ACTA MÉDICA PORTUGUESA
A Revista Científica da Ordem dos Médicos

Cox Regression in Survival Analysis: Practical Insights for Clinicians

Regressão de Cox em Análise de Sobrevida: Fundamentos Práticos para Clínicos

António GOMES [1], Bruna COSTA [2], Vitor NUNES [1], Constança COELHO [3]
Acta Med Port (In Press) ▪ https://doi.org/10.20344/amp.23078

ABSTRACT

Survival analysis is a fundamental tool in clinical research for evaluating time-to-event outcomes. While the Kaplan-Meier method remains a widely used univariable approach for estimating survival probabilities and comparing groups, it does not account for multiple risk factors simultaneously. To address this limitation, multivariable regression models are employed, with the Cox proportional hazards model (Cox regression) being the most commonly used. This paper provides a practical guide to Cox regression for clinicians, emphasizing its application in survival analysis rather than focusing on mathematical derivations. We discuss key concepts, including hazard ratios, model assumptions, variable selection, and interpretation of results. Additionally, we explore essential methodological considerations, such as assessing proportional hazards assumptions, handling missing data, and avoiding overfitting. By offering a step-by-step approach to implementing Cox regression in clinical research, this article aims to enhance understanding and improve the quality of survival analysis in medical studies. Practical examples illustrate how to interpret Cox regression results and their relevance in clinical decision-making.
Keywords: Investigative Techniques; Models, Statistical; Proportional Hazards Models; Regression Analysis; Survival Analysis

RESUMO

A análise de sobrevivência é uma ferramenta fundamental na investigação clínica para avaliar o tempo até à ocorrência de um acontecimento de interesse. Embora o método de Kaplan-Meier seja amplamente utilizado para estimar as probabilidades de sobrevivência e comparar grupos, este não permite avaliar simultaneamente múltiplos fatores de risco. Para contornar essa limitação, recorre-se a modelos de regressão multivariada, sendo o modelo de regressão de riscos proporcionais de Cox (regressão de Cox) o mais utilizado. Este artigo apresenta-se como um guia prático sobre a aplicação clínica da regressão de Cox, privilegiando a sua utilização na análise de sobrevivência, em detrimento de abordagens matemáticas mais complexas. São discutidos conceitos essenciais, incluindo a razão de risco, os pressupostos do modelo, a seleção de variáveis e a interpretação dos resultados. Adicionalmente, são abordadas questões metodológicas cruciais, como a verificação da suposição de riscos proporcionais, a gestão de dados omissos e a prevenção do *overfitting* do modelo. Através de uma abordagem passo a passo, este artigo visa melhorar a compreensão e a aplicação da regressão de Cox na investigação clínica. Exemplos práticos ilustram a interpretação dos resultados e a sua relevância para a tomada de decisão clínica.
Palavras-chave: Análise de Regressão; Análise de Sobrevivência; Modelos Estatísticos; Modelos de Riscos Proporcionais; Técnicas de Investigação

INTRODUCTION

Survival analysis is a fundamental tool in clinical research for evaluating outcomes over time. This paper is a follow-up to the article “Kaplan-Meier Survival Analysis: Practical Insights for Clinicians”,[1] which explored Kaplan-Meier analysis as one of the most commonly used methods in clinical research for univariate analysis.

While Kaplan-Meier analysis provides an intuitive approach for estimating survival probabilities and comparing different groups, its primary limitation is that it only allows for the assessment of one variable at a time, making it insufficient for evaluating complex interactions between multiple risk factors.

To address this limitation, a multivariable analysis is typically conducted after descriptive and univariable analyses to assess multiple variables (factors) that simultaneously best explain the outcomes (Fig. 1A, B, C). At this stage, regression models play a crucial role in this process by identifying significant variables and quantifying the magnitude of their influence (Fig. 1D).

By incorporating multiple covariates, such as age, treatment type, and comorbidities, the Cox proportional hazards model (commonly referred to as Cox regression) provides a more comprehensive understanding of survival dynamics compared to simpler methods. Its ability to adjust for confounders and provide an interpretable hazard ratio makes it an essential tool in clinical research, particularly for applications such as evaluating the effectiveness of interventions and identifying risk factors for disease progression.[2]

Besides Cox regression, other methods, such as the accelerated failure time model and the competing risks model, have alternative approaches with specific assumptions and applications.[3,4] However, these are rarely encountered in the medical literature, and so the focus of this paper will be on the Cox regression model, highlighting its practical applications in clinical research.

Despite its widespread use, many clinicians do not fully make use of its applications or understand its interpretation. This paper aims to bridge that gap by providing a clear and practical guide to multivariable analysis, specifically, how to

1. Surgery Department. Hospital Fernando Fonseca. Amadora. Portugal.
2. Champalimaud Centre for the Unknown. Lisbon. Portugal.
3. Genetics Laboratory. Faculdade de Medicina. Universidade de Lisboa. Lisbon. Portugal
✉ **Autor correspondente:** António Gomes. antonio.p.gomes@ulsasi.min-saude.pt
Recebido/*Received*: 13/03/2025 - **Aceite/*Accepted*:** 04/12/2025 - **Publicado *Online*/*Published Online*:** 13/02/2026

interpret these methods, recognize their alternatives, and be aware of their limitations. Rather than focusing on mathematical considerations, we will emphasize study design, result interpretation, and the implementation of Cox regression and multivariable survival analysis in clinical investigation, providing healthcare professionals with the necessary skills to apply these methods effectively.

METHODS

Understanding the Cox proportional hazards model

Consider a hypothetical study aimed at evaluating the role of the variable *X2* gene expression in gastric cancer disease progression. The collected data for all included gastric cancer patients were: *X2* expression, age at diagnosis, stage of the disease, resection margin, histological type, date of diagnosis, date of surgery, date of disease recurrence diagnosis (if applicable), date of death due to disease progression (if applicable), and the date of last observation.

