Proceedings of the 2026 International Conference on Biological Science and Public Health
DOI: 10.54254/2753-8818/2026.35456

Semiparametric Survival Models in Biostatistics: Cox, Transformation, Competing Risk, and Cure Models

Mingxuan Li

School of Life Science, Beijing Institute of Technology, Beijing, China
priusleeing@gmail.com

Abstract. Biomedical follow-up studies often record the time until an event occurs rather than only whether the event occurs. Because some event times are censored, special survival methods are needed to use incomplete observations correctly. This paper summarizes semiparametric survival models from a biostatistical and mathematical perspective. It begins with right-censored notation and the Cox proportional hazards model, then explains partial likelihood, large-sample justification, and the role of risk sets. The paper next discusses model extensions, including proportional odds, accelerated failure time, transformation models, and likelihood-based semiparametric estimation. A model demonstration using the Veterans' Administration lung cancer data illustrates Kaplan-Meier curves, Cox regression, proportional hazards diagnostics, and AFT comparison. Shorter sections introduce survival quantile regression, competing risks, and cure models to show how the target of inference changes when event structures become more complex. The central argument is that modern survival analysis builds on the Cox framework but must choose the model scale, assumptions, and event structure according to the scientific question.

Keywords: survival analysis, Cox regression, semiparametric models, competing risks, cure models

1. Introduction

Biomedical studies often follow patients over time and record when a clinically meaningful event occurs. The event may be death, recurrence, tumor progression, treatment failure, or another endpoint. These outcomes differ from ordinary continuous outcomes because the exact event time is not always observed. A patient may remain event-free when follow-up ends, or may leave the study before the event occurs. Such observations are censored, not useless. They show that the event did not occur before a known time. Survival analysis was developed to use this partial information rather than discarding it.

The Cox proportional hazards model is the usual starting point for survival regression [1, 2]. Its strength is that it separates the time pattern of baseline risk from the covariate effect. The baseline hazard remains unspecified, while covariates enter through an exponential relative-risk term. This makes the model semiparametric and gives interpretable hazard ratios without forcing the event time into a parametric distribution.

DOI: 10.54254/2753-8818/2026.35456

$$\lambda(t \mid Z) = \lambda_0(t) exp(\beta^T Z) \tag{1}$$

Note: $\lambda_0(t)$ is the baseline hazard and β is the regression coefficient vector. This formula is retained as the main starting point, but the paper does not reproduce the full derivation from Cox's 1972 article.

The Cox model is attractive because it balances interpretation and flexibility. Its hazard ratio is easy to report, and the unspecified baseline hazard avoids strong distributional assumptions. However, the proportional hazards assumption may fail when treatment effects change over follow-up or survival curves cross. In those situations, Cox regression remains a useful reference model, but other semiparametric models may describe the data more appropriately.

This paper follows a compact mathematical path. It begins with right-censored notation, then summarizes Cox regression, partial likelihood, and large-sample theory. It then moves to proportional odds, accelerated failure time, transformation models, and likelihood-based semiparametric estimation. The later sections discuss quantile regression, competing risks, and cure models. The goal is not to list every survival method, but to show how models extend the same core idea: finite-dimensional covariate effects are estimated while baseline functions or event structures remain only partly specified.

The focus is deliberately mathematical, but the discussion is kept at a level suitable for a biostatistics course paper. Instead of proving every result, the paper emphasizes what each model assumes, what parameter it estimates, and how the interpretation changes when the model scale changes. This approach makes it possible to connect classical Cox regression with later semiparametric models without turning the paper into a long technical monograph.

2. Survival notation

Suppose a follow-up study includes n subjects. For subject i, T_i is the true event time and C_i is the censoring time. Under right censoring, the observed data are summarized by the observed time and the event indicator.

$$Y_i = min(T_i, C_i), \Delta_i = I(T_i \leq C_i) \tag{2}$$

Note: Δ_i=1 means that the event is observed, while Δ_i=0 means that the subject is censored. The observed record is (Y_i, Δ_i, Z_i), where Z_i is the covariate vector.

