Evaluation and Comparison of Recidivism Modeling as a Survival Model

Andree Sulistio Chandra[1], Citra Nauli Al Fariz[1], Putri Anugrah Ramadhani[1], Rizky Uyuun Akfindarwan[1] & Utriweni Mukhaiyar[2,*]

[1]Master Program in Actuarial Science, Faculty of Mathematics and Natural Science, Bandung Technology Institute, Indonesia
[2]Statistic Research Division, Faculty of Mathematics and Natural Science, Bandung Technology Institute, Indonesia
*E-mail: utriweni.mukhaiyar@itb.ac.id

Abstract. Ex-offenders may reoffend after completing their sentences (recidivism). In this study, survival analysis was used to identify factors that affect time to reoffend. The Accelerated Failure Time (AFT) model was compared with the Cox Proportional Hazard (Cox PH) model. The effects of certain covariates on the acceleration or deceleration of the time to recidivism were then explained by the AFT model. Conversely, the Cox PH model provided estimates of the relative risk (hazard ratios) between groups. The model was evaluated using Concordance Index (CI), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). CI was used to evaluate the model's discrimination ability. The best-fitting model based on AIC and BIC was found to be the stratified Cox PH model after removal of several covariates. This was then used as the preferred model to estimate the relative failure risk between groups. Conversely, the AFT models gave higher CI values, indicating better predictive discrimination. The results suggest that the two models are of a different nature and should be selected depending on the analytical purpose and policy focus.

Keywords: *Accelerated Failure Time; Akaike Information Criterion (AIC); Bayesian Information Criterion (BIC); Concordance Index (CI); Cox Proportional Hazard; Recidivism; Stratified Cox PH; Survival Model.*

1 Introduction

Recidivism in criminal law refers to the tendency of former inmates to commit criminal acts again after serving a sentence for prior offenses [1]. Recidivism research using survival analysis has become an important part of multidisciplinary studies such as the analysis of risk factors that may lead former inmates to repeat criminal acts or studies on how the criminal justice system can act punitively or preventively toward the given numerical facts. Previous research has utilized survival models to understand recidivism dynamics in different

Received April 20th, 2026, Revised August 26th, 2026, Accepted for publication September 7th, 2026
Copyright © 2026 Published by ITB Institute for Research and Community Service, ISSN: 2337-5760, DOI: 10.5614/j.math.fund.sci.2026.58.1.5

community groups, including risk studies and studies on the influence of various factors on the probability of individuals reoffending, such as the influence of personal identity attribute factors [2,3,4,5] (for example, but not limited to, age, race, education level, financial capability), individual physical and mental health factors [6,7,8], external factors such as environment and community [7,9,10,11], as well as the form and level of punishment given to inmates [3,4,7,12,13,14], the influence of job quality obtained by parolees when released from custody [15], and criminal versatility factors (an index of various types of violations committed by an inmate during their lifetime) [16].

While previous studies commonly employed models such as the Kaplan-Meier estimator [2,6,8,9], Cox Proportional Hazards [2,3,6,7,9,10,12], and Accelerated Failure Time (AFT) [30], a critical methodological gap remains in how model efficacy can be evaluated. Traditionally, recidivism survival modeling prioritizes in-sample goodness-of-fit using penalized criteria like AIC and BIC. However, models achieving optimal theoretical fit do not necessarily guarantee high predictive discrimination on unseen data. This over-reliance on in-sample metrics creates a blind spot regarding the models' actual predictive stability and generalizability.

2 Methods

2.1 Survival Analysis

The statistical method of survival analysis looks at the time until an event happens, also known as time-to-event [18]. In this study, 'event' refers to the moment when former prisoners commit another crime that they are later convicted for. In mathematical terms, the time-to-event for an individual can be modeled as a random variable with an associated probability density function $f(t)$. Within survival models, the survival function $S(t)$ and the hazard function $h(t)$ are two prevalent functions that indicate the cumulative likelihood of an individual surviving beyond time t and the failure rate for individuals who have lived up to time t, respectively. These functions are mathematically expressed as follows:

$$S(t) = P(T \geq t) = \int_t^{\infty} f(s)ds \tag{1}$$

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr[t \leq T \leq t+\Delta t | T \geq t]}{\Delta t} = \frac{f(t)}{S(t)} \tag{2}$$

2.2 Accelerated Failure Time (AFT) Survival Model

The Accelerated Failure Time (AFT) regression model is a parametric survival model that assumes that the time-to-event variable is linearly correlated with its

covariates and follows a specific parameterized distribution. Conceptually, the model can be interpreted as 'accelerating' or 'decelerating' the occurrence of the event. Several common parametric distributions used to fit time-to-event data include the Exponential, Weibull, Logistic, Normal (Gaussian), Log-logistic, and Log-normal distributions.

Mathematically, the AFT model is defined as

$$S(t|\mathbf{X}) = S_0(t|\eta(\mathrm{X})) \tag{3}$$

$$\eta(\mathbf{X}) = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n) \tag{4}$$

where $S_0(t|\eta(X))$ represents the baseline survival function parameterized by the acceleration factor $\eta(X)$, which depends on the set of covariates $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$. The vector $\boldsymbol{\beta} = \{\beta_1, \beta_2, \ldots, \beta_n\}$ denotes the regression coefficients, with β_0 as the intercept term.

