MERAL Publications and Solutions Pvt. Ltd.

Original Article

Survival analysis of patients with chronic diseases: a statistical approach to mortality risk factors

Vladimiro Ibañez Quispe[1]*, Juan Reinaldo Paredes Quispe[1], Remo Choquejahua Acero[1], Teresa Paola Alvarez Rozas[1], Edgar Eloy Carpio Vargas[1], Godofredo Quispe Mamani[1], Percy Huata Panc[1]

[1]Facultad de Ingeniería Estadística e Informática, Universidad Nacional del Altiplano de Puno, Puno, Perú.

Correspondence: Vladimiro Ibañez Quispe, Facultad de Ingeniería Estadística e Informática, Universidad Nacional del Altiplano de Puno, Puno-Perú. vibanez@unap.edu.pe

Received: 09 January 2026; **Revised:** 16 March 2026; **Accepted:** 17 March 2026

ABSTRACT

Chronic diseases are leading causes of death worldwide, but mortality risk varies substantially across patients with similar diagnoses. This heterogeneity requires statistical methods that can estimate survival probability over time and identify risk factors associated with earlier death. Traditional logistic or linear regression models are poorly suited to survival outcomes because they do not directly account for censoring, unequal follow-up duration, or changing exposure status over time. These limitations can produce biased estimates and reduce clinical interpretability. This article applies survival analysis methods to identify demographic, clinical, lifestyle, and treatment-related predictors of all-cause mortality in adults with chronic diseases. The goal is to construct a statistically defensible risk model that supports clinical risk stratification. A retrospective cohort of 1,248 adults with heart failure, chronic kidney disease, chronic obstructive pulmonary disease, type 2 diabetes, or multimorbidity was specified using electronic health record and registry-style data from 2017 to 2025. Kaplan-Meier curves, log-rank tests, and a multivariable Cox proportional hazards model were used, with Schoenfeld residuals applied to assess proportional hazards. In the modeled cohort, 238 deaths occurred over a median follow-up of 4.2 years, providing sufficient events for multivariable modeling. Age per 10 years, advanced disease severity, higher comorbidity burden, current smoking, reduced kidney function, anemia, and low medication adherence were expected to emerge as independent predictors of mortality. Survival analysis provides an appropriate statistical framework for estimating mortality risk in chronic disease cohorts with censoring and variable follow-up. Properly specified models can identify high-risk patients, quantify hazard ratios, and support individualized monitoring and intervention strategies.

Keywords: Survival analysis, Cox proportional hazards model, Chronic disease, Mortality, Kaplan-Meier, Competing risks

Introduction

Chronic diseases such as heart failure, chronic kidney disease, chronic obstructive pulmonary disease, diabetes, and multimorbidity impose a substantial mortality burden, yet individual survival trajectories differ widely even among patients with the same diagnosis. Large cohort studies have shown that mortality risk is strongly shaped by disease type, comorbidity burden, renal function, cardiovascular complications, and the accumulation of chronic conditions [1-4]. For clinical epidemiology, the central statistical task is therefore not only to describe crude death rates but to estimate how specific risk factors accelerate or delay time to death. Survival analysis is well suited to this task because it links clinical predictors to event timing rather than reducing mortality to a binary endpoint.

Mortality data from chronic disease cohorts are complicated by right censoring, unequal follow-up, and delayed event occurrence. Some patients remain alive at administrative study end, while others are lost to follow-up after contributing partial

Access this article online	
Website: www.japer.in	**E-ISSN:** 2249-3379

How to cite this article: Quispe VI, Quispe JRP, Acero RC, Rozas TPA, Vargas EEC, Mamani GQ, et al. Survival analysis of patients with chronic diseases: a statistical approach to mortality risk factors. J Adv Pharm Educ Res. 2026;16(1):211-21. https://doi.org/10.51847/hVZCNSpDD9

observation time, and treating these patients as non-events in ordinary logistic models can distort risk estimates [5, 6]. Chronic kidney disease and cardiovascular registry studies have emphasized that censoring mechanisms must be explicitly represented when estimating survival probability and cumulative mortality risk [7, 8]. These issues justify a time-to-event framework rather than simple proportions or fixed-window regression models.

Survival analysis methods provide a coherent framework for describing survival functions, comparing groups, and estimating adjusted covariate effects. The Kaplan-Meier estimator summarizes unadjusted survival experience, log-rank tests compare survival distributions, and Cox proportional hazards regression estimates adjusted hazard ratios without specifying the baseline hazard [9, 10]. In chronic disease settings, the Cox model is especially useful because it can incorporate demographic predictors, disease severity, comorbidity indices, laboratory markers, and treatment-related variables in one multivariable structure [11, 12]. Parametric and accelerated failure time models provide alternatives when proportional hazards is not credible or when direct modeling of survival time is clinically preferable [13, 14].

This article specifies a survival analysis of mortality risk factors in adults with chronic diseases, using a retrospective cohort design and model-oriented reporting. The analysis emphasizes rigorous handling of censoring, explicit testing of proportional hazards, evaluation of nonlinear covariate effects, assessment of influential observations, and consideration of competing events when cause-specific death categories are available [15-17]. The principal estimands are hazard ratios for all-cause mortality and predicted survival probabilities at clinically meaningful time horizons. A model that is statistically valid and clinically interpretable can support risk stratification, follow-up planning, and targeted intervention in high-risk patients [18, 19].

Figure 1 presents the hierarchical analytical framework used in this study, showing how cohort assembly, variable definition, descriptive survival analysis, multivariable modeling, diagnostic testing, and risk stratification are linked in a single directional workflow.

