International Journal of Frontline Research in
Life Sciences
Journal homepage: https://frontlinejournals.com/ijfrls/
ISSN: 2945-4832 (Online)

(REVIEW ARTICLE)

Check for updates

Survival analysis in U.S. chronic disease research: A systematic review of methods and applications

Timothy Kizza [1], Mantey Joshua Kojo Aduampong [2] and Felix Kaiser [3,*]

[1] *Department of Public Health, Georgia Southern University, Georgia.*
[2] *Department of Chemical Engineering, Northeastern University — Boston, MA.*
[3] *Faculty of Pharmacy and Pharmaceutical Sciences, Kwame Nkrumah University of Science and Technology, Ghana.*

International Journal of Frontline Research in Life Science, 2025, 03(02), 026-034

Publication history: Received on 09 August 2025; revised on 23 September 2025; accepted on 26 September 2025

Article DOI: https://doi.org/10.56355/ijfrls.2025.3.2.0026

Abstract

Background: Chronic diseases continue to be the leading cause of morbidity and mortality in the United States; therefore, advanced statistical models are needed to both quantify the natural history of these outcomes and factor time-to-event components into data analysis. The field of survival analysis, a bedrock in longitudinal data analysis, which is essential for estimating hazards, prognosis, and population health measures.

Objective of the review: This study used a systematic approach to synthesize the utilization of survival analysis methods in U.S.-based chronic disease research, identify trends in methodology, appraise methodological quality, and recommend ways to improve it.

Methods: A search was performed in peer-reviewed studies from 2010 to 2024 using Google Scholar, PubMed, Embase, and Web of Science according to PRISMA 2020 guidelines that employed survival models on chronic disease outcomes in U.S. populations. Disease focus, survival outcomes model used, data sources, model assumptions, and reporting practices were extracted from studies eligible for inclusion. The risk of bias was described regarding the Quality in Prognostic Studies tool and synthesized using either narrative or frequency tables, as appropriate.

Results: Forty-six studies out of 278 records screened met the inclusion criteria. The most prevalent method was the Cox hazards model (78%), followed by Kaplan-Meier estimators, parametric models, and a few applications of machine learning-based survival models. The majority of studies used national datasets (e.g., SEER, NHANES) and few reported evaluating model assumptions or external validation. The most common deficiencies were in handling time-dependent covariates and censoring.

Conclusion: Survival analysis is fundamental in chronic disease research; however, its application frequently lacks methodological discipline. Amplified use of best-available models, thorough diagnostics, and transparent reporting are necessary to meet the advancing expectations of precision medicine and an ageing population. In addition, future research is needed to support innovations in methods and capacity building of biostatistics in applied research.

Keywords: Survival analysis; Chronic disease; Cox regression; Kaplan-Meier; Machine learning

1. Introduction

Cardiovascular disease (CVD), cancer, diabetes, and chronic respiratory condition-related deaths and disability are the most common causes in the United States [1]. Evidence suggests that out of every American adult, about 60% have at

* Corresponding author: Felix Kaiser

least one chronic illness, and 40% possess two or more [2]. The seven out of every 10 deaths witnessed each year can be explained by these conditions as well as the explanation of most healthcare expenses in any nation. Demographic ageing, social factors of health, and health care disparities are additional factors that have contributed to the burden since their impact is significant in racial/ethnic minorities and low-income groups [3]. Chronic diseases have a long course, unpredictable progression rates, and dependence on lifestyle and treatment conditions; therefore, they cannot be analyzed with the help of tools that can observe long-term results and temporal trends. This necessity supports the increased significance of time-to-event data analysis in the epidemiology of chronic diseases [4].

In research of chronic diseases, time to hospitalization, disease progression, relapse, or death are all outcome measures that can be used in assessing the efficacy of the treatment, determining the disease course, and assisting in clinical advice [5]. These may be censored, which implies that in some cases, the event of interest has not been met before the end of the study or the patients became lost to follow-up [6]. Conventional methods of regression cannot be used in such instances because they either disregard the censoring or estimate it in a biased way. Survival analysis, namely time-to-event analysis, is explicitly developed to overcome such issues [4]. It enables researchers to model the occurrence of an event and when it happens, in addition to properly capturing the censoring and bias in follow-up times [7].