The regression coefficients and the model's scale parameters are estimated using the Partial Maximum Likelihood Estimation (MLE) method. Subsequently, the covariates of the optimal model are tested using the Wald test to determine the statistical significance of each covariate's influence on the model.

2.3 Cox Proportional Hazard (Cox PH) Survival Model

The Cox Proportional Hazard (Cox PH) regression model is a semi-parametric survival model, which, unlike the AFT model, does not assume that the data follows a specific parametric distribution [19]. The Cox PH model focuses on modeling the hazard function as a function of covariates and is mathematically defined as

$$h(t|\boldsymbol{X}) = h_0(t)\exp(\boldsymbol{\beta}^T \boldsymbol{X}) = h_0(t)\exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n) \tag{5}$$

where $h(t|\boldsymbol{X})$ denotes the hazard function given covariates $\boldsymbol{X}$ and $h_0(t)$ represents the baseline hazard function. The Cox PH model uses the Proportional Hazard (PH) assumption, which states that the hazard ratio between any two individuals or groups is proportional and constant over time. The validity of this assumption can be tested using the Schoenfeld residual test. The Schoenfeld residual represents the difference between the covariate value of the *i*-th observation and the weighted mean of the covariate values across all individuals at risk at time t_i. Mathematically, it is expressed as

$$\boldsymbol{r}_i = \rho_i(\boldsymbol{X}_i - \overline{\boldsymbol{X}}_{\widehat{\beta}}(t_i)) \tag{6}$$

$$\overline{\boldsymbol{X}}_{\widehat{\beta}}(t_i) = \frac{\sum_{q \in R(t_i)} X_q \exp(\sum_{j=1}^{n} \widehat{\beta}_j X_{jq})}{\sum_{q \in R(t_i)} \exp(\sum_{j=1}^{n} \widehat{\beta}_j X_{jq})} \tag{7}$$

where $\boldsymbol{X}_i = \{X_{1i}, \ldots, X_{ni}\}$ is the vector of covariate values for the i-th individual who experiences the event at time t_i.

ρ_i : an event indicator (1 if the event occurs, 0 otherwise),

n : the number of covariates in the model

$R(t_i)$: represents the set of individuals still at risk at the time t_i.

Heuristically, testing Schoenfeld residuals involves performing a generalized least squares (GLS) regression [20] of each residual $\boldsymbol{r}_i$ against time t_i to determine whether the residuals are independent of time, thus satisfying the PH assumption. Mathematically, for N observations and each covariate $j \in \{1, \ldots, n\}$, the correlation statistic is computed as follows:

$$T^{(j)} = \frac{\operatorname{cov}(r_i^{(j)}, t_i)}{\operatorname{sd}(r_i^{(j)}) \operatorname{sd}(t_i)} = \frac{\sum_{i=1}^{N} \left(r_i^{(j)} - \overline{r^{(j)}}\right)(t_i - \bar{t})}{\sqrt{\sum_{i=1}^{N} \left(r_i^{(j)} - \overline{r^{(j)}}\right)^2} \sqrt{\sum_{i=1}^{N} (t_i - \bar{t})^2}} \tag{8}$$

The null hypothesis $T^{(j)} = 0$ can be evaluated using a two-tailed Z-test for each covariate or collectively using the Global Schoenfeld Residual Test, where $T_{\text{global}} = \sum_{j=1}^{n} (T^{(j)})^2$ is tested with a Chi-squared distribution with n degrees of freedom.

Mitigation strategies for PH assumption violations include elimination, removing the violating covariate, or stratification, treating it as a stratification variable instead of a predictor [21]. The stratified Cox PH model is expressed as

$$h_{\mathbf{Z}}(t|\boldsymbol{X}) = h_{0_{z_1, z_2, \ldots, z_m}}(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_{n-m} X_{n-m}) \tag{9}$$

The relative hazard risk (or hazard ratio) is the most common interpretation of the Cox PH model, representing the ratio of hazard rates between two groups for a given covariate [18,22]. For a covariate X_i that satisfies the PH assumption, the relative hazard risk is given by $\exp(\beta_i)$.

2.4 Model Fit and Accuracy Evaluation

Evaluation of model fit, or goodness-of-fit, is conducted to determine the best model for the empirical data. The criteria used for assessing goodness-of-fit are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion

(BIC), introduced by Akaike [31] and Schwarz [32], respectively. The criteria are formally expressed as

$$AIC = -2(\mathcal{L}) + 2(P + n) \tag{10}$$

$$BIC = -2(\mathcal{L}) + (P + n) \times \log(N) \tag{11}$$

where $P + n$ denotes the total number of independently estimated parameters in the model. For the AFT model, the log-likelihood function is defined as

$$\mathcal{L}_{AFT} = \sum_{i=1}^{N} \rho_i \log[f(t_i; X_i)] + \sum_{i=1}^{N} (1 - \rho_i) \log[S(t_i; X_i)] \tag{12}$$

For the Cox Proportional Hazard (Cox PH) model, the partial log-likelihood function is expressed as

$$\mathcal{L}_{Cox-PH} = \sum_{i \in D} \left[\beta^T X_i - \log(\sum_{q \in R(t_i)} e^{\beta^T X_q}) \right] \tag{13}$$

where D denotes the set of indices of events, β is the vector of model parameters, X_i is the covariate vector for individual i, and $R(t_i)$ denotes the risk set at time t_i. A theoretical limitation should be noted when contrasting these metrics, as direct comparisons of AIC and BIC between AFT, which uses full likelihood, and Cox PH, which uses partial likelihood. The models are strictly indicative, since their objective functions are defined on fundamentally different mathematical scales.