The main censoring assumption is that, after accounting for covariates, censoring should not carry extra information about the subject's future event risk. If high-risk subjects are more likely to be censored for unmeasured reasons, standard survival estimates may be biased. Under independent censoring, censored subjects still contribute information before their censoring times.

This notation is deliberately simple, but it is enough to support most models in the paper. The same observed record can be used for a Kaplan-Meier curve, a Cox model, an AFT model, or a transformation model. What changes is not the raw data structure, but the way the model connects covariates with the event-time distribution. This is why the notation section is important even though it looks elementary. It makes later formulas easier to interpret because each model is built from the same three objects: observed time, event status, and covariates.

Two functions describe the event-time distribution. The survival function gives the probability of remaining event-free beyond time t.

$$S(t) = P(T > t) \tag{3}$$

DOI: 10.54254/2753-8818/2026.35456

The hazard function describes the instantaneous event rate among subjects who have not yet experienced the event.

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t+\Delta t \mid T \ge t)}{\Delta t} \tag{4}$$

Note: The cumulative hazard is $\Lambda(t)=\int_0^t \lambda(u)du$, and for continuous event times $S(t)=exp\{-\Lambda(t)\}$. These note equations are used for interpretation and are not counted as numbered model equations.

Risk sets connect these functions to observed censored data. At an observed event time, the risk set contains subjects who are still followed and have not yet failed. The Kaplan-Meier estimator uses the event counts and risk sets to estimate the survival curve without covariates.

The risk-set idea also explains why censored subjects are not simply removed from the study. A subject censored at day 100 can still be part of the risk set at day 40 or day 80. The subject only stops contributing after the censoring time. This feature is one reason survival analysis differs from ordinary regression. The timing of censoring determines how much information a subject provides, and the order of events determines which comparisons are available to the model.

$$\widehat{S}(t) = \prod_{t_k \le t} (1 - \frac{d_k}{r_k}) \tag{5}$$

Note: d_k is the number of events at time t_k, and r_k is the number at risk just before that time. This estimator is descriptive and does not adjust for covariates.

3. Cox model

Cox regression adds covariates to survival analysis by comparing the subject who fails at an event time with all subjects still at risk at that same time. Conditional on one event occurring in a risk set, the probability that subject i is the failure is proportional to the subject's relative risk.

$$P(i fails \mid R(Y_i)) = \frac{exp(\beta^T Z_i)}{\sum_{j \in R(Y_i)} exp(\beta^T Z_j)} \tag{6}$$

Note: The baseline hazard $\lambda_0(Y_i)$ cancels because all subjects in the same risk set share it at that event time.

Multiplying these conditional probabilities over observed events gives the Cox partial likelihood. This is the main reason Cox regression can estimate covariate effects without selecting a parametric baseline survival distribution.

$$L(\beta) = \prod_{i:\Delta_i=1} \frac{exp(\beta^T Z_i)}{\sum_{j \in R(Y_i)} exp(\beta^T Z_j)} \tag{7}$$

The partial likelihood is not the full likelihood of the observed data. It focuses on relative comparisons inside risk sets. Even so, it gives a practical and theoretically justified estimator of β under standard assumptions. The score function compares the covariate vector of the observed failure with a model-weighted average covariate vector in the risk set.

$$U(\beta) = \sum_{i:\Delta_i=1} \{Z_i - \overline{Z}(Y_i, \beta)\} \tag{8}$$

Note: $\bar{Z}(Y_i,\beta)=\Sigma_{j \in R(Y_i)} Z_j \ exp(\beta^T Z_j)/\Sigma_{j \in R(Y_i)} exp(\beta^T Z_j)$. This note gives the definition without adding another numbered equation.

DOI: 10.54254/2753-8818/2026.35456

The observed information matrix is useful because it links risk-set covariate variation to statistical precision. Event times with little covariate variation contribute little information. Event times with diverse risk sets provide more evidence about covariate effects.