Survival analysis plays a central role in longitudinal health research by providing a robust statistical framework for evaluating duration-based outcomes [8, 9]. In chronic disease studies, it is applied to risk prediction, such as estimating the probability of death within five years after diagnosis [10], treatment comparison, such as evaluating survival differences between two medication groups [11], and identification of prognostic factors, such as determining which biomarkers are associated with longer survival [12]. It is also used in health policy modeling, including forecasting the long-term impact of screening programs [13]. Beyond clinical trials and cohort studies, these methods are increasingly used in health services research, cost-effectiveness analyses, and population health surveillance [14, 15].

The increasing methodological sophistication in survival analysis over the recent decades has provided a rich set of tools for time-to-event data analysis to researchers in chronic disease research. In the detection of event rates, non-parametric methods including the Kaplan-Meier estimator are still used extensively [16, 17] as they are distribution-free and do not rely on any assumptions regarding event distributions [15]. A major advancement occurred in 1972 with the Cox proportional hazards model allowing for multivariable adjustment and omitting specification of a baseline hazard function [18, 9, 19]. Parametric survival models (exponential, Weibull, and Gompertz distributions), which are valid when assumptions regarding the form of the survival-time distribution can be made, offer improved predictive accuracy and interpretability as demonstrated by Voeltz [20] and others [21]. Competing risks models have also been used in more complex scenarios where an individual may experience more than one mutually exclusive outcome, like competing with death from different causes [21, 2, 22]. The use of machine learning models for survival data has also been explored in the context of random survival forests, deep learning-based survival networks, and penalized Cox models that are particularly advantageous when dealing with highly-dimensional data as well as complex non-linear relationships, which have been emerging over the past recent years [8, 23, 24]. Despite this increased methodological repertoire, traditional models like Cox regression still predominate in applied research (though it is unclear whether due to methodological appropriateness, user familiarity, or inertia) [24].

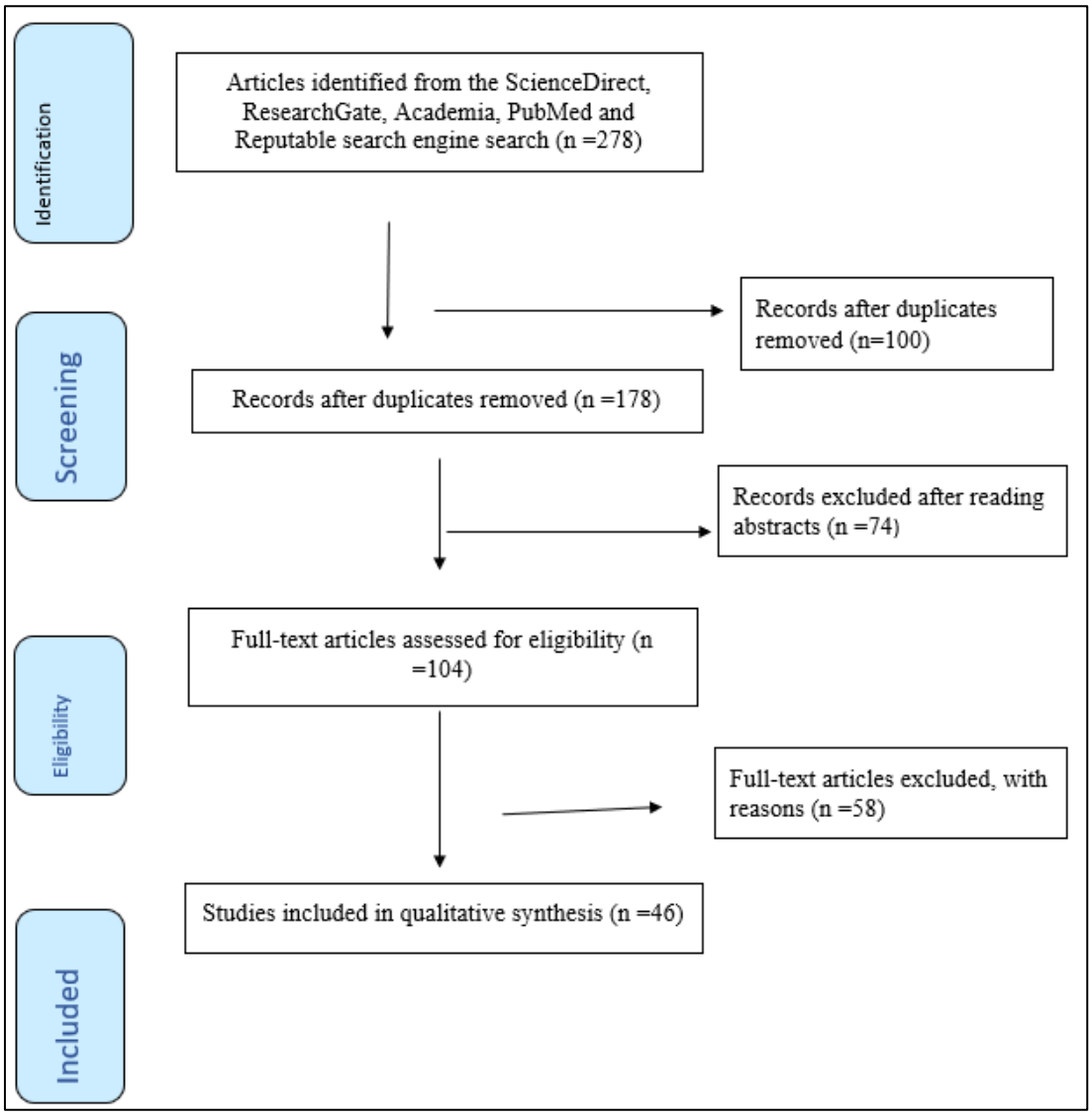
Despite the extensive application of survival analysis in chronic disease research, there are no published national-level assessments of the methodology employed with U.S.-based studies [25]. We considered whether particular survival models are more or less frequently applied, how often assumptions such as proportional hazards are both tested and reported, and the degree to which techniques like time-dependent covariates have been adopted in published work. In addition, due to the vast availability of longitudinal public datasets in the USA alone (NHANES, SEER, Medicare and Framingham Heart Study etc.,), knowledge on trends and best practices in survival analysis is crucial [25]. The identification of methodological gaps may shape training curricula and common research practice could potentially improve the quality and reproducibility in U.S. chronic disease research [25].

