association, 53(282), 457-481. https://doi.org/10.1080/01621459.1958.10501452

Kartsonaki, C. (2016). ”Survival analysis,” Diagnostics, 6(1), 11

M. Mondal och L. C. P, ”, "A framework for survival analysis based on modeling of cumulative hazard function using a hybrid neural network,” Expert Systems with Applications, vol. 85, pp. 106-118, 2017.

Momenimovahed, Z., Tiznobaik, A., Taheri, S., & Salehiniya, H. (2019). Ovarian cancer in the world: epidemiology and risk factors. International journal of women's health, 287-299. https://doi.org/10.2147/IJWH.S197604

Mukaromah, M. (2020). Analisis Survival pada Data Kanker Ovarium. MATHunesa: Jurnal Ilmiah Matematika, 8(2), 130-134.

Oken, M. M., Creech, R. H., Tormey, D. C., Horton, J., Davis, T. E., McFadden, E. T., & Carbone, P. P. (1982). Toxicity and response criteria of the Eastern Cooperative Oncology Group. American journal of clinical oncology, 5(6), 649-656.

Peto, R., Pike, M., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., ... & Smith, P. G. (1977). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. British journal of cancer, 35(1), 1-39.

Pocock, S. J., Clayton, T. C., & Altman, D. G. (2002). Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *The Lancet*, *359*(9318), 1686-1689.

S. P. Wright, C. B. Langman och C. A. Mitchell (2019). ” Proportional hazards assumption in survival analysis: Testing and consequences,” Journal of Applied Statistics. 46(5). 834–845,.