Although the score and information matrix look technical, their statistical meaning is intuitive. A coefficient is estimated from how failures are distributed across covariates inside each risk set. If high-risk covariate values repeatedly appear among the subjects who fail, the estimated coefficient moves in the direction of higher hazard. If the failures look similar to the remaining risk set, the model has weaker evidence for a covariate effect. This interpretation is useful in reading Cox output because it connects the table of hazard ratios back to the underlying risk-set comparisons.

$$I(\beta) = \sum_{i:\Delta_i=1}[\frac{\sum_{j\in R(Y_i)} Z_j Z_j^T exp(\beta^T Z_j)}{\sum_{j\in R(Y_i)} exp(\beta^T Z_j)} - \overline{Z}(Y_i,\beta)\overline{Z}(Y_i,\beta)^T] \tag{9}$$

The hazard ratio interpretation follows directly from the model. For two covariate values, the baseline hazard cancels.

$$\frac{\lambda(t|Z_a)}{\lambda(t|Z_b)} = exp\{\beta^T(Z_a - Z_b)\} \tag{10}$$

Because this expression does not depend on t, the model assumes proportional hazards. This assumption is convenient and often interpretable, but it should be checked in applications. When proportionality fails, the Cox coefficient can become an average summary rather than a complete description of the covariate effect.

In applied work, this interpretation is both helpful and risky. It is helpful because the hazard ratio gives a compact summary of covariate association. It is risky because the same summary may hide time-varying patterns. For example, two survival curves may be far apart early in follow-up and close together later. A single hazard ratio cannot show this shape directly. For that reason, the Cox model is best treated as a baseline semiparametric model whose assumptions must be examined rather than as an automatic final answer.

4. Cox theory

The partial likelihood requires large-sample justification because it is not a full likelihood. Later theory showed that, under independent censoring and regularity conditions, the partial likelihood estimator is consistent and asymptotically normal [3]. This makes standard inference possible.

$$\sqrt{n}(\widehat{\beta} - \beta_0) \Rightarrow N(0, \Sigma) \tag{11}$$

Note: This result supports Wald tests, score tests, confidence intervals, and standard errors for hazard ratios.

A counting-process view gives a clearer mathematical foundation. Let $N_i(t)$ record whether subject i has had an observed event by time t, and let $Y_i(t)$ indicate whether the subject is at risk at time t. The martingale residual process compares the observed event process with its model-based expected part.

$$M_i(t) = N_i(t) - \int_0^t Y_i(u) exp(\beta_0^T Z_i) d\Lambda_0(u) \tag{12}$$

Note: This representation treats the difference between observed events and expected events as mean-zero noise under the fitted model. It is the basis for martingale central limit arguments.

DOI: 10.54254/2753-8818/2026.35456

After β is estimated, the baseline cumulative hazard is usually estimated by a Breslow-type estimator. This estimator uses the same risk-set logic as partial likelihood: each event time adds a jump equal to the observed events divided by total weighted risk.

$$\widehat{\Lambda}_0(t) = \sum_{t_k \le t} \frac{d_k}{\sum_{j \in R(t_k)} exp(\widehat{\beta}^T Z_j)} \tag{13}$$

Predicted survival curves can then be obtained by combining the estimated baseline cumulative hazard with the fitted covariate effect.

This final step is important in applications because regression coefficients alone do not show the expected survival probability over time. A clinician or researcher often wants both relative measures, such as hazard ratios, and absolute summaries, such as predicted survival curves. Estimating the baseline cumulative hazard allows the Cox model to provide both kinds of information while still avoiding a parametric baseline distribution.

$$\widehat{S}(t \mid Z) = exp\{-\widehat{\Lambda}_0(t) exp(\widehat{\beta}^T Z)\} \tag{14}$$

This section is intentionally shorter than a full treatment of Cox's original article. The purpose is to retain the ideas needed for later sections: risk sets, partial likelihood, asymptotic inference, and baseline hazard estimation. These ideas reappear in broader semiparametric models with different assumptions and targets of inference.

The large-sample results also explain why Cox regression became a platform for later methods. Once the partial likelihood estimator could be studied with counting processes and martingales, researchers had a general language for censored data. The same language can be adapted when the hazard is not proportional, when the baseline function is replaced by an unknown transformation, or when the event process has several competing causes. Therefore, the theory is not only a justification for one model. It is part of the mathematical infrastructure of modern survival analysis.