In this context, the analysis examines how these different factors influence the likelihood of an event occurring over time, such as cancer recurrence or death. It estimates the risk of the event based on:

- The baseline hazard function (baseline risk level): this represents the starting probability of the event occurring when no other factors are considered. Although it is mathematically essential, it often lacks direct clinical interpretation and is not the scope of most studies.
- Covariates (predictor variables): these are patient characteristics used in the model, such as age, stage, margin, etc., that may influence the event. Careful consideration should be given to the number of covariates included in the regression model. Overfitting and reduced statistical power can occur when too many variables are used relative to the number of events, ultimately compromising the model's accuracy. Covariate selection is typically based on univariable analysis results, prioritizing variables with statistically significant associations with the outcome and/or strong biological plausibility.
- Regression coefficients: these estimate the effect of each covariate on the event. In this case, these coefficients quantify the influence of age at diagnosis, stage of the disease, resection margin, and histological type on time to cancer recurrence or time to death due to disease progression.

The clinical relevance of Cox regression arises from these regression coefficients.

Statistical software packages, such as SPSS, STATA, R, and Python, typically calculate the hazard ratio (HR) for each variable.

The HR compares the risk of an event occurring in one group to the risk of the same event in another group. An HR for dichotomous variables compares the instantaneous risk of an event between two groups over time. Usually, an "exposed" (or treatment) group and an "unexposed" (or control) group. An HR of 1.0 indicates no difference in risk, an HR > 1.0 means the exposed group has a higher risk, and an HR < 1.0 means the exposed group has a lower risk.

For example, in a clinical trial comparing two treatment regimens, an HR of 0.5 would suggest that patients receiving the new treatment have half the risk of the event occurring compared to those receiving the standard treatment. In Cox regression multivariable models, an HR of 1.5 for a given covariate means that a one-unit increase in that covariate is associated with a 50% increase in the hazard (or risk) of the event happening at any given time. For example, if the covariate is age (in years) and HR = 1.5, then for every one-year increase in age, the risk of recurrence increases by 50% (hypothetical scenario). If the covariate is sex, a reference category must first be defined (*e.g.*, female). In this case, an HR = 1.5 for male patients indicates that male patients have a 50% higher risk of the event occurring compared to female patients at any given time.

- Covariates with HR > 1 are generally associated with an increased risk of the event of interest, such as cancer recurrence or death occurring earlier in time, compared to a reference group.
- Covariates with HR < 1 are generally associated with a decreased risk of the event of interest, compared to a reference group. HR < 1 means that the covariate is a "protective factor".

Taking as an example the study "The Exploration of Surgery and Survival Prediction in Patients with Peritoneal Metastasis from Gastric Adenocarcinoma Based on the Surveillance, Epidemiology, and End Results (SEER) Database",[5] the authors preliminarily evaluated the possible treatment efficacy of surgical resection. For this purpose, the authors analyzed a dataset of 399 patients with peritoneal metastases from gastric carcinoma selected from the SEER database (2000 - 2022). Following a descriptive analysis, they presented Kaplan-Meier curves depicting overall survival according to each variable (univariable analysis).

Figure 2, reproduced from Shen Y *et al*,[5] illustrates a Kaplan-Meier curve for overall survival stratified by tumor grade

and T stage. Panels A and B depict survival according to surgical treatment in patients with low-grade (grade I - II) and high-grade (grade III - IV) tumors, respectively. Panels C and D show analogous comparisons stratified by early (T1 - T2) and advanced (T3 - T4) primary tumor stage. Across all strata, patients undergoing surgery demonstrate improved overall survival compared with those managed without surgery, with statistically significant differences observed on log-rank testing. The dashed lines indicate the time point at which 50% of the initially included patients within each age group have died (i.e., experienced the event), representing the median overall survival. These subgroup analyses illustrate how Kaplan–Meier curves allow intuitive visual assessment of survival differences within clinically relevant categories. However, despite the valuable descriptive information provided by Kaplan–Meier analysis and the quantitative comparison afforded by the log-rank test, neither approach quantifies the magnitude of effect nor accounts for the simultaneous influence of multiple prognostic variables. To assess the independent impact of each factor on survival and to adjust for potential confounding, multivariable analysis is required. It is in this context that Cox proportional hazards regression assumes a central role in clinical research.

Table 1 presents the results of a multivariable Cox regression analysis for the aforementioned cohort. This table includes the variables: age (categorized), sex, race, primary tumor staging and characteristics, and treatment-related variables, (reproduced from Shen Y *et al*).[5]

For each variable, the HR, 95% confidence interval (95% CI), and corresponding *p*-value are reported. Using age as an example, the 20 - 39 age group serves as the reference, with survival outcomes of all other age groups compared against it. An HR > 1 (as observed for the ≥ 80 years age group) indicates a higher risk of the event (death) relative to the reference. Conversely, HR < 1 (as observed for the 40 - 59 and 60 - 79 age groups) suggests a lower risk of death compared to the reference group.

Beyond HR values, interpretation of the 95% CI is essential. If the 95% CI includes the number one, the result is not statistically significant. In this case, although the HRs for the 40 - 59 and 60 - 79 age groups are < 1, their 95% CIs include 1. Consequently, these results are not statistically significant ($p > 0.05$), meaning these age groups cannot be confidently considered protective factors for survival.

Regarding the variable “surgery”, where “no surgery” is the reference category, the HR for patients undergoing surgery is 0.438 (95% CI: 0.301 - 0.632), which does not include 1, with a corresponding *p*-value of 0.001. Similar considerations apply to radiotherapy and chemotherapy.