In addition to penalized likelihood criteria, predictive performance is evaluated using the Concordance Index (CI) [23, 24]. CI measures the model's discriminatory ability by evaluating the proportion of concordant pairs among all evaluable pairs of subjects:

$$CI = P[\hat{y}_i > \hat{y}_j | y_i > y_j] = \frac{\sum_{i,j} I(y_i > y_j)\, I(\hat{y}_i > \hat{y}_j)}{\sum_{i,j} I(y_i > y_j)} \tag{14}$$

where $I(\circ)$ is an indicator, y_i, y_j and $\hat{y}_i, \hat{y}_j$ represent the observed and model-predicted survival times for individuals i and j. Under right-censoring, a pair is evaluable only if the observation with the shorter duration experienced the event, with $\rho = 0$, while non-informative pairs are omitted, such as those where both individuals are censored. A CI of 0.5 denotes random guessing, while 1.0 indicates perfect ranking. A time-dependent CI over intervals $[t_a, t_b]$ is also implemented, via resampling, to assess temporal stability across the follow-up period [23, 25].

3 Results and Discussions

3.1 Research Data and Variables

A total of 28,413 data on released individuals from fiscal years 2016 to 2024 were obtained from the Iowa Department of Corrections via the Iowa Open Data Program [17]. To handle right-censoring, non-recidivists were censored at the study's administrative observation cutoff date, September 13, 2024. Seven covariates were selected, as presented in Table 1, based on established criminological literature. Demographic traits, namely gender, race, and age, served as core baseline predictors [2, 3], while supervision type, duration, ending reason, and offense severity captured institutional and correctional factors shown to influence reoffense timing [12–14, 16]. Distributions of these covariates across the cohort are detailed in Table 2.

Table 1 Definition of variables as response variables, indicators, and covariates.

Variable	Name	Scale, Unit	Factor	Description
Y	Survival Time	Ratio, Days	N/A	The period from a recidivist's release into the community to their subsequent recidivism.
ρ	Recidivism Status	Nominal	1: Non-recidivist 0: Recidivist	Status of recidivism.
X_1	Gender	Nominal	1: Male, 0: Female	Gender of the recidivist.
X_2	Race	Nominal	1: White, 0: Other	Race/ethnicity of the recidivist.
X_3	Age	Ratio, Years	N/A	Age of the recidivist at the end of the supervision period.
X_4	Type of Supervision	Nominal	1: Prison 0: Work Release	The type of supervision assigned to recidivist. 'Prison' refers to custodial sentencing, while 'Work Release' refers to non-custodial supervision.
X_5	Supervision Duration	Ratio, Days	N/A	Duration of supervision assigned to the recidivist.
X_6	Reason for Ending Supervision	Nominal	1: Paroled 0: Expiration	Reason for the end of the supervision period. 'Paroled' refers to early release under conditions, while 'Expiration' refers to completing the sentence without early release.
X_7	Offense Level	Nominal	1: Felony 0: Misdemeanor	Classification of recidivist offense. 'Felony' refers to a serious offense, while 'Misdemeanor' refers to a minor offense.

The average age of the cohort was 35.98 years, with 34.13% between 21 and 30 years old and 35.03% between 31 and 40. The average supervision duration was

400.36 days, with 68.14% supervised for 0 to 365 days and 20.94% for 366 to 730 days. Among recidivists, the average age was 34.77 years, with 37.14% aged 21 to 30 and 36.29% aged 31 to 40. Their average supervision duration was 344.3 days, with 69.67% falling within 0 to 365 days and 21.51% within 366 to 730 days. Figure 1 presents these distributions by cohort, recidivist status, age, and detention length, illustrating the shape and skewness of each. Prior to the modeling stage, a Pearson correlation test [26, 27] on all random variables $\{X_1, \ldots, X_7\}$ showed every pairwise correlation below 0.3 in Table 3, indicating the variables were independent or weakly correlated.

Table 2 Frequency and proportion of recidivism by covariate.

Factor		Cohort		Non-Recidivist		Recidivist	
		Freq	**Prop**	**Freq**	**Prop**	**Freq**	**Prop**
(X_1) Gender	Male	24278	85.45%	14927	52.54%	9351	32.91%
	Female	4135	14.55%	2866	10.09%	1269	4.47%
(X_2) Race	White	19918	70.10%	12305	43.31%	7613	26.79%
	Other	8495	29.90%	5488	19.32%	3007	10.58%
(X_4) Type of Supervision	Prison	20643	72.65%	13392	47.13%	7251	25.52%
	Work Release	7770	27.35%	4401	15.49%	3369	11.86%
(X_6) Reason for Ending Supervision	Paroled	22286	78.44%	13248	46.63%	9038	31.81%
	Expiration	6127	21.56%	4545	15.99%	1582	5.57%
(X_7) Offense Level	Felony	23161	81.52%	14166	49.86%	8995	31.66%
	Misdemeanor	5252	18.48%	3627	12.76%	1625	5.72%

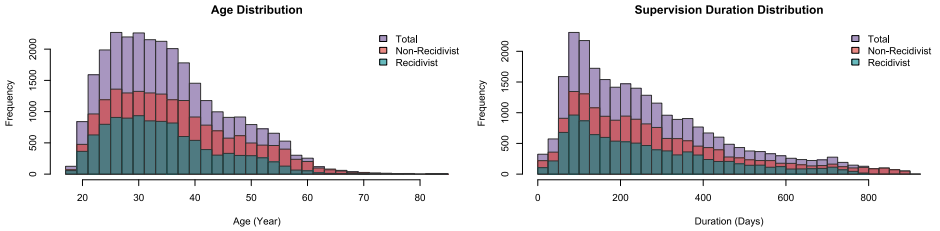


Figure 1 Histogram of recidivist distribution by age (left) and supervision duration (right).