5. Model extensions

The Cox model is useful, but it uses a constant hazard ratio. When proportional hazards are not plausible, other models may describe survival patterns on more suitable scales. A proportional odds model uses cumulative failure odds instead of instantaneous hazard.

$$\frac{1-S(t|Z)}{S(t|Z)} = \frac{1-S_0(t)}{S_0(t)} exp(\beta^T Z) \tag{15}$$

This model may be useful when the hazard ratio is not stable over time, but an odds-based summary of failure by time t is more natural. The accelerated failure time model gives another interpretation by modeling the event time itself.

The proportional odds model is especially useful as a contrast to Cox regression. Cox regression asks how covariates multiply the current event rate. The proportional odds model asks how covariates multiply the accumulated odds of having failed by a given time. These are different summaries of the same survival distribution. If the scientific question is closer to cumulative failure probability than to instantaneous risk, the proportional odds interpretation may be easier to explain to clinicians.

$$logT_i = \alpha + \beta^T Z_i + \varepsilon_i \tag{16}$$

DOI: 10.54254/2753-8818/2026.35456

Note: In an AFT model, exp(β$_j$) is a time ratio, not a hazard ratio. A coefficient may indicate delayed or accelerated survival time.

Linear transformation models place several common survival models into one semiparametric family [4-6]. They keep a finite-dimensional covariate effect while leaving a monotone transformation of time unspecified.

The value of the transformation model is that it turns several separate models into special cases of one framework. Instead of deciding only between Cox and one parametric AFT model, the analyst can think in terms of an unknown transformation of time plus a covariate shift. This point is important for a mathematical paper because it shows a hierarchy of assumptions. Cox regression fixes the hazard-ratio structure. AFT models impose a time-scale regression structure. Transformation models loosen the baseline time transformation while keeping a finite-dimensional regression coefficient.

$$H(T_i) = -\beta^T Z_i + \varepsilon_i \tag{17}$$

Note: With an extreme-value error distribution, this model corresponds to proportional hazards; with a logistic error distribution, it corresponds to proportional odds.

For censored data, transformation models require estimating both β and an unknown function. One way to express the estimating problem is through event processes and at-risk processes. The observed event increment is compared with a model-based expected increment.

The estimating-equation form is included because it shows how the Cox score idea generalizes. The model still compares what is observed with what is expected under the fitted model, but the expected increment now depends on a transformation function and an error distribution. This is more complicated than ordinary Cox regression, yet it follows the same broad principle. The observed failures provide information about the finite-dimensional coefficient, while the unspecified function absorbs the baseline event-time pattern.

$$\sum_{i=1}^{n} \int_0^{\infty} Z_i [dN_i(t) - Y_i(t) dL\{\beta^T Z_i + H(t)\}] = 0 \tag{18}$$

A related likelihood view is nonparametric maximum likelihood estimation, where the model contains finite-dimensional regression parameters and infinite-dimensional nuisance functions.

$$(\widehat{\beta}, \hat{\eta}) = \arg\max_{\beta,\eta} L(\beta, \eta) \tag{19}$$

Note: η may represent a baseline cumulative hazard, transformation function, or other unspecified component. This framework explains why semiparametric models are flexible but computationally more difficult than ordinary regression.

These extensions show that the Cox model is a foundation rather than the whole field. The model scale determines the interpretation: hazard ratio for Cox, odds ratio for proportional odds, and time ratio for AFT. Choosing among these models should depend on the data pattern and the scientific question, not only on mathematical convenience.

In practice, model choice should combine diagnostics, interpretation, and the research objective. If the proportional hazards diagnostic is acceptable and hazard ratios are meaningful, Cox regression remains a strong default. If the data show nonproportionality, a time-varying Cox model or a transformation model may be more appropriate. If the investigator wants to speak directly about delayed or shortened survival time, an AFT or quantile regression model may be easier to interpret. This is why the paper keeps several models in the same discussion rather than ranking them from simple to advanced.