Based on these findings, surgery, radiotherapy, and chemotherapy appear to be independent protective factors for survival. However, the discussion of the clinical significance of these results is beyond the scope of this paper.

Why Cox regression?

The Cox regression was first introduced by the British statistician Sir David Cox in 1972.[6] It revolutionized multivariable survival analysis, being, until today, the most widely used statistical method in this field:

- Cox regression does not assume a specific form for the baseline hazard function. We can modulate and predict survival data without knowing the baseline hazard risk.
- Cox regression deals with censored data (patients that enter after the beginning of the study, patients that are lost to follow-up)[1] with minimal bias.
- Cox regression introduced the concept of the hazard ratio (explained above), which quantifies how much a given covariate increases or decreases the risk of an event occurring at any point in time. This became a standard metric in survival analysis and medical statistics.

However, there are some restrictions (assumptions) to bear in mind: the Cox regression assumes that covariates act multiplicatively on the baseline hazard, meaning that each variable is associated with a constant proportional change in the hazard rather than an additive effect. In addition, the model assumes that these effects remain constant over time, corresponding to the proportional hazard’s assumption.

Step-by-step guide to implementing Cox regression

A. Definition of variables and data collection

The definition and coding of variables are essential to the success of the study. If not carried out correctly before data collection, it may compromise the entire work. For conducting a Cox regression, it is essential to clearly define the event under study. For each study participant, key information must be defined, including the time of study entry (relative to the study start), whether the event occurred, the time of event occurrence (if applicable), and the time of last observation or

study exit in the absence of an event. Additionally, it is essential to precisely define, code, and input covariate data at the study's outset.

B. Descriptive statistics and univariable analysis

This section typically includes Kaplan-Meier plots for covariates, where applicable. The approach will depend on the study's objective and the type and number of covariates involved.

C. Selection of variables for the multivariable model

The selection of covariates may vary depending on the study's objective and design. Generally, variables that show statistical significance in univariable analysis are included, along with those that have clinical relevance, even if they do not reach statistical significance in the univariable analysis.

D. Model building

Several software packages are available for model building. STATA and SPSS are among the most widely used. The survival package in R and the lifelines library in Python (as well as scikit-survival and statsmodels libraries) are also commonly applied. It is important to note that model robustness and power decrease as the number of covariates increases relative to the number of events. Similarly, categorical variables with many categories, especially when they involve small numbers of participants, can lead to data fragmentation and compromise accuracy, power, and potentially cause overfitting. A simple, generic strategy is to consider different subsets of three or four covariates (the number of variables depends on sample size and event count) and select the model that shows the best robustness.

E. Assessing model performance

Cox regression combines an unspecified baseline hazard with a parametric component that estimates the effect of covariates on the hazard. In practical terms, the model does not require specification of the underlying baseline risk of the event over time, but instead focuses on estimating relative effects between groups. For example, in a hypothetical study evaluating the effect of adjuvant chemotherapy on survival in patients with resected gastric cancer, the baseline hazard represents the underlying risk of death over time that would apply to all patients if no covariates were considered. Unlike fully parametric survival models, such as exponential or Weibull models – which assume a specific mathematical form for the baseline hazard – Cox regression leaves this baseline risk unspecified. The model therefore concentrates on estimating the hazard ratio between patients who did and did not receive chemotherapy, assuming that this relative effect remains constant over time (the proportional hazards assumption).[6] As a semi-parametric model, Cox regression is typically evaluated using likelihood-based measures and information criteria to compare different model specifications, rather than through tests of absolute goodness of fit. One of the most widely used approaches is the likelihood ratio test, which compares nested models – that is, models in which a simpler model is fully contained within a more complex one – by evaluating differences in their log-likelihoods, a measure of how well a given model explains the observed data. A larger chi-square (χ^2) statistic from the likelihood ratio test indicates that the more complex model explains additional variation in the data relative to the simpler model, while a p-value < 0.05 suggests that this improvement in model fit is statistically significant.[6,7]

In Appendix 1, Table 1, (https://www.actamedicaportuguesa.com/revista/index.php/amp/article/view/23078/15892) reproduced from Zhenguo Wu *et al*,[7] four Cox regression models are compared using several complementary metrics, including the χ^2 statistic, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC). While the χ^2 statistic reflects improvement in model fit, AIC and BIC additionally penalize model complexity, favoring models that achieve a better balance between explanatory power and parsimony. Lower AIC and BIC values therefore indicate superior relative model performance.

In this analysis, *"Model 3 + TyG + SUA"* demonstrated the most favorable overall performance across these criteria compared with the other candidate models. Other methods for model evaluation may be employed depending on the study's aims, including discrimination measures (*e.g.*, Harrell's C-index) and calibration assessments.

F. Verify the proportional hazards model assumptions

Cox regression, like other statistical models, relies on a set of underlying assumptions to ensure the validity of its results. A key assumption is the proportional hazards (PH) assumption, which states that the relative risk (hazard ratio) of an event between any two groups or individuals with different covariate levels remains constant over time. Violations of this

assumption may compromise both the internal and external validity of the model. The PH assumption can be assessed using both graphical and statistical approaches, which are routinely implemented in standard statistical software packages. Graphical methods include log[-log($S(t)$)] survival plots, which compare survival functions across covariate groups – if the curves are approximately parallel, the PH assumption is likely valid, whereas crossing or diverging curves suggest potential violations [Appendix 1, Fig. 1 (Appendix 1: https://www.actamedicaportuguesa.com/revista/index.php/amp/article/view/23078/15892)].[8]