Table 3 Correlation values between covariates.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1.000	-0.078	0.014	-0.025	0.089	-0.078	0.018
X_2	-0.078	1.000	0.101	-0.005	-0.019	0.037	-0.023
X_3	0.014	0.101	1.000	-0.056	0.102	0.059	0.039
X_4	-0.025	-0.005	-0.056	1.000	0.274	-0.221	-0.189
X_5	0.089	-0.019	0.102	0.274	1.000	-0.035	0.104
X_6	-0.078	0.037	0.059	-0.221	-0.035	1.000	0.290
X_7	0.018	-0.023	0.039	-0.189	0.104	0.290	1.000

3.2 Development of Parametric Accelerated Failure Time Model (AFT)

The regression coefficients and scale parameters (σ) for each AFT distribution are presented in Table 5. Based on the AIC and BIC values in Table 4, the Lognormal model provided the best fit, with all covariates statistically significant at $\alpha = 0.05$. Under this optimal model, the practical effects are interpreted through time ratios (e^{β}). Male offenders X_1 exhibited shorter survival times before recidivating than females, with a time ratio of 0.557. Each additional year of age X_3 is associated with a longer survival time, with a time ratio of 1.026. Paroled individuals X_6 reoffended substantially faster than those with expired sentences, with a time ratio of 0.373.

Table 4 AIC and BIC values for each parametric AFT distribution.

	Lognormal	**Loglogistic**	**Weibull**	**Exponential**	**Gaussian**	**Logistic**
AIC	184906.7	185974.0	187169.0	190366.0	205952.8	207455.7
BIC	184980.5	186047.9	187243.0	190432.0	206026.7	207529.5

Table 5 Estimated parameter values for each parametric AFT distribution (α = 5%).

	Lognormal		**Loglogistic**	**Weibull**	**Exponential**	**Gaussian**	**Logistic**
	value	$e^{-\beta_i}$	**value**	**value**	**value**	**value**	**value**
β_0	8.7131	0.00016	8.6983	9.3619	8.7276	3266.8182	3323.9154
β_1	-0.5856	1.79607	-0.6493	-0.6229	-0.4216	-572.9202	-674.0470
β_2	-0.2479	1.28133	-0.2760	-0.2613	-0.1805	-228.1815	-272.7590
β_3	0.0258	0.97453	0.0288	0.0271	0.0181	25.0900	29.3622
β_4	0.1445	0.86545	0.1520	0.1258	0.0807	102.8112	120.2353
β_5	0.0004	0.99960	0.0004	0.0004	0.0003	0.3669	0.4226
β_6	-0.9875	2.68451	-1.0268	-0.9591	-0.6505	-901.1055	-1048.2439
β_7	-0.1748	1.19101	-0.2506	-0.2742	-0.1937	-233.7186	-294.3498
σ	2.2280		1.3274	1.5970	1	2127.78	1324.802

3.3 Development of the Cox Proportional Hazard (Cox PH) Semi-Parametric Model

The Cox PH model was initially constructed by incorporating all random variables $\{X_1, \ldots, X_7\}$ as covariates, after which the p-values of the PH assumption test were calculated. The results from Table 6 indicate that 5 out of 7 included covariates did not satisfy the PH assumption; therefore, mitigation strategies were implemented to address the violation of the PH assumption.

Table 6 PH PH assumption test for the covariates in the Cox PH model (α = 5%).

Covariate	X_1	X_2	X_3	X_4	X_5	X_6	X_7
p	0.02	0.00	0.12	0.00	0.00	0.00	0.36

Given the substantial number of covariates violating the PH assumption, the construction of a new model was performed using an iterative method. Iterations of the Cox PH model with $n = \{2,3,4,5,6,7\}$ random variables included as covariates and $l = \{0,1,2\}$ random variables taken as strata were conducted. X_3 (age) and X_5 (supervision duration) were not included in the strata selection to avoid information loss. From the iterations performed, 665 combinations of covariates and strata were obtained, of which 29 Cox PH models satisfied the PH Assumption.

Table 7 Goodness-of-Fit tests for the Cox PH and AFT models.