DOI: 10.54254/2753-8818/2026.35456

For this reason, the paper uses formulas selectively. Core equations are numbered because they define the main models or estimators. Simpler related expressions are placed in notes so that the mathematical ideas remain visible without making the article look like a formula catalogue.

6. Veteran data

The model demonstration uses the no-prior-therapy subgroup from the Veterans' Administration lung cancer trial. The subgroup contains 97 patients, with 91 deaths and 6 censored observations. Survival time is measured in days. The covariates include tumor cell type, Karnofsky performance score, age, and treatment group. The example is included to connect the theoretical discussion with a real right-censored dataset rather than to make a new clinical claim.

All analyses in this section were implemented in R with standard survival-analysis functions. The Kaplan-Meier curves were used for visualization, the Cox model was used for covariate-adjusted hazard ratios, the Schoenfeld test was used for proportional hazards diagnostics, and AFT models were used as time-scale comparisons. The R output is summarized through figures and tables rather than through code, because the purpose is methodological illustration rather than software instruction.

Figure 2 shows the Kaplan-Meier treatment comparison, where median survival was 105 days in the standard group and 52 days in the test group. Figure 1 shows the corresponding curves by tumor cell type. Median survival was 128 days for squamous tumors, 51 days for small cell tumors, 51.5 days for adeno tumors, and 143 days for large cell tumors. These figure-based summaries suggest visible survival differences before covariate adjustment.

The Kaplan-Meier results should be read descriptively. They show how survival differs before covariate adjustment, but they do not explain whether the difference is due to treatment, tumor type, performance status, or age. In this dataset, tumor cell type and performance status are clinically meaningful sources of variation, so the regression model is needed to compare groups after accounting for multiple predictors. This also illustrates the typical workflow of survival analysis: start with nonparametric curves, then move to regression and diagnostics.

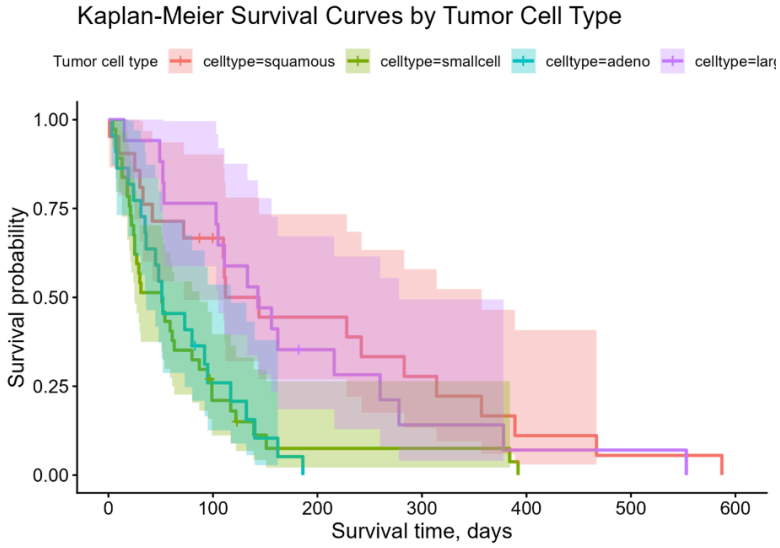

Figure 1. Kaplan Meier survival curves by tumor cell type in the veteran lung cancer data