Schoenfeld residual plots offer another visual approach: Schoenfeld residuals are calculated as the difference between the observed and expected covariate values at each event time, and they should display no systematic pattern when plotted against time if the PH assumption holds [Appendix 1, Fig. 2 (Appendix 1: https://www.actamedicaportuguesa.com/revista/index.php/amp/article/view/23078/15892)].[9,10]

For a more formal assessment, the Grambsch-Therneau test examines the relationship between Schoenfeld residuals and time; a statistically significant result (*e.g.*, p-value < 0.05) indicates a violation of the PH assumption for the corresponding covariate.[11]

G. Time-dependent covariates

In scenarios where a covariate's value changes during follow-up or its effect on the hazard is not constant over time, the covariate is considered time-dependent.[12,13] Both situations violate the proportional hazards assumption of the Cox regression model and require explicit modelling to avoid biased hazard ratio estimates.[10,11]

Consider the following simple scenario, illustrated in Fig. 3, in which three patient groups (A, B, and C) are followed over time. The underlying recurrence hazard changes at approximately 12 and 24 months, resulting in the crossing of Kaplan-Meier curves (Fig. 3A). Initially, during follow-up, group B has a lower hazard than group A. Still, this effect decreases and then reverses after two years. Group C has intermediate early risk and improves late. This pattern is common when an initial treatment effect wanes over time or late adverse effects emerge.

To test this formally, the proportional hazards assumption is examined using Schoenfeld residuals (Fig. 3B). The residual plot shows a clear upward trend (instead of fluctuating randomly around zero), and the associated Grambsch-Therneau test is statistically significant (p-value < 0.05).

Once time-dependence is detected, there are two main strategies:

1. Piecewise modelling: divide the follow-up into meaningful periods (*e.g.*, 0 - 12, 12 - 24, > 24 months) and allow the hazard ratio to vary between periods. This approach is intuitive and produces period-specific hazard ratios, as shown in Fig. 3C. Group B *versus* A is protective in the first year, neutral in the second, and harmful after two years. This is often the most accessible approach for clinicians and is especially useful when the change points have biological or clinical meaning.[13,14]
2. Continuous time-covariate interactions: when changes are gradual rather than abrupt, or when pre-specified periods are not obvious, a more flexible strategy is to include an interaction between the covariate and a function of time [*e.g.*, covariate × log(time)] in the Cox regression model. This extended model allows the hazard ratio to evolve smoothly over time without artificially imposing cut-points.

A practical clinical example is a trial in which some patients started a second-line therapy several months after enrolment. If this treatment is included in the model as a fixed baseline variable, each patient contributes follow-up time both before and after starting therapy, but the model treats them as if they had been on treatment from day one. This creates what is known as immortal time bias, because the period before treatment initiation is incorrectly attributed to the "treated" group, leading to distorted hazard ratio estimates.[15] The correct approach is to treat treatment status as a time-dependent covariate. Each patient's follow-up is divided into two intervals – before and after starting the second-line therapy – so that their treatment status is accurately represented at each time point. Technically, this is done by splitting follow-up time into multiple intervals per patient and using the start-stop format (also called counting process notation) in the Cox model. In each interval, the covariate values are updated to reflect the patient's current status.[14,16]

Assessing influential observations and outliers

Outliers or highly influential observations can distort hazard ratio estimates. DFBETA values measure the change in regression coefficients when an observation is removed (compares the model with and without a particular observation to see how it changes): large DFBETA values indicate observations that disproportionately influence the model [Appendix 1,

Fig. 3 (Appendix 1: https://www.actamedicaportuguesa.com/revista/index.php/amp/article/view/23078/15892)].[17]

Dealing with missing data and informative censoring

Missing values are inevitable in real-world clinical studies. Traditional approaches like complete case analysis may introduce bias, particularly if missingness is related to prognosis. Techniques such as multiple imputation or inverse probability weighting can help mitigate these biases, but their implementation requires careful consideration to avoid introducing new sources of error.

Informative censoring occurs when patients with severe disease have a higher probability of dropping out, while healthier patients are more likely to remain in the study. This results in bias toward better survival outcomes, as the likelihood of a subject being censored is related to their risk of event. This remains a challenge that standard Cox regression does not inherently address.[18]

Sample size calculation in Cox regression models: practical insights

Sample size calculation in Cox regression analysis is essential to ensure adequate statistical power for detecting meaningful associations between covariates and survival outcomes.

Unlike traditional statistical methods, survival analysis is event-driven, meaning the required sample size depends primarily on the number of observed events rather than the total number of participants. A commonly used guideline suggests at least 10 events per predictor variable to reduce the risk of overfitting.[19-21] Several factors influence sample size determination, including expected event rates, effect sizes, follow-up duration, and potential loss to follow-up, all of which impact the precision of the hazard ratio.[19-21] Several freely available online tools can be used to estimate sample size under these assumptions, including web-based calculators implementing Schoenfeld's formula (*e.g.*, QuesGen, powerandsamplesize.com) as well as open-source statistical software such as R, which allows both analytical and simulation-based approaches. Schoenfeld's formula is widely used for studies involving proportional hazards assumptions, while simulation-based methods are particularly useful in complex study designs.[22-24] When reading a paper, one must take this into account to know how to translate the findings into clinical practice. While clinical trials must include a sample size calculation in their methods, non-interventional studies often lack this information.

As an example, in the study "Risk and Predictors of Dementia and Parkinsonism in Idiopathic REM Sleep Behaviour Disorder: A Multicentre Study",[25] the authors explicitly state in the methods section how they estimated the sample size for a future neuroprotective trial. Their calculation was based on a categorical definitive endpoint (disease phenoconversion), with two groups (placebo *versus* a single-dose treatment), a two-sided alpha = 0.05, and 80% power. They used time-to-event analysis (http://www.quesgen.com/SSSurvival.php) for a two-year trial, assuming an agent that reduces phenoconversion with HR = 0.5. The sample size was calculated for the entire population, as well as stratified by prodromal marker testing, using directly observed conversion rates, and adjusting for center effects based on the current study's hazard ratio. In the results section, the authors showed the differences in sample size calculations according to different assumptions of the previously mentioned factors.