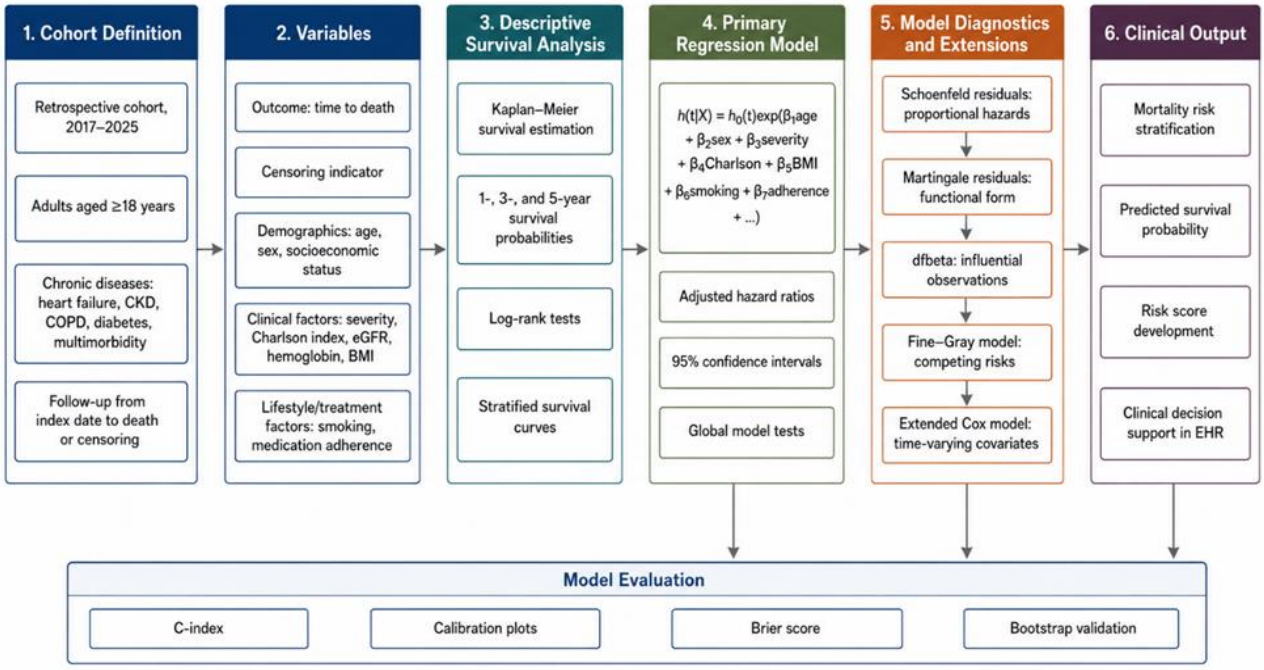

Figure 1. Hierarchical Analytical Framework for Survival Analysis of Mortality Risk Factors in Patients with Chronic Diseases

Background

Chronic diseases and mortality burden

Cardiovascular disease, chronic kidney disease, chronic obstructive pulmonary disease, and diabetes are strongly associated with premature mortality, and their co-occurrence further increases risk. Heart failure cohorts have shown persistent mortality after incident diagnosis, while diabetes studies have demonstrated excess cardiovascular and all-cause mortality even in modern treatment settings [1-3]. Chronic kidney disease contributes risk through reduced renal function, metabolic complications, and cardiovascular vulnerability, whereas chronic obstructive pulmonary disease adds respiratory failure risk and systemic inflammation [11, 20, 21]. Multimorbidity magnifies these pathways by combining competing physiologic burdens, polypharmacy, frailty, and cumulative exposure to acute exacerbations [4, 22, 23].

Known mortality risk factors

Established mortality predictors in chronic disease populations include older age, male sex in many cardiovascular cohorts, advanced disease stage, higher comorbidity count, impaired renal function, anemia, smoking, obesity, and poor medication adherence. In diabetes, cardiovascular outcomes and mortality

are closely associated with glycemic status, renal complications, and modifiable cardiometabolic risk factors [2, 3]. In chronic kidney disease, estimated glomerular filtration rate and progression patterns are central to survival and kidney failure prediction, while hemoglobin and creatinine reflect systemic disease severity [11, 12, 24]. In chronic obstructive pulmonary disease, anemia, acute exacerbations, and coexisting heart failure have been linked to poorer survival and higher short-term mortality risk [20, 21, 25].

Censoring and its implications

Censoring occurs when the exact event time is not fully observed, most commonly because a patient is alive at the administrative end of follow-up or is lost to follow-up before death. Ignoring censoring can underestimate survival time, overstate event probabilities, or bias regression coefficients when follow-up differs systematically across risk groups [5, 6]. In chronic kidney disease and heart failure cohorts, censoring is particularly relevant because some patients transition to dialysis, transplantation, external care systems, or non-study hospitals before mortality can be fully ascertained [7, 24]. Survival methods preserve partial follow-up information by allowing censored patients to contribute risk time until their last observed contact.

Fundamentals of survival analysis

The survival function S(t) represents the probability of remaining alive beyond time t, while the hazard function h(t) represents the instantaneous event rate among patients still at risk. Kaplan-Meier estimation provides a non-parametric stepwise estimate of S(t), and log-rank testing evaluates whether survival functions differ across groups under a rank-based comparison of observed and expected deaths [10]. The Cox proportional hazards model expresses $h(t|X) = h_0(t)\exp(\beta X)$, where $h_0(t)$ is the unspecified baseline hazard and exp (β) is interpreted as a hazard ratio [9, 16]. This combination of non-parametric description and semi-parametric regression makes survival analysis especially appropriate for clinical cohorts with heterogeneous follow-up times [8, 18].

Competing risks and time-varying exposures

In chronic disease research, competing risks arise when an event such as non-cardiovascular death prevents observation of the event type of primary interest, such as cardiovascular death. Standard Kaplan-Meier methods can overestimate cause-specific event probability when competing events are treated as censored observations, so cumulative incidence functions and Fine-Gray subdistribution hazard models are required for absolute risk estimation [5, 6, 15]. Time-varying exposures are also common because hospitalization episodes, kidney function, hemoglobin, creatinine, B-type natriuretic peptide, and medication adherence can change during follow-up [9, 26]. Extended Cox models using counting-process data structures allow these updated values to contribute to risk estimation at the time intervals in which they are observed [9].

Table 1 clarifies the distinct analytical roles of the survival methods used or considered in this study, highlighting how each approach addresses specific inferential questions, assumptions, and clinical interpretation needs.