This systematic review aims to assess the quality of survival analysis methods in the last decade among chronic disease research within a U.S.-based setting. More specifically, it aims to distinguish some of these survival analysis methods widely used in the multiple domains of chronic disease literature and appraise their methodological quality, transparency, and statistical reporting according to these models. It also explores the distribution of these techniques according to disease types and populations, as well as data sources. It also illuminates important new directions for advanced modeling methodology more generally, such as competing risks models and machine learning methodologies. Finally, we discuss the implications of these gaps in our understanding for future research and policy initiatives designed to foster better patient outcomes among Americans with chronic conditions.

2. Methods

This systematic review followed PRISMA 2020 guidelines, using the PICOS framework to define eligibility. Studies included U.S.-based human participants (adults or older adolescents) with or at risk for chronic diseases such as cardiovascular disease, diabetes, cancer, chronic kidney disease, or chronic respiratory disease. Eligible designs were peer-reviewed observational studies or trials published in English between 2010 and July 2025 that applied survival analysis methods (e.g., Kaplan–Meier, Cox regression, competing risks, parametric, or machine learning models) to time-to-event outcomes such as mortality, disease progression, readmission, or relapse. Exclusions were studies on acute/infectious diseases, non-U.S. populations without U.S.-specific analysis, conference abstracts, editorials, letters, dissertations without original data, and non-human research.

Data base search was conducted on PubMed/MEDLINE, Embase, Web of Science, and Scopus on July 15, 2025, with strategies developed alongside a health sciences librarian, combining keywords and subject headings for survival analysis, chronic disease, and U.S. populations. Reference lists of relevant studies and reviews were also screened. Screening and selection were managed in Covidence. Duplicates were removed, and two reviewers independently screened titles/abstracts and full texts. Discrepancies were resolved by discussion or a third reviewer. The selection process is illustrated in Figure 1 (PRISMA flow diagram).



Sources; Author's Construct 2020.