Machine learning, deep learning, and the future of survival analysis

Cox regression remains a fundamental tool in clinical research, but its application in modern datasets presents practical challenges that clinicians and researchers must be aware of.

Recent advancements in machine learning (ML) and deep learning (DL) offer powerful alternative survival models, but they lack interpretability.[26-28]

A practical approach is not to replace Cox regression but to enhance it through hybrid modeling. Artificial intelligence-driven methods can assist in feature selection, interaction detection, and non-linear modeling while maintaining a Cox-based framework for interpretability.

For example, machine learning can identify clusters of patients with distinct survival patterns, which can then be analyzed using a Cox model to generate clinically meaningful hazard ratios.

Deep learning in particular has shown promise in survival analysis through models such as DeepSurv, which extends the Cox proportional hazards model by using neural networks to learn complex, non-linear relationships between covariates and survival outcomes. Unlike traditional Cox regression, DeepSurv does not require the assumption of proportional hazards and can adaptively model interactions without explicit specification. However, its major limitations include the need for large datasets, potential overfitting, and reduced interpretability.[26-29]

One promising direction is combining deep learning with Cox regression to balance predictive accuracy and clinical

interpretability. For instance, neural networks can be used for feature extraction, transforming raw data into meaningful representations that can then be analyzed using a Cox model. This hybrid approach preserves the advantages of both methods, allowing for better risk stratification while maintaining the ability to report hazard ratios that are clinically actionable.

For example, in the article "Explainable Machine Learning Can Outperform Cox Regression Predictions and Provide Insights in Breast Cancer Survival",[30] the authors compare the performance of Cox regression and different hybrid models. The authors conclude that, based on their data, ML-based approaches can perform as well as the conventional Cox regression model or, in the case of the XGB model, even better. However, this improvement comes with increased complexity and reduced interpretability.[30]

CONCLUSION

Cox regression is a powerful tool for analyzing survival data, allowing researchers to account for multiple covariates simultaneously.

This article aims to simplify the key aspects of its application, including its potential, limitations, and pitfalls. It is essential to identify which points need to be reported so that others can critically assess the applicability of our results, as well as to develop the ability to analyze and apply findings from the wider scientific community. Recognizing the inherent limitations associated with each method and study is fundamental to practicing evidence-based medicine responsibly and ensuring that academic research remains both valuable and sustainable.

ACKNOWLEDGMENTS

The authors declare that Chat GPT was employed to improve English grammar and clarity.

AUTHOR CONTRIBUTIONS

AG: Study conceptualization, writing of the manuscript.
BC: Critical review of the manuscript.
VN: Critical review of the manuscript, supervision.
CC: Study conceptualization, critical review of the manuscript, supervision.
All authors approved the final version to be published.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

FUNDING SOURCES

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Gomes AP, Costa B, Marques R, Nunes V, Coelho C. Kaplan-Meier survival analysis: practical insights for clinicians. Acta Med Port. 2024;37:280-5.
2. Bugnard F, Ducrot C, Calavas D. Advantages and inconveniences of the Cox model compared with the logistic model: application to a study of risk factors of nursing cow infertility. Vet Res. 1994;25: 134-9.
3. Haushona N, Esterhuizen TM, Thabane L, Machekano R. An empirical comparison of time-to-event models to analyse a composite outcome in the presence of death as a competing risk. Contemp Clin Trials Commun. 2020;19:100639.
4. Swindell WR. Accelerated failure time models provide a useful statistical framework for aging research. Exp Gerontol. 2009;44:190-200.
5. Shen Y, Yu S, Zhang Y, Wang S, Chen J, Xu Y, et al. The exploration of surgery and survival prediction in patients with peritoneal metastasis from gastric adenocarcinoma based on the SEER database. J Gastrointest Oncol. 2024;15:597-611.
6. Cox DR. Regression models and life-tables. J R Stat Soc B. 1972;34:187-202.
7. Kleinbaum DG, Klein M. Survival analysis: a self-learning text. Berlin: Springer Nature; 2012.
8. Akinyemi JO, Afolabi RF, Awolude OA. Semi-parametric model for timing of first childbirth after HIV diagnosis among women of childbearing age in Ibadan, Nigeria. PLoS One. 2020;15:e0240247.
9. Chen J, Cheng Z, Yao Y, Wang S. Variation of all-cause mortality with fat-free mass index (FFMI) and fat mass index (FMI) in individuals with asthma: results from the NHANES retrospective cohort study. J Epidemiol Glob Health. 2024;14:1555-68.
10. Park SH, David J. Reassessing Schoenfeld residual tests of proportional hazards in political science event history analyses. Am J Polit Sci. 2015;59:1072-87.
11. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994;81:515-26.
12. Beyersmann J, Schumacher M. Time-dependent covariates in the proportional hazards model: background, structure, and interpretation. Lifetime Data Anal. 2007;13:401-18.
13. Fisher LD, Lin DY. Time-dependent covariates in the Cox proportional-hazards regression model. Annu Rev Public Health. 1999;20:145-57.
14. Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. Berlin: Springer Nature; 2000.
15. Suissa S. Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf. 2007;16:241-7.