Model	p	AIC	BIC	Model	p	AIC	BIC
$X_{\{1,3,7\}} + Z_{\{2,4\}}$	0.08	186662.8	186684.6	AFT Loglogistic	-	192611.7	192686.0
$X_{\{3,5,7\}} + Z_{\{2,4\}}$	0.32	186680.3	186702.1	AFT Weibull	-	193828.1	193902.4
$X_{\{3,7\}} + Z_{\{2,4\}}$	0.19	186758.1	186772.6	$X_{\{1,3\}} + Z_{\{6,7\}}$	0.41	195132.9	195147.4
$X_{\{3,5\}} + Z_{\{2,4\}}$	0.72	186795.4	186809.9	AFT Exponential	-	197221.7	197287.7
$X_{\{5,7\}} + Z_{\{2,4\}}$	0.20	186928.3	186942.9	$X_{\{3,5,7\}} + Z_{\{4\}}$	0.20	199378.6	199400.4
$X_{\{1,3,5\}} + Z_{\{4,6\}}$	0.13	190571.9	190593.7	$X_{\{3,7\}} + Z_{\{4\}}$	0.11	199461.1	199475.7
$X_{\{1,3\}} + Z_{\{4,6\}}$	0.29	190666.6	190681.1	$X_{\{3,5\}} + Z_{\{4\}}$	0.45	199489.8	199504.4
$X_{\{1,3\}} + Z_{\{2,6\}}$	0.56	190717.5	190732.0	$X_{\{5,7\}} + Z_{\{4\}}$	0.18	199605.5	199620.0
$X_{\{3,5\}} + Z_{\{4,6\}}$	0.20	190719.5	190734.0	$X_{\{1,3,7\}} + Z_{\{2\}}$	0.18	200065.6	200087.4
$X_{\{1,5\}} + Z_{\{4,6\}}$	0.08	190832.0	190846.6	$X_{\{3,7\}} + Z_{\{2\}}$	0.47	200165.6	200180.1
$X_{\{3,5\}} + Z_{\{4,7\}}$	0.53	190841.7	190856.3	$X_{\{1,3\}} + Z_{\{2\}}$	0.09	200197.4	200211.9
$X_{\{1,3\}} + Z_{\{2,7\}}$	0.10	190980.3	190994.8	$X_{\{1,7\}} + Z_{\{2\}}$	0.18	200334.4	200348.9
$X_{\{3,5,7\}} + Z_{\{1,4\}}$	0.22	191510.4	191532.2	$X_{\{1,3\}} + Z_{\{6\}}$	0.31	203410.5	203425.0
AFT Lognormal	-	191534.0	191608.3	$X_{\{3,7\}} + Z_{\{1\}}$	0.24	205003.8	205018.3
$X_{\{3,7\}} + Z_{\{1,4\}}$	0.12	191611.3	191625.9	$X_{\{3,7\}}$	0.23	212868.1	212882.7
$X_{\{3,5\}} + Z_{\{1,4\}}$	0.51	191620.9	191635.4	AFT Gaussian	-	213395.1	214949.1
$X_{\{5,7\}} + Z_{\{1,4\}}$	0.20	191736.9	191751.4	AFT Logistic	-	213469.4	215023.4
$X_{\{3,7\}} + Z_{\{2,1\}}$	0.48	192339.5	192354.0				

*$X_{\{i,j,k\}}$ *denotes variables used as covariates and* $Z_{\{a,b\}}$ *denotes variables used as strata in the Cox PH model.*

The AIC and BIC values for all variations of Cox PH models satisfying the PH assumption were calculated with details presented in Table 7. The AIC and BIC values from the six AFT models from Table 5 were also calculated and presented in the same table as a comparison of the goodness-of-fit across all model variations.

3.4 Model Accuracy Evaluation Based on the Concordance Index Metric

The Concordance Index was calculated using a resampling procedure performed 2048 times, where 70% of the cohort data were randomly selected without event distribution balancing. In this study, the time-dependent CI was evaluated for

annual survival intervals over the first three years, specifically within the intervals [0, 365], [366, 730], and [731, 1,095] days. The results are presented in Figure 2, with the numerical summaries of the CI for the overall survival period shown in Table 8. It can be observed that all parametric AFT models exhibit substantially higher predictive accuracy compared to all Cox PH models.

Table 8 CI values for the overall survival time.