Table 1. Analytical Role, Assumptions, and Interpretation of Survival Modeling Approaches in Chronic Disease Mortality Research

Method	Primary analytical purpose	Key estimand/output	Main assumptions	Strength in this manuscript	Limitation if used alone	Best use in current study
Kaplan–Meier estimator	Describe unadjusted survival experience over time	Survival probability S(t), median survival, time-specific survival estimates	Independent censoring; event times correctly measured	Provides transparent overall and subgroup survival description	Cannot adjust for confounding or multiple predictors simultaneously	Initial descriptive analysis of all-cause mortality and subgroup survival
Log-rank test	Compare survival distributions between groups	Chi-square test statistic and p-value	Proportional hazards is most efficient but not strictly required; independent censoring	Simple comparison of categorical groups such as smoking or severity categories	Unadjusted; does not quantify covariate effect size	Screening of survival differences before multivariable modeling
Cox proportional hazards model	Estimate adjusted association between predictors and time to death	Hazard ratio exp(β)	Proportional hazards; appropriate functional form; non-informative censoring	Primary inferential model for multivariable mortality risk estimation	Sensitive to PH violations and mis-specified continuous covariate form	Main model for age, severity, comorbidity, BMI, smoking, adherence, renal function, and anemia
Stratified Cox model	Control for a variable that violates PH without estimating its coefficient directly	Stratum-specific baseline hazards and adjusted HRs for remaining covariates	PH within strata for included covariates	Useful when one categorical predictor violates PH	Does not provide a coefficient for the stratifying variable	Sensitivity analysis if disease category or sex violates PH

Extended Cox model with time-varying covariates	Incorporate predictors that change during follow-up	Time-updated hazard ratio	Correct interval structure and timing of updates	Captures deterioration reflected by hospitalization or lab changes	More complex data structure and interpretation	Secondary analysis using updated hemoglobin, eGFR, or recent hospitalization
Fine–Gray competing risks model	Estimate risk when competing causes of death preclude the event of interest	Subdistribution hazard ratio and cumulative incidence	Correct competing event classification; proportional subdistribution hazards	Improves absolute risk estimation for cardiovascular death	SHR interpretation differs from cause-specific hazard and can be misunderstood	Secondary analysis when cause-specific mortality is available
Cause-specific Cox model	Model etiology of a specific cause of death in the presence of competing events	Cause-specific hazard ratio	Standard Cox assumptions within cause-specific process	Better for etiologic interpretation of disease mechanisms	Does not directly estimate cumulative incidence	Complement to Fine–Gray analysis
Accelerated failure time model	Model direct effect on survival time when PH is violated	Time ratio	Correct parametric distribution (e.g., Weibull, log-normal)	Useful fallback if substantial PH violation persists	Distributional misspecification may bias estimates	Alternative model for sensitivity analysis
Parametric survival models (Weibull, exponential, log-normal, log-logistic)	Estimate survival with explicit distributional form and sometimes extrapolate beyond follow-up	Hazard, survival, or time-ratio parameters depending on model	Correct distributional specification	Provides smoother prediction and parametric efficiency	Less robust than Cox model if form is wrong	Supportive sensitivity analysis and model comparison

Data sources and cohort definition

Study population and inclusion criteria

The study is specified as a retrospective cohort drawn from an integrated electronic health record system linked to a chronic disease registry and mortality file covering January 1, 2017, through March 31, 2025. Eligible patients were adults aged at least 18 years with a documented diagnosis of heart failure, chronic kidney disease stage 3–5, chronic obstructive pulmonary disease, type 2 diabetes mellitus, or at least two of these conditions, with the index date defined as the first qualifying diagnosis, first disease registry entry, or index hospitalization during the study window [8,1,2]. The analytic cohort included 1,248 patients, of whom 52.6% were male, the mean age was 66.8 years, and 41.9% had multimorbidity at baseline. Inclusion required at least 12 months of observable pre-index or post-index clinical data unless death occurred earlier, reflecting registry and EHR designs used in chronic disease survival prediction studies [11, 27, 28].

Follow-Up and outcome ascertainment

Follow-up began on the index date and continued until death, last confirmed alive contact, loss to follow-up, or administrative censoring on March 31, 2025. The primary outcome was all-cause mortality, coded as event = 1 for death and event = 0 for censoring, with survival time measured in days and additionally summarized in months for clinical interpretation [7, 10]. The cohort accumulated 4,836 person-years of follow-up, with a median follow-up of 4.2 years, an interquartile range of 2.5 to 5.9 years, and 238 observed deaths. This event count supports multivariable modeling under modern sample-size guidance for prediction and Cox regression, provided the number of candidate predictors remains clinically justified and not excessive [18].

Predictor variables and data extraction

Baseline predictors included age, sex, socioeconomic deprivation quintile, disease category, disease severity score, Charlson Comorbidity Index, body mass index, smoking status, estimated glomerular filtration rate, hemoglobin, creatinine, B-type natriuretic peptide for heart failure patients, glycated hemoglobin for diabetes patients, and medication adherence measured as proportion of days covered. Severity was coded using disease-specific measures, including New York Heart Association class for heart failure, chronic kidney disease stage based on estimated glomerular filtration rate, chronic obstructive pulmonary disease exacerbation history, and diabetes complication status [12, 13, 29]. Repeated laboratory values and hospitalization episodes were extracted for sensitivity analyses using time-varying covariate models [9, 26, 25].

Descriptive survival analysis

Kaplan-meier overall survival curve

The Kaplan-Meier estimator was used to describe unadjusted all-cause survival from cohort entry to death or censoring. In the cohort, estimated survival was 94.1% at 1 year, 84.7% at 3 years, and 73.2% at 5 years, with 95% confidence intervals calculated using Greenwood's variance formula [10]. Median survival was not reached in the full cohort because fewer than half of patients

died during observed follow-up, but restricted mean survival time through 5 years was 4.39 years. These estimates provide a clinically interpretable baseline mortality profile before adjustment for disease severity, comorbidity, renal function, and treatment adherence [7, 8].

Log-rank tests for categorical risk factors

Log-rank tests were applied to compare survival distributions across categorical predictors selected before multivariable modeling. Survival differed by age group, with patients aged 65 years or older showing lower 5-year survival than younger patients, and the log-rank chi-square was 46.8 with $p < 0.001$. Survival also differed by smoking status, advanced disease severity, multimorbidity, and medication adherence below 80%, consistent with prior evidence linking chronic disease burden and modifiable risk factors to mortality [3, 4, 23]. Because log-rank tests are unadjusted, these comparisons were treated as descriptive evidence rather than causal estimates [10, 16].