DOI: 10.54254/2753-8818/2026.35456

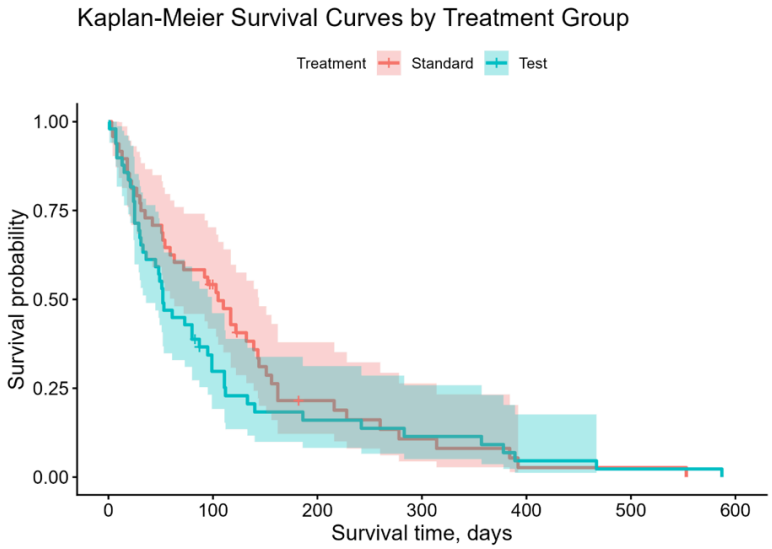

Figure 2. Kaplan Meier survival curves by treatment group in the veteran lung cancer data

Table 1 summarizes the fitted Cox model with tumor cell type, Karnofsky score, age, and treatment as covariates. Compared with squamous tumors, small cell tumors had a hazard ratio of 3.03, and adeno tumors had a hazard ratio of 3.13. Karnofsky score had a hazard ratio of 0.973 per one-point increase, indicating lower mortality risk with better performance status. Age was not statistically significant, while the test treatment group had a hazard ratio of 1.97 relative to the standard group.

Table 1. Cox proportional hazards model fitted to the veteran lung cancer data

Variable	Hazard Ratio	P_value	Lower_95_CI	Upper_95_CI
Cell-type small	3.03	0.000948	1.569714	5.838055
Cell-type adeno	3.13	0.001024	1.583997	6.182635
Cell-type large	1.55	0.221325	0.767827	3.130113
Karno	0.973	9.41E-06	0.961244	0.984836
Age	0.986	0.205869	0.963829	1.007971
Test treatment	1.97	0.010000	1.170000	3.310000

Table 2 reports the Schoenfeld residual diagnostic for the proportional hazards assumption. The global proportional hazards test was significant, suggesting that the fitted Cox model may not fully satisfy proportional hazards. Cell type and Karnofsky score also showed evidence of nonproportionality, while age and treatment did not. This result supports the main argument that Cox regression is a useful starting point, but its assumptions still need to be checked.

The diagnostic result is not a failure of the analysis. It is part of the analysis. A model that reports hazard ratios without checking proportional hazards may give a misleading impression of precision. Here, the significant global test strengthens the motivation for discussing transformation models and AFT models. The data example therefore connects the theoretical sections to a practical modeling decision: the Cox model is informative, but additional models may help describe patterns that a constant hazard ratio cannot capture.

DOI: 10.54254/2753-8818/2026.35456

Table 2. Schoenfeld residual test for the proportional hazards assumption

Variable	chisq	df	p
Cell-type	9.908184	3	0.019363
Karno	12.68224	1	0.000369
Age	0.765727	1	0.381542
Treatment	0.510206	1	0.47505
GLOBAL	22.68615	6	0.000909

A Cox model with a time-varying Karnofsky effect was also considered. The significant time-varying term suggests that the effect of performance status may vary over follow-up rather than remain constant. As a comparison, Weibull and log-normal AFT models were fitted. These AFT models gave a similar substantive pattern to the Cox analysis, but their coefficients are interpreted through time ratios rather than hazard ratios. The Weibull AFT model had an AIC of 1025.772, while the log-normal AFT model had a slightly lower AIC of 1025.676. Because the two AIC values were nearly identical, this comparison should be interpreted as illustrative rather than decisive.

The veteran data are not used for competing risks or cure modeling because the dataset does not clearly separate event types and does not show a stable long-term survival plateau. Its role is narrower but still important: it demonstrates how Cox regression, proportional hazards diagnostics, and AFT comparison can be applied to a standard right-censored biomedical dataset.

This separation between model demonstration and methodological extension keeps the analysis coherent. The same dataset should not be forced into every model simply because those models are discussed. A competing-risks model requires event-type information, and a cure model requires evidence of a nonsusceptible subgroup. Since the veteran dataset does not clearly provide either feature, limiting the applied analysis to standard right-censored models is more defensible.