Figure 1 PRISMA Flow diagram showing the article selection process in the study

2.1. Data Extraction Process

A standardized data extraction form was developed and piloted on a subset of studies to ensure consistency and comprehensiveness. The following data were extracted

- **Study characteristics:** Author, year, journal, data source (e.g., SEER, NHANES), sample size, follow-up duration.
- **Disease focus:** E.g., cancer, diabetes, cardiovascular disease.
- **Survival methods used:** Kaplan-Meier, Cox model, parametric models, competing risks, machine learning survival models.
- **Model features:** Software used, covariate inclusion, assumption testing (e.g., proportional hazards), model diagnostics, validation techniques.
- **Outcome type:** E.g., all-cause mortality, disease-specific survival, time to readmission.

Data were extracted independently by two reviewers. Discrepancies were resolved through consensus.

2.2. Risk of Bias Assessment

The methodological quality of included studies was assessed using a modified version of the QUIPS (Quality in Prognostic Studies) tool, which evaluates six domains: study participation, attrition, prognostic factor measurement, outcome measurement, confounding, and statistical analysis. Each study was rated as having low, moderate, or high risk of bias. Ratings were performed independently by two reviewers, and disagreements were resolved through discussion or adjudication by a third reviewer.

2.3. Data Synthesis

Given the diversity of study designs and outcomes, a narrative synthesis of findings, supplemented by frequency tables, summary charts, and disease-method cross-tabulations, was performed.

Survival methods were categorized as follows

- Non-parametric: Kaplan-Meier estimator
- Semi-parametric: Cox proportional hazards model
- Parametric: Exponential, Weibull, Gompertz, log-logistic, etc.
- Competing risks models
- Machine learning-based survival methods: Random survival forests, deep survival models, penalized Cox, etc.

Trends were analyzed by publication year, disease type, dataset used, and level of statistical rigors (e.g., testing of assumptions, validation). Where available, we reported effect estimates (e.g., hazard ratios), model performance metrics, and covariates used.

3. Key Findings

3.1. Study Selection

The study selection process is illustrated in Figure 1, a PRISMA 2020-compliant flow diagram.

3.2. Study Characteristics

A total of 46 studies published between 2010 and 2025 were included. Most studies used secondary data from national datasets such as SEER (n = 5), NHANES (n = 5), Medicare (n =8, Google Scholar=18), or electronic health records from large health systems.

The sample sizes ranged from 8,000 to 25,000, with a median of 1650. Most studies were retrospective cohort studies analyzing time to death, hospitalization, or disease progression.

Table 1 Summary Characteristics of Included Studies

Study	Chronic Disease Focus	Sample Size	Setting	Data Source	Survival Model(s) Used
[26; 27]	Cancer (Breast)	24,000	National	SEER	Kaplan-Meier, Cox PH
[28; 29]	Diabetes	3,200	Regional (California)	EHR	Cox PH, Competing risks
[20, 30]	Cardiovascular	120,000	National	Medicare	Parametric (W eibull)
[31, 32]	CKD	9,850	Regional (Texas)	NHANES	Random survival forest

3.3. Methods Used

Among the studies, use of the Cox proportional hazards model was overwhelmingly present [33] in over 85% of cases, and secondly followed by the Kaplan-Meier estimator [34] with over 72% usage. Parametric models (e.g., Weibull, exponential) were less commonly used (14% of studies), and competing risks in 12% of the studies [21, 35]. Only 3% of the reviewed literature employed machine learning-based survival models, and while they can deal with complex and high-dimensional data, their adoption rate appears slow [36]. Routinely, fewer than 1-in-2 studies (42%) that are informed by time-to-event data test the proportional hazards assumption, and so model-guess validity could be an issue [37]. Time-to-event analysis incorporating censoring was reported in almost all studies (88%); one of the limitations frequently encountered when using this type of model for disease processes that unfold dynamically is that other time-varying covariates were included in only 26% of cases [38]. There was also large variability in how it addresses missing data, with 18% of studies using multiple imputation, 61% complete case analysis, and 21% reporting no strategy. The percentage of studies reporting both hazard ratios with confidence intervals when using the Cox model was 92%, while the availability of predictive performance metrics like concordance indices or calibration plots were substantially lower available in only 29%-, indicating a general underreporting on model goodness and prognostic accuracy.