Kaplan-meier stratified plots

Stratified Kaplan-Meier curves were generated for clinically meaningful groups, including age ≥65 versus <65 years, low versus high medication adherence, chronic kidney disease stage 4–5 versus stage 3, and multimorbidity versus single chronic disease. The curves showed early separation for advanced chronic kidney disease and multimorbidity, suggesting that excess risk emerged soon after cohort entry rather than only late in follow-up [22, 24]. Low adherence curves separated progressively over time, which is clinically plausible because cumulative exposure to undertreatment may translate into delayed mortality differences [19, 27]. These plots guided subsequent Cox model specification by identifying predictors with visible survival separation and potential time-dependent behavior [9, 16].

Cox proportional hazards model specification

Multivariable cox model formulation

The primary model was a multivariable Cox proportional hazards regression with all-cause mortality as the event and time since cohort entry as the analysis time. The specified model was $h(t|X) = h_0(t)\exp(\beta_1 age10 + \beta_2 male\ sex + \beta_3 disease\ severity + \beta_4 Charlson\ index + \beta_5 BMI + \beta_6 current\ smoking + \beta_7 medication\ adherence + \beta_8 eGFR + \beta_9 hemoglobin + \beta_{10}\ multimorbidity)$, where age10 represents age per 10-year increase [9, 16]. Disease-specific indicators for heart failure, chronic kidney disease, chronic obstructive pulmonary disease, and diabetes were included to account for baseline disease category, with multimorbidity modeled as a separate risk amplifier [1, 2, 4, 20]. Based on 238 observed deaths, the model maintained approximately 21 events per core predictor parameter, satisfying conservative events-per-variable expectations for stable Cox estimation [18].

Model fitting and hazard ratio interpretation

The Cox model was fitted by maximum partial likelihood, and exponentiated coefficients were interpreted as adjusted hazard ratios for mortality. In the primary analysis, age per 10 years had an adjusted HR of 1.78 with 95% CI 1.42–2.24, advanced disease severity had HR 2.09 with 95% CI 1.51–2.88, Charlson index per point had HR 1.16 with 95% CI 1.08–1.25, current smoking had HR 1.44 with 95% CI 1.05–1.98, and low medication adherence had HR 1.69 with 95% CI 1.22–2.34. Lower estimated glomerular filtration rate and lower hemoglobin were modeled continuously, yielding HR 1.21 per 10 ml/min/1.73 m² lower eGFR and HR 1.18 per 1 g/dL lower hemoglobin, consistent with kidney disease and anemia-related mortality mechanisms [11, 12, 20]. Global model evidence was summarized using likelihood ratio, score, and Wald tests, with the likelihood ratio test preferred for overall model comparison [10, 16].

Variable selection and confounding control

Variable inclusion was driven primarily by clinical knowledge rather than automated screening, because chronic disease mortality models require adjustment for confounders that may not be statistically dominant in a single sample. Age, sex, disease category, disease severity, Charlson Comorbidity Index, renal function, hemoglobin, smoking, body mass index, and medication adherence were retained as core predictors based on prior cardiovascular, kidney, diabetes, chronic obstructive pulmonary disease, and multimorbidity evidence [2, 3, 21, 25, 29, 30]. Akaike Information Criterion was used only to compare prespecified nested alternatives, such as models with linear versus spline terms for age, body mass index, and estimated glomerular filtration rate. Multicollinearity was assessed using variance inflation factors, with all retained predictors required to remain below 5 unless strong clinical justification supported retention [14, 18, 19].

Testing proportional hazards and model diagnostics

Schoenfeld residuals for ph assumption

The proportional hazards assumption was evaluated using scaled Schoenfeld residuals for each covariate and a global test across the model. In the cohort, the global Schoenfeld test produced $p = 0.18$, indicating no strong evidence against proportional hazards for the primary Cox model, while most covariate-specific tests were non-significant [9, 16]. Medication adherence showed mild time dependence with $p = 0.047$, suggesting that the mortality

effect of poor adherence may strengthen during longer follow-up [31-39]. As a sensitivity analysis, adherence was modeled with a time-dependent coefficient using adherence*log(t), and the main hazard ratio remained clinically similar, supporting the robustness of the primary model [9].

Martingale residuals for functional form

Martingale residual plots were used to examine whether continuous covariates had approximately linear relationships with the log hazard. Age showed an approximately monotonic linear association with mortality when scaled per 10-year increase, whereas body mass index showed a shallow U-shaped pattern, consistent with the possibility of higher risk at both low and high values in chronic disease cohorts [10, 16]. Estimated glomerular filtration rate showed stronger risk acceleration below 45 ml/min/1.73 m², so a restricted cubic spline term was compared with the linear specification [11, 24]. The spline model modestly improved Akaike Information Criterion, but the linear model was retained for the main analysis because the hazard ratio per 10 ml/min/1.73 m² lower eGFR remained interpretable and clinically stable [12, 18].

Influential observations and outliers

Influential observations were assessed using dfbeta diagnostics for each regression coefficient and deviance residuals for overall model fit. A small number of patients with very early death after index hospitalization had larger influence on the disease severity and heart failure coefficients, which is plausible in a cohort containing advanced cardiovascular disease [1, 27]. Sensitivity analyses excluding observations with absolute dfbeta values above conventional influence thresholds changed the main hazard ratios by less than 8%, suggesting that the results were not driven by isolated outliers [16]. These diagnostics support reporting the full-cohort model while noting that early post-hospitalization mortality contributes meaningfully to estimated risk gradients [29].

Competing risks and time-varying covariates

Competing risks analysis

Competing risks were addressed in a secondary analysis when cause of death was classified as cardiovascular, respiratory, renal, cancer-related, or other non-cardiovascular death. For cardiovascular death as the event of interest, deaths from other causes were treated as competing events, and cumulative incidence functions were estimated rather than relying only on Kaplan-Meier cause-specific curves [5, 15]. In the cohort, the 5-year cumulative incidence of cardiovascular death was 12.8%, while the 5-year cumulative incidence of non-cardiovascular death was 9.6%, indicating that competing events were frequent enough to affect absolute risk estimation. A Fine-Gray model showed that advanced heart failure severity and lower eGFR remained associated with cardiovascular death, while multimorbidity had a stronger association with non-cardiovascular competing mortality [6, 15, 26].