16. Andersen PK, Gill RD. Cox's regression model for counting processes: a large-sample study. Ann Stat. 1982;10:1100-20.
17. Arslan J, Benke KK. Application of machine learning to ranking predictors of anti-VEGF response. Life. 2022;12:1176060.
18. Jin M. Imputation methods for informative censoring in survival analysis with time-dependent covariates. Contemp Clin Trials. 2024;136:107401.
19. Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. J Clin Epidemiol. 2016;76:175-82.
20. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis: II. Accuracy and precision of regression estimates. J Clin Epidemiol. 1995;48:1503-10.
21. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165:710-8.
22. Hsieh FY, Lavori PW. Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. Control Clin Trials. 2000;21:552-60.
23. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE Jr, Moons KG, et al. Minimum sample size for developing a multivariable prediction model: part II – binary and time-to-event outcomes. Stat Med. 2019;38:1276-96.
24. Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. Biometrics. 1983;39:499-503.
25. Postuma RB, Iranzo A, Hu M, Högl B, Boeve BF, Manni R, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain. 2019;142:744-59.
26. Spreafico M, Hazewinkel AD, van de Sande MA, Gelderblom H, Fiocco M. Machine learning versus Cox models for predicting overall survival in osteosarcoma: analysis of the EURAMOS-1 trial. Cancers. 2024;16:2880.
27. Kantidakis G, Hazewinkel AD, Fiocco M. Neural networks for survival prediction in medicine: a review and critical appraisal. Comput Math Methods Med. 2022;2022:1176060.
28 Wiegrebe S, Kopper P, Sonabend R, Bischl B, Bender A. Deep learning for survival analysis: a review. Artif Intell Rev. 2024;57:65.
29. Hazewinkel AD, Gelderblom H, Fiocco M. Prediction models with survival data: a comparison between machine learning and the Cox model. medRxiv. 2022:2022.03.29.22273112.
30. Moncada-Torres A, van Maaren MC, Hendriks MP, Siesling S, Geleijnse G. Explainable machine learning can outperform Cox regression and provide insights in breast cancer survival. Sci Rep. 2021;11:6968.
31. Wu Z, Cheng C, Sun X, Wang J, Guo D, Chen S, et al. The synergistic effect of the triglyceride-glucose index and serum uric acid on major adverse cardiovascular events after CABG: a multicentre cohort study. Cardiovasc Diabetol. 2023;22:103.

Table 1 – Multivariate Cox regression analysis for patient survival after propensity score matching. Reproduced from Shen Y *et al.*[5]

Characteristics	HR (95% CI)	*p*-value
Age (years)		
20 – 39	Reference	
40 – 59	0.8252 (0.3588, 1.8977)	0.65
60 – 79	0.7020 (0.3084, 1.5979)	0.40
≥ 80	1.1708 (0.4685, 2.9261)	0.74
Sex		
Male	Reference	
Female	0.8709 (0.6114, 1.2404)	0.44
Race		
White	Reference	
Black	0.9495 (0.5329, 1.6916)	0.86
Others	1.2767 (0.7712, 2.1137)	0.34
Grade		
I – II	Reference	
III – IV	0.9957 (0.6538, 1.5165)	0.98
T stage		
T1 – T2	Reference	
T3 – T4	0.9389 (0.5986, 1.4728)	0.78
N stage		
N0	Reference	
N1	1.4114 (0.9319, 2.1375)	0.10
N2	1.3273 (0.7547, 2.3341)	0.33
N3	1.0959 (0.5423, 2.2143)	0.80
Surgery		
No	Reference	
Yes	0.4382 (0.3037, 0.6324)	< 0.001***
Radiation		
No	Reference	
Yes	0.5463 (0.3355, 0.8896)	0.02*
Chemotherapy		
No	Reference	
Yes	0.3782 (0.2514, 0.5689)	< 0.001***
Tumor size (mm)		
2 – 43	Reference	
44 – 70	1.0131 (0.6915, 1.4843)	0.95
71 – 245	0.8655 (0.5417, 1.3829)	0.55

*: $p < 0.05$; ***: $p < 0.001$, multivariate Cox regression. PSM: propensity score matching; HR: hazard ratio; CI: confidence interval