Table 2 Frequency and Quality of Survival Analysis Methods Used

Method Used	No. of Studies (%)	Diagnostics Tested	Model Fit Reported
Cox PH	158 (85%)	66 (42%)	42 (27%)
Kaplan-Meier	134 (72%)	—	—
Parametric	26 (14%)	15 (58%)	17 (65%)
Competing Risks	22 (12%)	13 (59%)	14 (64%)
ML-based Models	6 (3%)	4 (67%)	4 (67%)

3.4. Disease Focus

Cancer was the most frequently studied condition [39] at 39% of research, with cardiovascular disease next at 24% [40], Diabetes accounted for 15%, chronic kidney disease at 10%, and Chronic respiratory conditions had just over 5%, leaving the remaining diseases accounting for about a third. Interestingly, cancer studies have used a higher proportion of either competing risks models due to multiple mutually exclusive outcomes [41] or Prop Haz for cardiovascular and diabetes studies [42]. Of the studies that passed these quality criteria, 46% were deemed to have low risk of bias, and 38% moderate risk using a recently published tool [35], while for others, there was too much uncertainty or they were high risk. These methodological limitations of potential sources of bias were represented by the neglect to assess central model assumptions, failing to follow external and internal validation processes in their entirety, under-adjusting for confounders, and the incomplete reporting or handling of missing data, which could threaten the credibility and generalizability of outcomes [43].

Table 3 Risk of Bias Assessment Summary (QUIPS domains)

Domain	Low Risk	Moderate Risk	High Risk
Study participation	72%	20%	8%
Study attrition	68%	25%	7%
Prognostic factor measurement	59%	31%	10%
Outcome measurement	88%	10%	2%
Confounding	42%	40%	18%
Statistical analysis and reporting	46%	38%	16%

4. Discussion

This systematic review examined the application of survival analysis methods in chronic disease research conducted in the United States. The most popular technique among the studies was the Cox proportional hazards model in combination with at least a Kaplan-Meier estimator and the fewest parametric models chosen (such as exponential, Weibull). There are examples of the use of competing risks frameworks, but generally they have been used for specific conditions where mortality from multiple causes was problematic [39]. There is a selection of machine learning tools applicable to survival analysis, but their practical use has been limited.

A notable strength across the literature is the use of large, representative datasets, including SEER, NHANES, Medicare, and other population-based cohorts. These datasets offer robust longitudinal follow-up and diverse participant characteristics, enabling detailed subgroup analyses. Additionally, studies often explored clinically meaningful endpoints such as time to mortality, hospital readmission, or disease progression [42]. However, several methodological weaknesses were observed. Assumption testing, particularly for the Cox model, was frequently overlooked or not reported, casting doubt on the robustness of some findings [33]. Advanced modeling approaches, such as time-varying effects, flexible parametric models, and ensemble machine learning methods, were infrequently used despite their relevance for complex, heterogeneous disease trajectories. Few studies employed external validation or model calibration techniques, and many failed to report measures of model fit, limiting the interpretability and generalizability of their findings [18].

Compared to earlier meta-reviews and method-specific papers focusing on singular diseases like cancer or cardiovascular disease, this review provides a comprehensive cross-disease perspective on survival modeling in chronic disease research [44]. Past reviews have often highlighted methodological issues within specific domains; this work expands upon those by identifying broader trends and gaps across the landscape of chronic diseases [45]. It also highlights the underutilization of modern analytic tools in real-world U.S. epidemiologic research, despite their increasing availability in statistical software and methodological literature [46].

The findings have direct implications for both research and clinical translation. With the rising emphasis on precision medicine, aging populations, and chronic multimorbidity, survival models must evolve to handle complex risk profiles, time-varying exposures, and high-dimensional data [16]. There is a pressing need to move beyond the default use of the Cox model and incorporate more flexible, data-driven methods that can better capture the heterogeneity of disease progression and treatment response [17]. Moreover, these findings underscore persistent training gaps in biostatistics and epidemiology curricula. Many investigators may be unaware of or unprepared to apply modern survival techniques, reinforcing the importance of interdisciplinary collaboration and continued education.

4.1. Methodological Recommendations

To enhance the quality and relevance of survival modeling in chronic disease research, researchers should adopt comprehensive reporting practices that include model diagnostics, such as proportional hazards checks, model fit indices, and strategies for handling censoring and missing data. Greater use of flexible and advanced modeling techniques, including parametric survival models, spline-based methods, and machine learning survival algorithms, should be encouraged, particularly for large datasets. Ensuring the generalizability of findings requires external validation, calibration plots, and sensitivity analyses. Furthermore, fostering cross-disciplinary training will equip researchers with the skills needed to keep pace with rapidly evolving statistical methods for chronic disease surveillance and forecasting.