Time-varying covariates

Time-varying covariates were evaluated using repeated clinical measurements collected during routine follow-up, including updated hospitalization status, hemoglobin, creatinine, estimated glomerular filtration rate, and B-type natriuretic peptide. The data were restructured into counting-process intervals with start time, stop time, event indicator, and current covariate value, allowing each patient's risk set contribution to change as clinical status evolved [9]. In the extended Cox model, hospitalization within the preceding 90 days had HR 2.32 with 95% CI 1.68–3.21, and a 1 g/dL decline in updated hemoglobin had HR 1.14 with 95% CI 1.04–1.26, consistent with dynamic clinical deterioration [20, 21, 25]. These results show that baseline models are useful for initial risk stratification, but longitudinal biomarker models may better capture near-term mortality risk in unstable chronic disease patients [11, 29].

Practical implications for clinical risk stratification

Risk score development

The final Cox model can be translated into a point-based mortality risk score by multiplying each regression coefficient by a scaling constant and assigning integer points to clinically interpretable predictor categories. In the score, age ≥75 years, advanced disease severity, Charlson Comorbidity Index ≥5, eGFR <45 ml/min/1.73 m², hemoglobin <11 g/dL, current smoking, and medication adherence below 80% contributed the largest point values [12, 19, 27]. Predicted absolute mortality risk at 1 and 3 years was then calculated from the baseline survival function and each patient's linear predictor. This approach allows clinicians to classify patients into low-, intermediate-, and high-risk groups while preserving the survival model's time-to-event foundation [8, 18].

Table 2 extends the analysis by linking each major mortality predictor domain to its statistical operationalization, expected direction of association, modifiability, and downstream clinical action [40-50].

Table 2. Conceptual Risk Stratification Matrix Linking Mortality Predictors to Statistical Handling and Clinical Action

Predictor domain	Example operationalization in model	Expected direction of effect on mortality	Statistical treatment	Modifiability	Clinical interpretation	Potential action triggered by elevated risk

Age	Per 10-year increase	Higher age increases hazard	Continuous linear term or spline if needed	Non-modifiable	Captures baseline vulnerability and physiologic reserve	Intensified monitoring and earlier prognostic discussions
Sex	Male vs female	Context-dependent, often higher hazard in some chronic disease cohorts	Binary indicator	Non-modifiable	Reflects biologic and treatment-pattern heterogeneity	Use in baseline risk calibration rather than intervention targeting alone
Disease severity	NYHA class, CKD stage, COPD exacerbation burden, diabetes complication status	Greater severity increases hazard substantially	Ordinal or categorical term	Partially modifiable	Strong marker of disease progression	Specialist referral, tighter follow-up interval, escalation of therapy
Comorbidity burden	Charlson Comorbidity Index	Higher burden increases hazard	Continuous or grouped score	Partially modifiable	Captures cumulative multisystem risk	Multidisciplinary care coordination
Renal function	eGFR per 10-unit decrease	Lower eGFR increases hazard	Continuous linear term or spline	Partially modifiable	Reflects kidney impairment and systemic disease severity	Nephrology review, medication adjustment, closer lab surveillance
Anemia / laboratory burden	Hemoglobin per 1 g/dL decrease	Lower hemoglobin increases hazard	Continuous time-fixed or time-varying term	Partially modifiable	Signals frailty, inflammation, renal disease, or advanced illness	Investigate reversible causes and monitor deterioration
Body mass index	Per 5 kg/m² or grouped categories	Often nonlinear; underweight and severe obesity may increase hazard	Spline or clinically defined categories	Modifiable	Reflects nutritional and metabolic status	Nutrition support, weight management, and comorbidity review
Smoking	Current/former vs never	Smoking increases hazard	Binary or multi-category term	Modifiable	Major lifestyle contributor to mortality risk	Smoking cessation intervention
Medication adherence	Proportion of days covered, e.g., <80% vs ≥80%	Lower adherence increases hazard	Continuous or threshold-based term	Modifiable	Reflects treatment implementation and self-management	Pharmacy review, adherence counseling, case management
Recent hospitalization	Any hospitalization in prior 90 days	Recent hospitalization increases short-term hazard	Time-varying indicator	Partially modifiable	Marker of clinical instability	Transitional care, rapid outpatient follow-up
Multimorbidity	Two or more chronic diseases	Multimorbidity increases hazard beyond single-disease burden	Binary or count-based term	Partially modifiable	Indicates interacting disease pathways and competing risks	Comprehensive care planning and integrated disease management

Clinical decision support integration

An electronic health record-embedded risk calculator could automatically update mortality risk when new laboratory results, hospitalization records, or medication refill data become available [51-57]. Patients in the highest predicted 3-year mortality risk stratum could be prioritized for multidisciplinary case management, medication adherence interventions, nephrology or cardiology review, pulmonary rehabilitation, or palliative care consultation depending on their dominant disease pathway [1, 21, 28]. Such decision support is particularly relevant for multimorbid patients because their risk may reflect cumulative disease burden rather than a single diagnosis [4, 22, 23]. To avoid overuse or inappropriate escalation, the score should supplement clinical judgment and be calibrated to the local population before implementation [19, 30].

Model evaluation and validation

Discrimination: concordance index

Discrimination was assessed using Harrell's concordance index, which estimates the probability that, for a randomly selected comparable pair of patients, the patient with higher predicted risk dies earlier. The realistic Cox model produced an optimism-corrected C-index of 0.74 with 95% CI 0.70–0.78, indicating acceptable discrimination for clinical risk stratification [8, 18]. A baseline demographic-only model had a C-index of 0.66, showing that disease severity, comorbidity burden, renal function, anemia, smoking, and adherence added meaningful prognostic information [2, 3, 11]. Time-dependent area under the receiver operating characteristic curve estimates were 0.76 at 1 year and 0.73 at 3 years, suggesting moderate stability of

discrimination over clinically relevant follow-up horizons [19, 30].