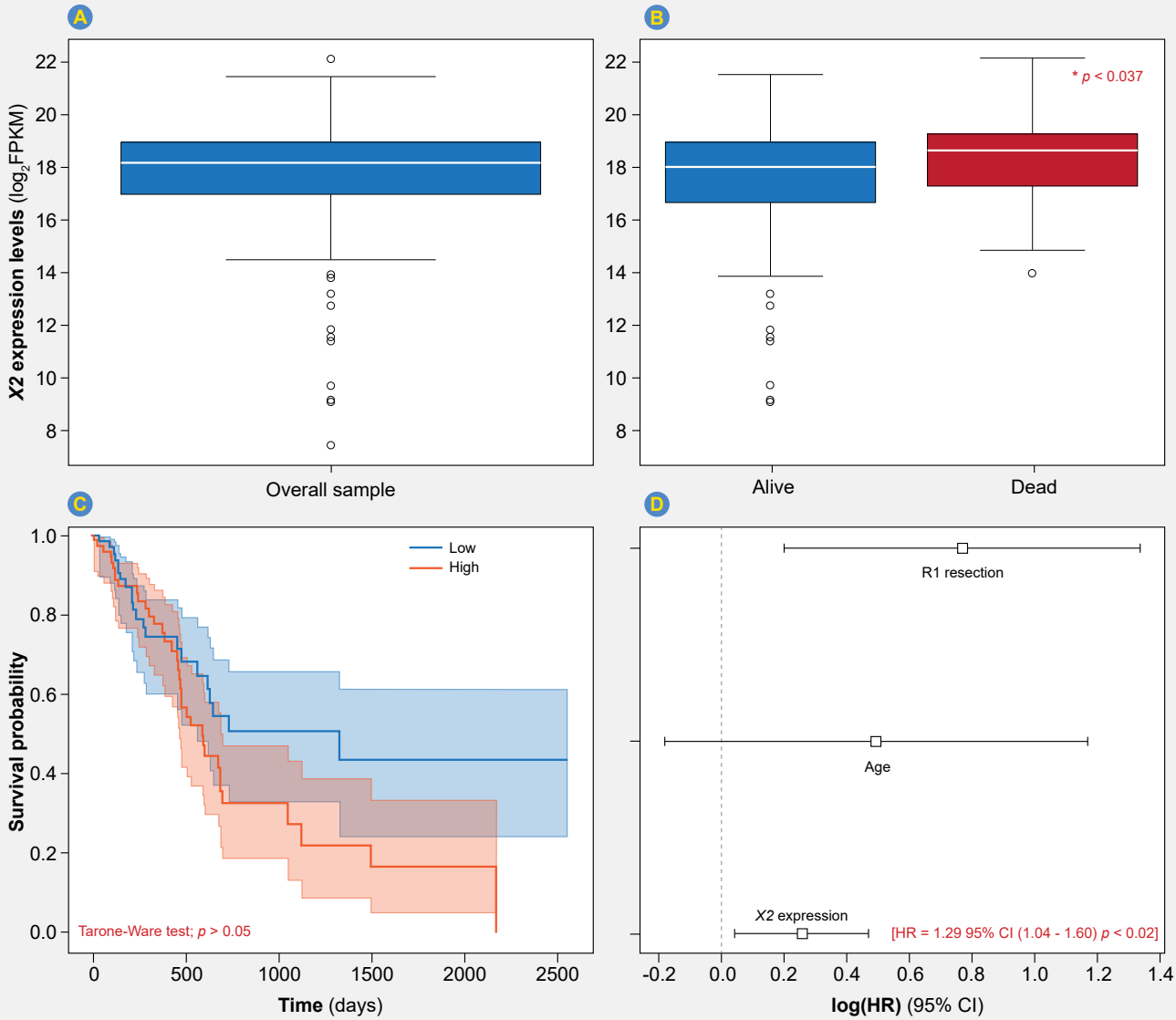

Figure 1 – Considering a hypothetical study exploring the role of *X2* gene expression in gastric cancer progression: A) descriptive statistics with a boxplot of *X2* gene expression of the whole sample. B) univariate statistics comparing the *X2* gene expression according to "no progression" (blue) group and disease progression (red) group, showing that *X2* expression is higher among patients with disease progression (*p*-value < 0.05). C) Kaplan-Meier curves comparing time to disease progression in patients with low *X2* expression (blue) and high *X2* expression (orange). It seems that the orange group has a lower time to event (higher risk of recurrence) compared to the blue group although *p*-value > 0.05 (no statistical significance). D) Results of a multivariate Cox regression model for 3 covariates (R1 resection; Age; *X2* expression). The left and right limits of each line represent the 95% confidence interval limits. *X2* expression confidence interval shows that higher *X2* expression is a risk factor for disease progression [HR = 1.29 95%CI (1.04 - 1.60), *p*-value < 0.02], adjusted for type of resection margin and age; R1 resection is an independent risk factor for disease progression. Author's own work.