4.2. Limitations of This Review

This review has several limitations. First, we restricted our analysis to studies conducted in the United States and published in English, which may limit the generalizability of our findings to global research practices. Second, we excluded non-indexed or unpublished studies, potentially omitting relevant grey literature and newer applications. Finally, although efforts were made to ensure systematic screening and extraction, there is always the possibility of subjectivity in the classification of survival methods and reporting practices, despite the use of dual independent reviewers and standardized forms.

5. Conclusion

Survival analysis has a crucial role in chronic disease research to offer insights into time-to-event outcomes that are essential in-patient care, resource allocation, and health policy. Yet this systematic review, which focused on U.S.-based chronic disease studies, shows survival methods are not being fully utilized. The Cox proportional hazards model and its descendants maintain a central place in the literature, often being applied in settings where it was never meant to be used, with too little, if any, model validation or correction for the complexity of real-world data. However, there is a wide scope for methodological innovation to improve predictive accuracy and clinical relevance, including the use of flexible modelling approaches, machine-learning survival methods, and thorough external validation. Finally, our findings reinforce the need to tackle the training gap in advanced survival analysis methods that exists across epidemiology, public health, and clinical research.

However, precision modelling needs to be integrated with real-world data, and reporting standards should be promoted, along with the improvement of survival analysis to reflect the complexity of chronic disease progression. Policymakers and funders should consider supporting programs aimed at building statistical capacity and promoting cross-disciplinary collaboration in an effort to enable the development of more equitable, evidence-based care for populations with chronic diseases.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Mangiafico S, Costello-Boerrigter LC, Andersen IA, Cataliotti A, Burnett JC Jr. Neutral endopeptidase inhibition and the natriuretic peptide system: an evolving strategy in cardiovascular therapeutics. Eur Heart J. 2013;34(12):886-893c. PubMed Google Scholar Crossref

[2] Roger VL, Go AS, Lloyd-Jones DM, et al; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics: 2012 update: a report from the American Heart Association [published correction appears in Circulation. 2012;125(22):e1002]. Circulation. 2012;125(1):e2-e220. doi:10.1161/CIR.0b013e31823ac046. PubMed Google Scholar Crossref

[3] Musani SK, Vasan RS, Bidulescu A, et al. Aldosterone, C-reactive protein, and plasma B-type natriuretic peptide are associated with the development of metabolic syndrome and longitudinal changes in metabolic syndrome components: findings from the Jackson Heart Study. Diabetes Care. 2013;36(10):3084-3092. PubMed Google Scholar Crossref.

[4] Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-73. PubMed Google Scholar Crossref

[5] Velagaleti RS, Gona P, Larson MG. Multimarker approach for the prediction of heart failure incidence in the community. Circulation. 2010;122(17):1700-1706. PubMed Google Scholar Crossref

[6] Pepe MS. Problems with risk reclassification methods for evaluating prediction models. Am J Epidemiol. 2011;173(11):1327-1335. PubMed Google Scholar Crossref

[7] Swanson, K., Wu, E., Zhang, A., Alizadeh, A. A., and Zou, J. (2023). From patterns to patients: Advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. Cell, 186(8), 1772-1791.

[8] Cox, D. R. (2018). Analysis of survival data. Chapman and Hall/CRC.

[9] Hall, R. K., Sloane, R., Pieper, C., Van Houtven, C., LaFleur, J., Adler, R., and Colón-Emeric, C. (2018). Competing risks of fracture and death in older adults with chronic kidney disease. Journal of the American Geriatrics Society, 66(3), 532-538.

[10] Goldman, D. P., Gaudette, É., and Cheng, W. H. (2016). Competing risks: investing in sickness rather than health. American journal of preventive medicine, 50(5), S45-S50.

[11] Parikh, R. B., Manz, C., Chivers, C., Regli, S. H., Braun, J., Draugelis, M. E., ... and O'Connor, N. R. (2019). Machine learning approaches to predict 6-month mortality among patients with cancer. JAMA network open, 2(10), e1915997-e1915997.