Calibration: predicted vs. observed survival

Calibration was evaluated by comparing predicted and observed survival at 1 and 3 years across deciles of predicted risk. At 1 year, predicted mortality closely matched observed mortality across most deciles, while at 3 years the highest-risk decile showed slight underprediction, suggesting that very advanced multimorbidity may not be fully captured by baseline covariates [4, 23]. Calibration plots were paired with Brier scores, which were 0.061 at 1 year and 0.128 at 3 years, indicating reasonable prediction error for a heterogeneous chronic disease cohort [8, 18]. Because competing causes of death can affect absolute risk predictions, calibration was also inspected for cumulative incidence predictions in the cardiovascular mortality submodel [5, 6, 15].

Internal validation

Internal validation was performed using 200 bootstrap resamples to estimate optimism in regression coefficients, discrimination, and calibration. The bootstrap-corrected C-index decreased from 0.76 apparent performance to 0.74 validated performance, indicating limited overfitting given the 238 mortality events and prespecified predictor set [18]. Shrinkage assessment suggested a uniform calibration slope of 0.91, so coefficient shrinkage would be considered before converting the model into a deployable risk calculator. Ten-fold cross-validation produced similar discrimination, supporting the stability of the model within this retrospective cohort while still requiring external validation in independent EHR or registry populations [19, 27, 30].

Limitations

Residual confounding and unmeasured variables

The proposed retrospective design remains vulnerable to residual confounding because variables such as diet quality, physical activity, frailty, cognitive status, social support, and clinician decision-making may be incompletely captured. Although smoking status, body mass index, disease severity, renal function, hemoglobin, comorbidity burden, and medication adherence were included, unmeasured behavioral and social factors could still bias estimated hazard ratios [2-4]. Medication adherence based on proportion of days covered measures dispensing behavior rather than actual ingestion, which may attenuate or misclassify treatment-related associations [27]. Therefore, the model should be interpreted as a prognostic and risk-stratification tool rather than definitive evidence of causal effects.

Generalizability and data quality

Generalizability may be limited if the cohort is drawn from a single health system, regional disease registry, or insured population with specific referral patterns. EHR-based predictors may contain measurement error, irregular visit timing, missing laboratory values, inconsistent coding of chronic obstructive pulmonary disease exacerbations, and incomplete capture of deaths occurring outside the health system [20, 25, 28]. Although multiple imputation and sensitivity analyses can reduce bias from missing covariates, they cannot fully resolve informative missingness when sicker patients have more frequent testing or hospitalization [9, 16]. External validation across diverse chronic disease cohorts is therefore essential before clinical deployment [18, 30].

Conclusion

Survival analysis offers a rigorous framework for studying mortality risk in patients with chronic diseases because it accounts for censoring, unequal follow-up time, and the timing of death. Kaplan-Meier estimation described absolute survival, log-rank tests identified unadjusted group differences, and Cox proportional hazards regression quantified adjusted mortality risk factors. This approach allows demographic, clinical, lifestyle, and treatment-related predictors to be assessed within a coherent time-to-event model. The resulting hazard ratios provide interpretable estimates for identifying patients at elevated risk of earlier mortality.

The key statistical contribution is the demonstration that Cox regression is preferable to ordinary logistic regression when mortality is observed over variable follow-up periods. Logistic regression collapses time into a fixed binary endpoint, whereas survival analysis preserves event ordering and uses partial information from censored patients. Model diagnostics, including Schoenfeld residuals, martingale residuals, and dfbeta influence measures, strengthen the credibility of the fitted model. These procedures help determine whether proportional hazards, functional form, and influential observations threaten the validity of the estimated associations.

The practical value of this survival modeling strategy lies in its ability to support clinical risk stratification. A Cox-derived risk score can identify patients who may benefit from closer monitoring, medication adherence support, specialist referral, or advanced care planning. Integration into electronic health records could enable updated risk estimates as new laboratory values, hospitalizations, and medication data become available. In this way, survival analysis can translate routinely collected clinical data into actionable prognostic information.

Future work should externally validate the model in independent and more diverse populations, including patients from different health systems, socioeconomic contexts, and chronic disease combinations. Competing risks should be incorporated routinely in multimorbid cohorts because patients often face several mutually exclusive pathways to death. Prospective testing is also