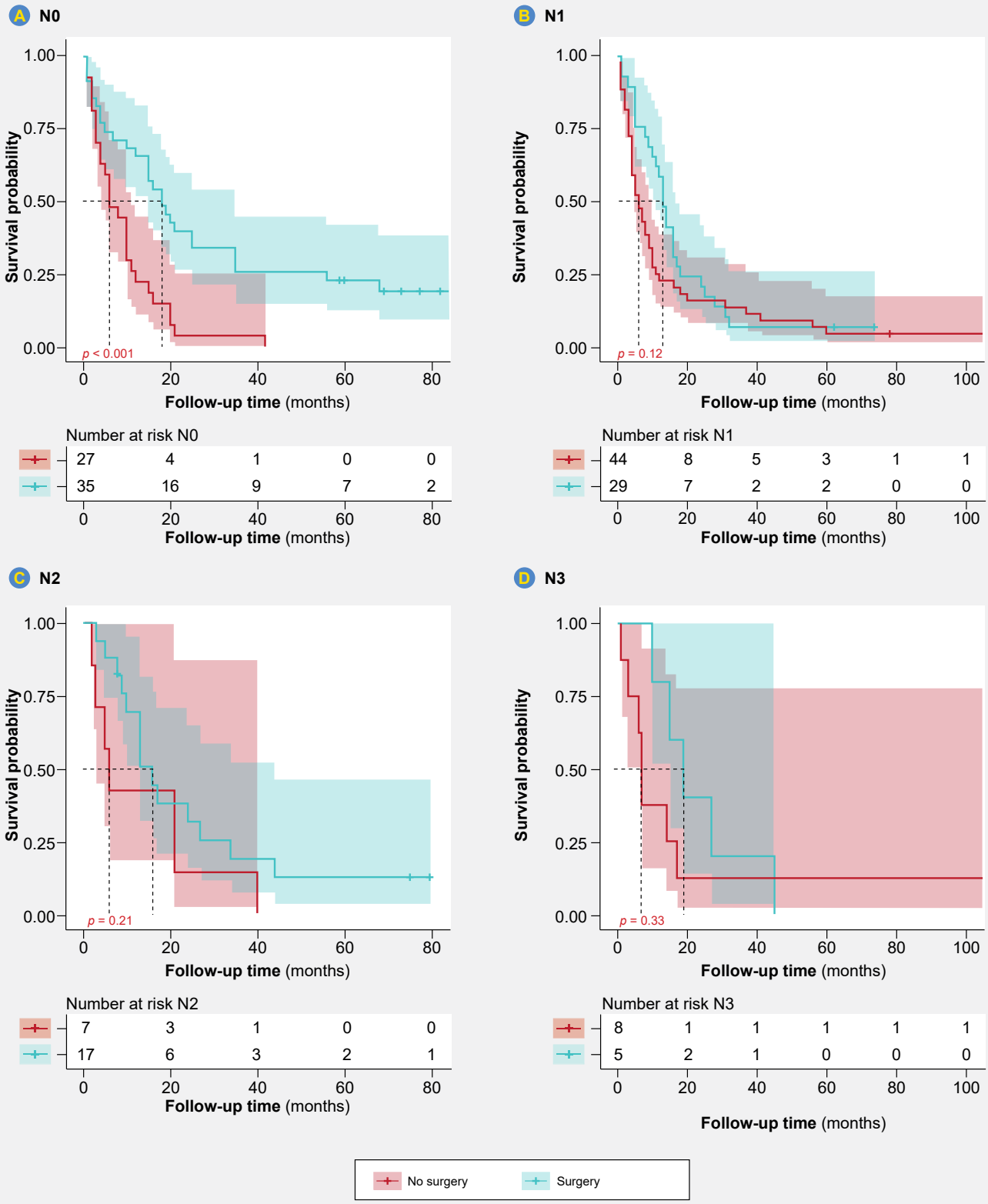

Figure 2 – Kaplan-Meier survival curve of GA with PMs in different grades and T stages after PSM. (A) grades I – II; (B) grades III – IV; (C) T1 – T2 stages; (D) T3 – T4 stages.

GA, gastric adenocarcinoma; PM, peritoneal metastasis; PSM, propensity score matching.

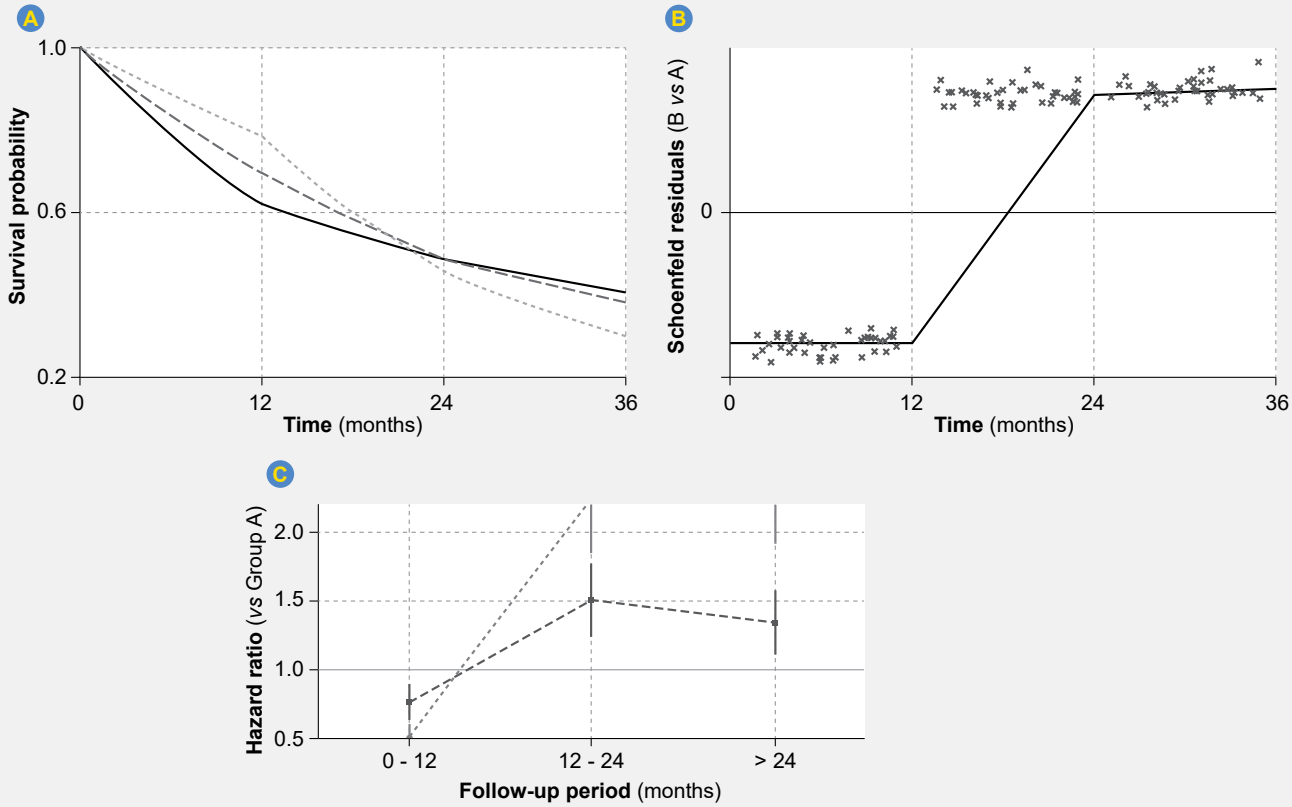

Figure 3 – Illustration of time-dependent effects. (A) Kaplan-Meier curves for three groups showing clear crossings at approximately 12 and 24 months, suggesting violation of the proportional hazards assumption. (B) Schoenfeld residuals for Group B *versus* Group A, demonstrating a clear time trend. The Grambsch–Therneau test for proportionality was statistically significant (p-value < 0.05). (C) Piecewise Cox regression model with three follow-up periods (0 - 12, 12 - 24, > 24 months), showing how the hazard ratio evolves over time: Group B *versus* A changes from protective to harmful, while Group C *versus* A shows gradual improvement. Together, these plots illustrate how visual inspection (A), formal diagnostics (B), and appropriate modelling (C) can be combined to address time-dependent effects in survival analysis. Author's own work.