Model	Mean	Interval	Model	Mean	Interval	Model	Mean	Interval
AFT loglogistic	0.59712	(0.59, 0.61)	$X_{\{3,7\}} + Z_{\{2,4\}}$	0.55488	(0.54, 0.57)	$X_{\{3,5\}} + Z_{\{2,4\}}$	0.54650	(0.53, 0.56)
AFT weibull	0.59706	(0.59, 0.61)	$X_{\{3,7\}} + Z_{\{2\}}$	0.55410	(0.54, 0.56)	$X_{\{1,3\}} + Z_{\{6,7\}}$	0.54642	(0.54, 0.56)
AFT exponential	0.59706	(0.59, 0.61)	$X_{\{3,7\}} + Z_{\{2,1\}}$	0.55282	(0.54, 0.56)	$X_{\{3,7\}} + Z_{\{1\}}$	0.54613	(0.53, 0.56)
AFT lognormal	0.59701	(0.59, 0.61)	$X_{\{1,3\}} + Z_{\{2\}}$	0.55215	(0.54, 0.56)	$X_{\{3,5\}} + Z_{\{4,7\}}$	0.54384	(0.53, 0.56)
AFT logistic	0.59695	(0.59, 0.61)	$X_{\{1,3\}} + Z_{\{6\}}$	0.55171	(0.54, 0.56)	$X_{\{3,5\}} + Z_{\{1,4\}}$	0.54256	(0.53, 0.56)
AFT gaussian	0.59692	(0.59, 0.61)	$X_{\{3,5,7\}} + Z_{\{4\}}$	0.55086	(0.54, 0.56)	$X_{\{3,5\}} + Z_{\{4\}}$	0.54169	(0.53, 0.55)
$X_{\{1,3,7\}} + Z_{\{2,4\}}$	0.56189	(0.55, 0.57)	$X_{\{1,3\}} + Z_{\{2,7\}}$	0.55019	(0.54, 0.56)	$X_{\{1,7\}} + Z_{\{2\}}$	0.53183	(0.52, 0.54)
$X_{\{1,3,7\}} + Z_{\{2\}}$	0.56151	(0.55, 0.57)	$X_{\{3,5,7\}} + Z_{\{1,4\}}$	0.55018	(0.54, 0.56)	$X_{\{1,5\}} + Z_{\{4,6\}}$	0.53054	(0.52, 0.54)
$X_{\{1,3\}} + Z_{\{2,6\}}$	0.55868	(0.55, 0.57)	$X_{\{3,7\}} + Z_{\{4\}}$	0.54866	(0.54, 0.56)	$X_{\{5,7\}} + Z_{\{2,4\}}$	0.52795	(0.52, 0.54)
$X_{\{1,3,5\}} + Z_{\{4,6\}}$	0.55786	(0.55, 0.57)	$X_{\{3,7\}}$	0.54816	(0.54, 0.56)	$X_{\{5,7\}} + Z_{\{1,4\}}$	0.52690	(0.52, 0.54)
$X_{\{3,5,7\}} + Z_{\{2,4\}}$	0.55574	(0.54, 0.57)	$X_{\{3,5\}} + Z_{\{4,6\}}$	0.54707	(0.53, 0.56)	$X_{\{5,7\}} + Z_{\{4\}}$	0.52690	(0.51, 0.54)
$X_{\{1,3\}} + Z_{\{4,6\}}$	0.55512	(0.54, 0.57)	$X_{\{3,7\}} + Z_{\{1,4\}}$	0.54696	(0.53, 0.56)			

The time-dependent CI evaluation of the AFT models indicates a notable decline in predictive accuracy over time, with the highest predictive performance observed during the survival interval [0, 365] and the lowest performance during [731, 1,095]. This suggests that the presence of temporal limitations affects prediction accuracy as the survival duration increases. In contrast, such a pattern was not observed in the time-dependent CI results for the Cox PH models. All Cox PH models displayed their highest CI values within the [731, 1,095] interval and their lowest CI values within the [0, 365] interval.

Furthermore, as shown in Figure 2, the CI values for the overall survival period tended to be lower than those observed for the interval [366, 1,095]. This indicates that the model's predictive performance in the aggregated evaluation is weaker compared to its performance when assessed within specific time intervals. However, even the best-performing Cox PH model, specifically the model $X_{\{1,3,7\}} + Z_{\{2\}}$ within the [731, 1,095] interval, achieved CI values in the range of (0.56, 0.58), which remained relatively low and approached the level of random chance prediction.

4 Conclusions

The evaluation results of model fit and accuracy based on AIC, BIC, and the Concordance Index yielded several models that can be used to represent

recidivism as a survival process for the population captured by the data cohort employed in this study. The Cox proportional hazards model with covariates $X_{\{1,3,7\}} + Z_{\{2,4\}}$ demonstrated the strongest model fit according to AIC and BIC among all Cox proportional hazards and AFT specifications tested, reflecting the greater flexibility of the Cox proportional hazards model in performing fitting without assuming a particular parametric distribution, which makes it the strongest model for estimating the relative failure risk or hazard ratio between two groups.

The evaluation of model prediction accuracy using the Concordance Index revealed that AFT models built on all six common parametric distributions produced substantially better prediction accuracy than every Cox proportional hazard model tested. The Concordance Index results for AFT models across these six parametric distributions, evaluated over the overall time horizon and within the survival time intervals of zero to 365 days, 366 to 730 days, and 731 to 1095 days, fell respectively within the ranges of 0.59 to 0.61, 0.61 to 0.64, 0.6 to 0.63, and 0.6 to 0.62. These results contrast sharply with the Cox proportional hazards models, which exhibited considerably lower Concordance Index ranges throughout.

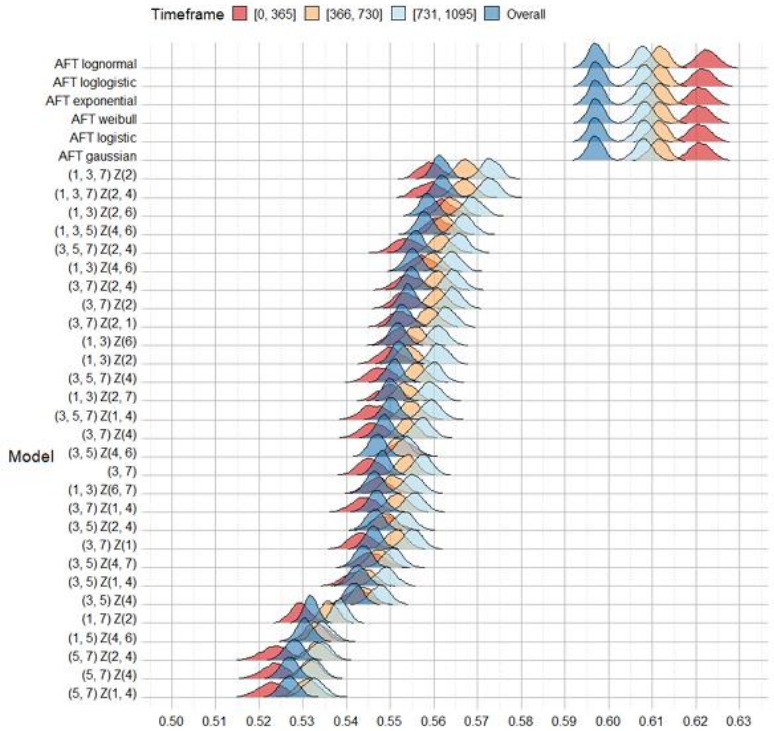


Figure 2 Prediction accuracy test results of the model.