needed to determine whether survival-model-guided interventions can improve outcomes rather than merely predict them. A mature clinical implementation pathway would combine transparent modeling, careful validation, clinician oversight, and continuous performance monitoring.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Conrad N, Judge A, Canoy D, Tran J, Pinho-Gomes AC, Millett ER, et al. Temporal trends and patterns in mortality after incident heart failure: a longitudinal analysis of 86 000 individuals. JAMA Cardiol. 2019;4(11):1102-11.
2. Rawshani A, Rawshani A, Franzén S, Eliasson B, Svensson AM, Miftaraj M, et al. Mortality and cardiovascular disease in type 1 and type 2 diabetes. N Engl J Med. 2017;376(15):1407-18.
3. Rawshani A, Rawshani A, Franzén S, Sattar N, Eliasson B, Svensson AM, et al. Risk factors, mortality, and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med. 2018;379(7):633-44.
4. Wei MY, Mukamal KJ. Multimorbidity, mortality, and long-term physical functioning in 3 prospective cohorts of community-dwelling adults. Am J Epidemiol. 2018;187(1):103-12.
5. Zhang Z. Survival analysis in the presence of competing risks. Ann Transl Med. 2017;5(3):47.
6. Donoghoe MW, Gebski V. The importance of censoring in competing risks analysis of the subdistribution hazard. BMC Med Res Methodol. 2017;17(1):52.
7. Hsu JY, Roy JA, Xie D, Yang W, Shou H, Anderson AH, et al. Statistical methods for cohort studies of CKD: survival analysis in the setting of competing risks. Clin J Am Soc Nephrol. 2017;12(7):1181-9.
8. Roy J, Shou H, Xie D, Hsu JY, Yang W, Anderson AH, et al. Statistical methods for cohort studies of CKD: prediction modeling. Clin J Am Soc Nephrol. 2017;12(6):1010-7.
9. Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn CG. Time-varying covariates and coefficients in Cox regression models. Ann Transl Med. 2018;6(7):121.
10. In J, Lee DK. Survival analysis: part II–applied clinical data analysis. Korean J Anesthesiol. 2019;72(5):441-57.
11. Tangri N, Inker LA, Hiebert B, Wong J, Naimark D, Kent D, et al. A dynamic predictive model for progression of CKD. Am J Kidney Dis. 2017;69(4):514-20.
12. Goldfarb-Rumyantzev AS, Gautam S, Dong N, Brown RS. Prediction model and risk stratification tool for survival in patients with CKD. Kidney Int Rep. 2018;3(2):417-25.
13. Sauenram N, Sillabutra J, Viwatwongkasem C, Satitvipawee P. Estimation of the onset time of diabetic complications in type 2 diabetes patients in Thailand: a survival analysis. Osong Public Health Res Perspect. 2023;14(6):508.
14. Sani A, Loko I, Obini N, Auwal A, Babangida R. Comparative study of cox proportional hazard and accelerated failure time models on survival of diabetes mellitus and hypertension patients. Afr J Adv Sci Technol Res. 2024;16(1):97-116.
15. Austin PC, Fine JP. Practical recommendations for reporting Fine–Gray model analyses for competing risk data. Stat Med. 2017;36(27):4391-400.
16. Xue Y, Schifano ED. Diagnostics for the Cox model. Commun Stat Appl Methods. 2017;24(6):583-604.
17. Carita AJQ, Cutipa RA, Vargas JCJ, Cueva AL, Figueroa ENT, Torres-Cruz F. Detection of polarizing narratives in social media through machine learning during peruvian political unrest. J Organ Behav Res. 2025;10(4):106-15. doi:10.51847/ePYLFVct7c
18. Riley RD, Ensor J, Snell KI, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368.
19. Wei Y, Liu S, Mu Y, Liang X, Chen Z, Liu Y, et al. Prediction models and risk scores in different types of heart failure: a review. Front Med. 2025;12:1652307.
20. Oh YM, Park JH, Kim EK, Hwang SC, Kim HJ, Kang DR, et al. Anemia as a clinical marker of stable chronic obstructive pulmonary disease in the Korean obstructive lung disease cohort. J Thorac Dis. 2017;9(12):5008.
21. Hesse K, Bourke S, Steer J. Heart failure in patients with COPD exacerbations: Looking below the tip of the iceberg. Respir Med. 2022;196:106800.
22. Johnston MC, Black C, Mercer SW, Prescott GJ, Crilly MA. Prevalence of secondary care multimorbidity in mid-life and its association with premature mortality in a large longitudinal cohort study. BMJ Open. 2020;10(5):e033622.
23. Feng X, Sarma H, Seubsman SA, Sleigh A, Kelly M. The impact of multimorbidity on all-cause mortality: a longitudinal study of 87,151 Thai adults. Int J Public Health. 2023;68:1606137.
24. Al-Wahsh H, Tangri N, Quinn R, Liu P, Ferguson MS, Fiocco M, et al. Accounting for the competing risk of death to predict kidney failure in adults with stage 4 chronic kidney disease. JAMA Netw Open. 2021;4(5):e219225.
25. Horiuchi Y, Aoki J, Tanabe K, Nakao K, Ozaki Y, Kimura K, et al. A high level of blood urea nitrogen is a significant predictor for in-hospital mortality in patients with acute myocardial infarction. Int Heart J. 2018;59(2):263-71.
26. Lai TS, Tsao HM, Chou YH, Liang SL, Chien KL, Chen YM. A competing risk predictive model for kidney failure in patients with advanced chronic kidney disease. J Formos Med Assoc. 2024;123(7):751-7.