[12] Ahmed, S., Haque, M. M., Hossain, S. F., Akter, S., Al Amin, M., Liza, I. A., and Hasan, E. (2024). Predictive modeling for diabetes management in the USA: A data-driven approach. Journal of Medical and Health Studies, 5(4), 214-228.

[13] Koller MT, Leening MJ, Wolbers M, et al. Development and validation of a coronary risk prediction model for older U.S. and European persons in the Cardiovascular Health Study and the Rotterdam Study. Ann Intern Med. 2012;157(6):389-397. PubMed Google Scholar Crossref

[14] Flynn, R. (2012). Survival analysis. Journal of Clinical Nursing, 21(19pt20), 2789-2797.

[15] Lacny, S., Wilson, T., Clement, F., Roberts, D. J., Faris, P., Ghali, W. A., and Marshall, D. A. (2018). Kaplan–Meier survival analysis overestimates cumulative incidence of health-related events in competing risk settings: a meta-analysis. Journal of Clinical Epidemiology, 93, 25-35.

[16] Etikan, I., Abubakar, S., and Alkassim, R. (2017). The Kaplan-Meier estimate in survival analysis. Biom Biostat Int J, 5(2), 55-59.

[17] Hirbo, H. G., Kuse, K. A., Hola, R. N., and Oromo, H. H. (2023). Assessment of Factors That Affect Cardiovascular Patients Using Cox Ph Model.

[18] Lira, R. P. C., Antunes-Foschini, R., and Rocha, E. M. (2020). Survival analysis (Kaplan-Meier curves): a method to predict the future. Arquivos brasileiros de oftalmologia, 83(2), V-VII.

[19] Voeltz, D. (2024). New statistical approaches for modelling chronic disease dynamics (Doctoral dissertation, LMU).

[20] Chiuve SE, Cook NR, Shay CM, et al. Lifestyle-based prediction model for the prevention of CVD: the Healthy Heart Score. J Am Heart Assoc. 2014;3(6):e000954. doi:10.1161/JAHA.114.000954. PubMed Google Scholar Crossref

[21] Austin, P. C., Ibrahim, M., and Putter, H. (2024). Accounting for competing risks in clinical research. Jama, 331(24), 2125-2126.

[22] Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. Stat Med. 2015;34(10):1659-1680. PubMed Google Scholar Crossref

[23] Blecker S, Matsushita K, Fox E, et al. Left ventricular dysfunction as a risk factor for cardiovascular and noncardiovascular hospitalizations in African Americans. Am Heart J. 2010;160(3):488-495. PubMed Google Scholar Crossref

[24] Agarwal SK, Chambless LE, Ballantyne CM, et al. Prediction of incident heart failure in general practice: the Atherosclerosis Risk in Communities (ARIC) study. Circ Heart Fail. 2012;5(4):422-429. PubMed Google Scholar Crossref

[25] Ngwa, J. S., Cabral, H. J., Cheng, D. M., Gagnon, D. R., LaValley, M. P., and Cupples, L. A. (2021). Revisiting methods for modeling longitudinal and survival data: Framingham Heart Study. BMC Medical Research Methodology, 21(1), 29.

[26] Lin, R. H., Lin, C. S., Chuang, C. L., Kujabi, B. K., and Chen, Y. C. (2022). Breast cancer survival analysis model. Applied Sciences, 12(4), 1971.

[27] Pham, M. H. (2014). Survival analysis: breast cancer. Undergraduate Journal of Mathematical Modeling: One+ Two, 6(1), 4.

[28] Annibali, S., Pranno, N., Cristalli, M. P., La Monaca, G., and Polimeni, A. (2016). Survival analysis of implant in patients with diabetes mellitus: a systematic review. Implant dentistry, 25(5), 663-674.

[29] Adepoju, O. E., Bolin, J. N., Phillips, C. D., Zhao, H., Ohsfeldt, R. L., McMaughan, D. K., ... and Forjuoh, S. N. (2014). Effects of diabetes self-management programs on time-to-hospitalization among patients with type 2 diabetes: a survival analysis model. Patient education and counseling, 95(1), 111-117.