The Cox proportional hazards models gave the best in-sample fit as measured by AIC and BIC but performed very poorly in terms of predictive accuracy, only marginally better than random guesses. The parametric assumptions underlying the AFT models likely contributed to their superior predictive performance. The results suggest that it is a dangerous practice to rely solely on goodness-of-fit measures such as AIC and BIC, which only measure the fit of the model to the current data, while the Concordance Index measures the more important ability to rank future risks and can lead to the selection of practically useless models. Therefore, out-of-sample evaluation is necessary and we suggest AFT frameworks rather than Cox proportional hazards models for robust recidivism prediction based on this study.

The absolute Concordance Index values obtained from the best model in this study were 0.61-0.64, which surely can be improved. The predictive power of the covariates available in the present dataset is relatively weak compared to similar studies that used up to twenty-five covariates [28] or focused on high-risk cohorts such as sexual offenders [29], which showed Concordance Index ranges of 0.69 to 0.74. This limitation, however, does not impact the methodological comparison presented here. Future work should evaluate this framework on a richer set of predictors to improve absolute accuracy, including dynamic risk factors such as employment stability following release or a history of substance abuse, while maintaining a focus on out-of-sample discrimination.

That the fit of the model and predictive accuracy do not align has significant practical implications for criminal justice policy and practitioner risk assessments. Models that fit historical data well but do not discriminate individual risks could result in misallocation of supervision resources or unfair parole decisions. By applying AFT models that are designed for true predictive discrimination, parole boards and correctional practitioners can make more reliable evidence-based assessments and focus intensive supervision on those who are at truly higher risk of early recidivism.

5 Acknowledgement

The authors would like to express their gratitude to Institut Teknologi Bandung (ITB) for providing the necessary facilities and support to conduct this research. This research was also supported by PPMI KK FMIPA ITB 2026.

6 References

[1] National Institute of Justice, *Recidivism*, National Institute of Justice, (Online), 2025, Available:

https://nij.ojp.gov/topics/corrections/recidivism, (Accessed 1 January 2025).

[2] Cuervo, K., Villanueva, L. & Basto-Pereira, M., *Prediction of Youth and Adult Recidivism Among Spanish Juveniles Involved in Serious Offenses*, Criminal Justice and Behavior, **47**(4), pp. 399-418, 2020. 10.1177/0093854819896395.

[3] Yu, R., Långström, N., Forsman, M., Sjölander, A., Fazel, S. & Molero, Y., *Associations between Prisons and Recidivism: A Nationwide Longitudinal Study*, PLoS ONE, **17**(5), e0267941, 2022. 10.1371/journal.pone.0267941.

[4] Ostermann, M., Link, N.W. & Hyatt, J. M., *Reframing the Debate on Legal Financial Obligations and Crime: How Accruing Monetary Sanctions Impacts Recidivism*, Criminology, **62**(2), pp. 331-363, 2024. 10.1111/1745-9125.12375.

[5] Beek, G.V., de Vogel, V. & Mheen, D.V.D., *Financial Problems and Debt as Predictive Factors for Recidivism,* Criminal Justice and Behavior, **50**(9), pp. 1341-1360, 2023. 10.1177/00938548231177708.

[6] Butler, A., Nicholls, T.L., Samji, H., Fabian, S. & Lavergne, M.R., *Mental Health Needs, Substance Use, and Reincarceration: Population-level Findings From a Released Prison Cohort*, Criminal Justice and Behavior, **51**(7), pp. 1054-1071, 2024. 10.1177/00938548241238327.

[7] Gibbs, D., Colledge-Frisby, S., Farnbach, S., Doyle, M., Shakeshaft, A. & Larney, S., *Associations between Supported Accommodation and Health and Re-offending Outcomes: A Retrospective Data Linkage Study*, Journal of Urban Health, **101**(1), pp. 80-91, 2024. 10.1007/s11524-023-00824-w.

[8] Verdugo-Martinez, A., Ronzon-Tirado, R. & Redondo-Rodriguez, N., *Personality Traits and Their Role in Intimate Partner Violence Recidivism: A 15-year Follow-up Study within A Prison Sample*, Personality and Individual Differences, **235**, 112969, 2025. 10.1016/j.paid.2024.112969.

[9] Campedelli, G.M. & Yaksic, E*., Survival of the Recidivistic? Revealing Factors Associated with the Criminal Career Length of Multiple Homicide Offenders*, Homicide Studies, **26**(3), pp. 244-268, 2022 10.1177/10887679211010882.

[10] Jacobs, L.A. & Skeem, J.L., *Neighborhood Risk Factors for Recidivism: for Whom Do They Matter*?, American Journal of Community Psychology, **67**(1-2), pp. 103-115, 2021. 10.1002/ajcp.12463.