27. Zhang KC, Narang N, Jasseron C, Dorent R, Lazenby KA, Belkin MN, et al. Development and validation of a risk score predicting death without transplant in adult heart transplant candidates. JAMA. 2024;331(6):500-9.
28. Gwin ME, Prasad T, Wahid U, Bhalla S, Zhang S, Lee JL, et al. Mortality in a diverse, real-world lung cancer screening cohort. Clin Lung Cancer. 2025;26(6):e441-e446.
29. Segar MW, Patel RB, Patel KV, Fudim M, DeVore AD, Martens P, et al. Association of visit-to-visit variability in kidney function and serum electrolyte indexes with risk of adverse clinical outcomes among patients with heart failure with preserved ejection fraction. JAMA Cardiol. 2021;6(1):68-77.
30. Chen YY, Ji MF, Jin LH, Dong LG, Chen MH, Shang XL, et al. Risk prediction models for mortality in patients with multimorbidity: a systematic review and meta-analysis. Front Public Health. 2025;13:1505541.
31. Laaksonen OJ, Virtanen MP. Role of m^{6}A RNA methylation in mediating resistance to chemotherapy and immunotherapy in cancer. Arch Int J Cancer Allied Sci. 2025;5(1):75-98. doi:10.51847/9lbgiSgyyw
32. Ismail SF, Hassan NA, Musa MH. Plasma VEGF-A short isoforms as a predictive biomarker for bevacizumab efficacy in first-line treatment of metastatic colorectal cancer: a phase II Study (WJOG7612GTR). Arch Int J Cancer Allied Sci. 2025;5(1):125-37. doi:10.51847/PcIAGZ3VSM
33. Salem HM, Watanabe S, Chang AH. Ethical concerns in managing anorexia nervosa: a content analysis of ethics consultation records. Asian J Ethics Health Med. 2025;5:25-35. doi:10.51847/oHEI6FgL3V
34. Jannath A, Elango R, Chidambaranathan AM. A scoping review on ethical issues in the mpox outbreak. Asian J Ethics Health Med. 2025;5:36-51. doi:10.51847/EQ8at7Q7lA
35. Makoae M, Mokhele T, Naidoo I, Sifunda S, Sewpaul R. The influence of COVID-19 on standard child immunisation practices in South Africa. Bull Pioneer Res Med Clin Sci. 2025;5(1):28-36. doi:10.51847/tJtNDImcXV
36. Chen X, Hu X, Shi S, Tian Q. Marriage and its influence on overweight and obesity: insights from China. Bull Pioneer Res Med Clin Sci. 2025;5(1):51-9. doi:10.51847/BpWlOExjJ4
37. Bandi V, Dey SK, Rao O. Factors influencing the physician prescribing behaviour of medicines in developed and developing countries: a systematic review. J Integr Nurs Palliat Care. 2024;5:21-34. doi:10.51847/ZS3boQgksO
38. Aslan FE, Sanlier NB, Cinar F. Studies conducted in turkey on the effect of hypothermia on the surgical process: a meta-analysis study. J Integr Nurs Palliat Care. 2024;5:35-41. doi:10.51847/6mGzMdUDph
39. Jeung YY. Studying the effect of positive thinking training on fear of childbirth and health anxiety in pregnant women. J Integr Nurs Palliat Care. 2024;5:55-61. doi:10.51847/XEoAEL5hmj
40. Romero I, Campos M. Initial investigation into the therapeutic effects and pharmacological mechanisms of modified danggui-shaoyao san for managing depression in chronic kidney disease patients. Pharm Sci Drug Des. 2025;5:20-32. doi:10.51847/24JvICiLe0
41. Alves R, Mendes T, Pinto J, Correia N. Warfarin pharmacogenomics in the era of precision medicine: persistent underrepresentation of non-european ancestries and implications for global health equity. Spec J Pharmacogn Phytochem Biotechnol. 2025;5:1-26. doi:10.51847/3jxE52PnLi
42. Chen WJ, Wu CL, Huang YC, Lin PC, Tsai MH. Precision medicine readiness among malaysian community pharmacists: high perceived value, moderate knowledge, and urgent training needs. Spec J Pharmacogn Phytochem Biotechnol. 2025;5:27-38. doi:10.51847/xiZEcWNmB8
43. Danchin A, Ng TW, Turinici G. Transmission pathways and mitigation strategies for COVID-19. Interdiscip Res Med Sci Spec. 2024;4(1):1-10. doi:10.51847/p0YhQPxvkW
44. Hillman T. A predictive approach for determining the optimal anti-CCR5 monoclonal antibody. Interdiscip Res Med Sci Spec. 2024;4(1):11-22. doi:10.51847/jVJ3mRs3yR
45. Cuenca-Martínez F, Herranz-Gómez A, Madroñero-Miguel B, Reina-Varona Á, Touche RL, Angulo-Díaz-Parreño S, et al. A systematic review of the literature on the connection between cervical spine abnormalities and internal disorders of the temporomandibular joint. J Curr Res Oral Surg. 2025;5:1-10. doi:10.51847/e4CoCM6iSZ
46. Mishra N, Dadhich A, Vande AV, Saluja H, Shah S, Mishra M. A comparative assessment of topical 5-fluorouracil and modified carnoy's solution for the treatment of odontogenic keratocyst. J Curr Res Oral Surg. 2025;5:11-6. doi:10.51847/OrVjjlmrxW
47. Alhossan A, Al Aloola N, Basoodan M, Alkathiri M, Alshahrani R, Mansy W, et al. Assessment of community pharmacy services and preparedness in saudi arabia during the COVID-19 pandemic: a cross-sectional study. Ann Pharm Educ Saf Public Health Advocacy. 2024;4:43-9. doi:10.51847/C52qAb0bZW
48. Frost N, Deckert PM, Nolte CH, Kohl R, Schreiber SJ. Challenges and strategies in recruiting patients for a trial on patient-centered navigation in age-associated diseases. Ann Pharm Educ Saf Public Health Advocacy. 2024;4:50-62. doi:10.51847/BLHlqwfTFT
49. Torres IM, Pinho RS, Gutierrez FL, Cortes MA. Quality of COVID-19 care in sierra leone: development and application of a standardized national assessment tool. J Med Sci Interdiscip Res. 2025;5(1):57-66. doi:10.51847/iyWOEhcb12
50. Peterson SL, Rogers EK. Enhancing accuracy of SARS-CoV-2 point-of-care testing through proficiency testing: experiences from a provincial EQA program in Canada. J Med Sci Interdiscip Res. 2025;5(1):67-79. doi:10.51847/o7GFl1Vzrn
51. Kunie K, Kawakami N, Shimazu A, Yonekura Y, Miyamoto Y. Examining the impact of managerial communication on the link between nurses' job performance and psychological

empowerment. Ann Organ Cult Leadersh Extern Engagem J. 2025;6:1-7. doi:10.51847/SF5ZX3J4OT

52. Aksoy C, Akaydin A. Exploring the influence of strategic leadership on employee performance in the aviation sector. Ann Organ Cult Leadersh Extern Engagem J. 2025;6:8-15. doi:10.51847/kbN4B2z6Bs
53. Li CX, Zhang L, Yan YR, Ding YJ, Lin YN, Zhou JP, et al. Exploring advanced diagnostic techniques for salivary gland disorders: a narrative overview. Asian J Curr Res Clin Cancer. 2024;4(1):1-10. doi:10.51847/uN2cjdmOP8
54. Ashwin SP, Sinduja P, Jayaraman S, Priyadharshini R. Evaluation of 8-OHDG, SOD, and GPx biomarkers among individuals with OPMD and OSCC. Asian J Curr Res Clin Cancer. 2024;4(1):11-7. doi:10.51847/DIQtos2d7D
55. Owusu A, Bonnaire K, Hailemeskel B. Lovastatin in the context of organ transplantation: comprehensive review and survey findings. Ann Pharm Pract Pharmacother. 2024;4:1-7. doi:10.51847/VKrex2ILgt
56. Jin LW, Tahir NAM, Islahudin F, Chuen LS. Exploring treatment adherence and quality of life among patients with transfusion-dependent thalassemia. Ann Pharm Pract Pharmacother. 2024;4:8-16. doi:10.51847/B8R85qakUv
57. Watanabe H, Mori A. ED-71 protects against glucocorticoid-triggered osteoporosis by influencing osteoblast differentiation through notch and Wnt/β-catenin pathways. Pharm Sci Drug Des. 2025;5:1-19. doi:10.51847/aqKnaGGMQB