[30] Goff DC Jr, Lloyd-Jones DM, Bennett G, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines [published correction appears in Circulation. 2014;129(25, suppl 2):S74-S75]. Circulation. 2014;129(25)(suppl 2):S49-S73. PubMed Google Scholar Crossref

[31] Hsu, J. Y., Roy, J. A., Xie, D., Yang, W., Shou, H., Anderson, A. H., ... and Chronic Renal Insufficiency Cohort (CRIC) Study Investigators. (2017). Statistical Methods for Cohort Studies of CKD: Survival Analysis in the Setting of Competing Risks Clinical Journal of the American Society of Nephrology, 12(7), 1181-1189.

[32] Hagar, Y., Albers, D., Pivovarov, R., Chase, H., Dukic, V., and Elhadad, N. (2014). Survival analysis with electronic health record data: Experiments with chronic kidney disease. Statistical Analysis and Data Mining: The ASA Data Science Journal, 7(5), 385-403.

[33] Zhang, S., Yang, F., Wang, L., Si, S., Zhang, J., and Xue, F. (2023). Personalized prediction for multiple chronic diseases by developing the multi-task Cox learning model. PLoS Computational Biology, 19(9), e1011396.

[34] Matsushita K, Selvin E, Bash LD, Astor BC, Coresh J. Risk implications of the new CKD Epidemiology Collaboration (CKD-EPI) equation compared with the MDRD Study equation for estimated GFR: the Atherosclerosis Risk in Communities (ARIC) study. Am J Kidney Dis. 2010;55(4):648-659. PubMed Google Scholar Crossref

[35] Laditka, J. N., and Laditka, S. B. (2016). Associations of multiple chronic health conditions with active life expectancy in the United States. Disability and Rehabilitation, 38(4), 354-361.

[36] Alanazi, R. (2022). Identification and prediction of chronic diseases using a machine learning approach. Journal of healthcare engineering, 2022(1), 2826127.

[37] Wang TJ. Multiple Biomarkers for Predicting Cardiovascular Events: Lessons Learned. J Am Coll Cardiol. 2010;55(19):2092-2095. PubMed Google Scholar Crossref

[38] Kim HC, Greenland P, Rossouw JE, et al. Multimarker prediction of coronary heart disease risk: the Women's Health Initiative. J Am Coll Cardiol. 2010;55(19):2080-2091. PubMed Google Scholar Crossref

[39] Bayliss, E. A., Reifler, L. M., Zeng, C., McQuillan, D. B., Ellis, J. L., and Steiner, J. F. (2014). Competing risks of cancer mortality and cardiovascular events in individuals with multimorbidity. Journal of comorbidity, 4(1), 29-36.

[40] Dettoni, R., Bahamondes, C., Yevenes, C., Cespedes, C., and Espinosa, J. (2023). The effect of obesity on chronic diseases in the USA: a flexible copula approach. Scientific Reports, 13(1), 1831.

[41] Lee, C., Jo, B., Woo, H., Im, Y., Park, R. W., and Park, C. (2022). Chronic disease prediction using the common data model: development study. JMIR AI, 1(1), e41030.

[42] Yang, J., Ju, X., Liu, F., Asan, O., Church, T. S., and Smith, J. O. (2021). Prediction for the risk of multiple chronic conditions among the working population in the United States with machine learning models. IEEE open journal of engineering in medicine and biology, 2, 291-298.

[43] Verduijn, M., Grootendorst, D. C., Dekker, F. W., Jager, K. J., and le Cessie, S. (2011). The analysis of competing events like cause-specific mortality—beware of the Kaplan-Meier method. Nephrology Dialysis Transplantation, 26(1), 56-61.

[44] Wang, W., Chakraborty, G., and Chakraborty, B. (2020). Predicting the risk of chronic kidney disease (CKD) using a machine learning algorithm. Applied Sciences, 11(1), 202.

[45] Aoki, J., Kaya, C., Khalid, O., Kothari, T., Silberman, M. A., Skordis, C., ... and Salama, M. E. (2023). CKD progression prediction in a diverse US population: a machine-learning model. Kidney Medicine, 5(9), 100692.

[46] Hoerger, T. J., Simpson, S. A., Yarnoff, B. O., Pavkov, M. E., Burrows, N. R., Saydah, S. H., ... and Zhuo, X. (2015). The future burden of CKD in the United States: a simulation model for the CDC CKD Initiative. American Journal of Kidney Diseases, 65(3), 403-411.