[11] Pyrooz, D.C., Clark, K.J., Tostlebe, J.J., Decker, S.H. & Orrick, E., *Gang Affiliation and Prisoner Reentry: Discrete-time Variation in Recidivism by Current, Former, and Non-gang Status*, Journal of Research in Crime and Delinquency, **58**(2), pp. 192-234, 2021. 10.1177/0022427820949895.

[12] Monnery, B., Wolff, F.C. & Henneguelle, A., *Prison, Semi-liberty and Recidivism: Bounding Causal Effects in A Survival Model*, International

Review of Law and Economics, **61**, 105884, 2020. 10.1016/j.irle.2019.105884.

[13] Spivak, A.L., & Damphousse, K.R., *Who Returns to Prison? A Survival Analysis of Recidivism among Adult Offenders Released in Oklahoma, 1985–2004*, Justice Research and Policy, **8**(2), pp. 57-88, 2006. 10.3818/JRP.8.2.2006.57.

[14] Smith, L.G. & Akers, R.L., *A Comparison of Recidivism of Florida's Community Control and Prison: A Five-Year Survival Analysis*, Journal of Research in Crime and Delinquency, **30**(3), pp. 267-292, 1993. 10.1177/0022427893030003002.

[15] LaBriola, J., *Post-prison Employment Quality and Future Criminal Justice Contact*, RSF: The Russell Sage Foundation Journal of the Social Sciences, **6**(1), pp. 154-172, 2020. 10.7758/RSF.2020.6.1.08.

[16] Walters, G.D., *Does Criminal Versatility Predict Recidivism? Results from Cox Proportional Hazards and Binomial Logistic Regression Analyses*, Crime & Delinquency, **71**(13-14), pp. 4096-4115, 2025. 10.1177/00111287241282411.

[17] Iowa Department of Corrections, *Districts & Prisons*, 2024, (Online), Available: https://doc.iowa.gov/districts-prisons, (Accessed 10 December 2024).

[18] Kleinbaum, D.G. & Klein, M., *Survival Analysis: A Self-learning Text, Third Edition*, Blackwell, New York, 2012.

[19] Hosmer Jr, D.W., Lemeshow, S. & May, S., *Applied Survival Analysis: Regression Modeling of Time-to-event Data*, Wiley, New Jersey, 2008.

[20] Grambsch, P.M. & Therneau, T.M., *Proportional Hazards Tests and Diagnostics Based on Weighted Residuals*, Biometrika, **81**(3), pp. 515-526, 1994. 10.1093/biomet/81.3.515.

[21] Therneau, T.M. & Grambsch, P.M., *Modeling Survival Data: Extending the Cox Model*, Springer, New York, 2000.

[22] Sutradhar, R. & Austin, P.C., *Relative Rates Not Relative Risks: Addressing A Widespread Misinterpretation of Hazard Ratios*, Annals of Epidemiology, **28**(1), pp. 54-57, 2018. 10.1016/j.annepidem. 2017.10.014.

[23] Therneau, T. & Atkinso, E., *Concordance*, [Online]. Available: https://cran.r-project.org/web/packages/survival/vignettes/concordance.pdf, 2024, (Accessed 1 January 2025).

[24] Therneau, M.T. & Watson, D.A., *The Concordance Statistic and the Cox Model*, *Department of Health Science Research*, **85**, pp. 1-18, 2017.

[25] Longato, E., Vettoretti, M. & Di Camillo, B., *A Practical Perspective on the Concordance Index for the Evaluation and Selection of Prognostic Time-to-event Models*, Journal of Biomedical Informatics, **108**, 103496, 2020. 10.1016/j.jbi.2020.103496.

[26] Schober, P., Boer, C. & Schwarte, L.A., *Correlation Coefficients: Appropriate Use and Interpretation*, Anesthesia & Analgesia, **126**(5), pp. 1763-1768, 2018. 10.1213/ANE.0000000000002864.

[27] Lee Rodgers, J. & Nicewander, W.A., Thirteen Ways to Look at the Correlation Coefficient, *The American Statistician*, **42**(1), pp. 59-66, 1988. 10.1080/00031305.1988.10475524.

[28] Tollenaar, N. & Van Der Heijden, P.G., *Optimizing Predictive Performance of Criminal Recidivism Models Using Registration Data with Binary and Survival Outcomes*, PLoS ONE, **14**(3), e0213245, 2019. 10.1371/journal.pone.0213245.

[29] de la Cruz, R., Padilla, O., Valle, M.A. & Ruz, G.A., *Modeling Recidivism through Bayesian Regression Models and Deep Neural Networks*, Mathematics, **9**(6), 639, 2021. 10.3390/math9060639.

[30] Arfan, N., Irfanullah, A., Hamidi, M. R., & Mukhaiyar, U., *Comparison of Survival Analysis Using Accelerated Failure Time Model and Cox Model for Recidivist Case*, BAREKENG: J. Math. & App*.,* **19**(1), pp. 629-642, 2025. 10.30598/barekengvol19iss1pp0629-0642.

[31] Akaike, H., *A New Look at the Statistical Model Identification,* IEEE Transactions on Automatic Control, **19**(6), pp. 716-723, 1974. 10.1109/TAC.1974.1100705.

[32] Schwarz, G., *Estimating the Dimension of a Model*, *The Annals of Statistics*, **6**(2), pp. 461-464, 1978. 10.1214/aos/1176344136.