7. Quantile regression

Survival quantile regression provides a time-scale alternative to hazard-based modeling [7]. Instead of summarizing covariate effects through a single hazard ratio, it models conditional quantiles of event time. This is useful when covariate effects differ across early, median, and later survival.

Quantile regression also helps interpret heterogeneity. A covariate may have a large association with early failures but a smaller association among subjects with longer survival. Hazard models can sometimes obscure this distributional pattern. By allowing $\beta(\tau)$ to change with the quantile level, survival quantile regression provides a more detailed view of how covariates relate to different parts of the event-time distribution.

$$Q_{logT}(\tau \mid Z) = \beta_0(\tau) + Z^T\beta(\tau) \tag{20}$$

This method is not the central focus of the paper, but it emphasizes an important point. Survival modeling does not have to be built only around hazards. If the scientific question concerns how a covariate shifts survival time at different parts of the distribution, quantile regression can be more direct than a hazard ratio.

DOI: 10.54254/2753-8818/2026.35456

8. Competing risks

Competing risks arise when several mutually exclusive event types are possible. Let T be the time to the first event and J be the event type, where J can take values from 1 to m. A cancer study may distinguish cancer-related death from death due to other causes. A subject who has one event type can no longer experience the other event types in the same way.

The cause-specific hazard for event type j describes the instantaneous rate of that event among subjects who are still free of all events.

$$\lambda_j(t \mid Z) = \lim_{\Delta t \to 0} \frac{P(t \le T < t+\Delta t,\, J=j \mid T \ge t,\, Z)}{\Delta t} \tag{21}$$

A Cox-type model can then be written for each event type.

$$\lambda_j(t \mid Z) = \lambda_{0j}(t) exp(\beta_j^T Z) \tag{22}$$

This model is useful for mechanism-focused questions. However, clinicians often want the probability of a particular event type by time t. That probability is the cumulative incidence function.

$$F_j(t \mid Z) = P(T \le t, J = j \mid Z) \tag{23}$$

Note: Equivalently, $F_j(t \mid Z) = \int_0^t S(u \mid Z)\lambda_j(u \mid Z)du$. The cumulative incidence depends on all event types through the overall survival term.

Fine and Gray proposed a model for the subdistribution hazard, which is more closely tied to cumulative incidence [8, 9].

$$\tilde{\lambda}_j(t \mid Z) = \tilde{\lambda}_{0j}(t) exp(\tilde{\beta}_j^T Z) \tag{24}$$

Cause-specific and Fine-Gray models answer different questions. Cause-specific hazards are often better for mechanisms, while Fine-Gray models are often better for predicting the cumulative probability of a specific event. The choice should be driven by the scientific target rather than by treating one method as generally superior.

This distinction is important in biomedical reporting. A treatment can reduce one cause-specific hazard while increasing a competing event, so the cumulative incidence may not move in the direction suggested by one hazard alone. Therefore, competing risks require the analyst to define the estimand before fitting the model. The question may be biological, focusing on the rate of a cause among event-free subjects, or clinical, focusing on the probability that a patient will experience a specific event within a time window.

9. Cure models

Standard survival models often assume that every subject will eventually experience the event if follow-up is long enough. Cure models relax this assumption by allowing the survival curve to approach a positive limit, which represents a nonsusceptible fraction [10, 11].

$$\lim_{t \to \infty} S(t) > 0 \tag{25}$$

DOI: 10.54254/2753-8818/2026.35456

A mixture cure model separates the probability of being susceptible from the survival distribution among susceptible subjects.

Note: The population survival function is $S_{pop}(t \mid X,Z)=1-p(X)+p(X)S_u(t \mid Z)$. In a Cox cure model, $S_u(t \mid Z)=\exp\{-\Lambda_{u0}(t)\exp(\gamma^T Z)\}$. These formulas are placed in the note to keep the main numbered formula count controlled.

This model separates incidence from latency. Incidence asks whether the event can occur at all, while latency asks when the event occurs among susceptible subjects. Cure models should be used only when the data and biology support a nonsusceptible subgroup. They generally need long follow-up and a clear survival plateau. The veteran lung cancer data are not suitable for cure modeling because they do not provide such evidence.

The cure-model distinction between incidence and latency is useful because two patients can differ in different ways. One covariate may affect the probability of ever being susceptible to relapse, while another may affect the timing of relapse among susceptible patients. A single standard Cox model cannot separate these two mechanisms. However, this separation also makes cure models more demanding, because censored subjects may be cured or may simply not have failed yet.

10. Discussion

The models in this paper are connected by a common problem: event times are often censored, and the baseline event-time structure is rarely known exactly. Cox regression addresses this by estimating finite-dimensional covariate effects while leaving the baseline hazard unspecified. Its partial likelihood is important because it transforms censored survival regression into risk-set comparisons. This gives interpretable hazard ratios without requiring a full parametric distribution.

At the same time, the Cox model should not be treated as a universal solution. Its proportional hazards assumption is strong, and the veteran data example showed why diagnostics matter. The significant Schoenfeld tests suggest that some covariate effects may change over time. In that setting, AFT models, transformation models, and quantile regression provide alternative interpretations on different scales.

The more advanced extensions also change the target of inference. Competing risks require the analyst to decide between mechanism-oriented cause-specific hazards and cumulative incidence. Cure models require a different long-term population structure, where some subjects may never experience the event. These choices show that survival analysis is not only a collection of formulas. It is a process of matching the statistical model to the biomedical question, the censoring structure, and the event process.

This perspective also helps control the scope of the article. Cox regression is covered in enough detail to explain the foundation, but the paper does not reproduce the full original set of formulas. The more advanced models are kept in view because they show where the field moves when proportional hazards, single-event endpoints, or eventual failure assumptions are not enough. The emphasis is therefore on a connected modeling logic rather than on exhaustive derivation.

11. Conclusion

The Cox proportional hazards model remains the main foundation for semiparametric survival analysis because it combines an unspecified baseline hazard with interpretable covariate effects. Its partial likelihood approach explains why regression parameters can be estimated from censored data without fixing a full survival distribution. However, the Cox model is only one part of modern survival analysis. Proportional odds models, AFT models, linear transformation models, and

DOI: 10.54254/2753-8818/2026.35456

survival quantile regression show that covariate effects can be studied on several scales, not only through hazard ratios. The veteran lung cancer example also shows the practical importance of model checking, since proportional hazards diagnostics suggested possible nonproportionality. Competing risks and cure models extend the framework further by changing the event structure or the long-term population assumption. Overall, the central lesson is that survival modeling requires both mathematical structure and scientific judgment. A good model should match the censoring pattern, the event type, the scale of interpretation, and the question being asked. In this sense, Cox regression is best understood as a starting point for a wider semiparametric system rather than as a final answer for all survival data.

References

[1] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B, 34*(2), 187–220.
[2] Kalbfleisch, J. D., & Schaubel, D. E. (2023). Fifty years of the Cox model. *Annual Review of Statistics and Its Application, 10*, 1–23.
[3] Tsiatis, A. A. (1981). A large sample study of Cox's regression model. *The Annals of Statistics, 9*(1), 93–108.
[4] Chen, K., Jin, Z., & Ying, Z. (2002). Semiparametric analysis of transformation models with censored data. *Biometrika, 89*(3), 659–668.
[5] Zeng, D., & Lin, D. Y. (2007). Maximum likelihood estimation in semiparametric regression models with censored data. *Journal of the Royal Statistical Society: Series B, 69*(4), 507–564.
[6] Guo, S., & Zeng, D. (2014). An overview of semiparametric models in survival analysis. *Journal of Statistical Planning and Inference, 151*, 1–16.
[7] Peng, L. (2021). Quantile regression for survival data. *Annual Review of Statistics and Its Application, 8*, 413–437.
[8] Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association, 94*(446), 496–509.
[9] Geskus, R. B. (2024). Competing risks: Concepts, methods, and software. *Annual Review of Statistics and Its Application*.
[10] Sy, J. P., & Taylor, J. M. G. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics, 56*(1), 227–236.
[11] Amico, M., & Van Keilegom, I. (2018). Cure models in survival analysis. *Annual Review of Statistics and Its Application, 5*, 311–342.