Have Cox-Proportional Hazard models been used to evaluate healthcare costs?

Yes, eight studies have applied Cox-proportional hazards models to evaluate healthcare costs across diverse clinical settings including diabetes, myocardial infarction, and cardiac surgery.

Abstract

Eight studies have applied Cox-proportional hazards models to evaluate healthcare costs in diverse clinical settings. In diabetes care, one study reported that Cox models, when compared with Weibull and Aalen additive approaches, yielded hospitalisation cost estimates that differed by as much as 20%; another noted that Cox models overestimated high-risk and underestimated low-risk costs by roughly 20%. In myocardial infarction and cardiac surgery settings, Cox models produced lower mean cost estimates and in one analysis were the most accurate predictors of mean, median, and high costs based on mean squared and absolute errors. A study in geriatric care employed a Coxian phase-type distribution for scenario testing, while a methodological investigation proposed a novel Cox approach with temporal covariates that delivered consistent, asymptotically normal estimators. Finally, a cost-effectiveness review highlighted that the proportional hazards assumption was seldom checked in such analyses.

Key points:

1. Multiple studies confirm that Cox-proportional hazards models have been used for cost estimation across conditions such as diabetes, myocardial infarction, and cardiac surgery.
2. Model choice can yield differences in estimated costs of up to 20%, and performance varies by context.
3. Alternative approaches (e.g., Aalen additive regression and novel Cox models) sometimes perform better, particularly when standard model assumptions are challenged.

These findings document the application of Cox-proportional hazards models in the analysis of healthcare costs without asserting universal superiority.

Paper search

We performed a semantic search using the query ”Have Cox-Proportional Hazard models been used to evaluate healthcare costs?” across over 138 million academic papers from the Elicit search engine, which includes all of Semantic Scholar and OpenAlex.

We retrieved the 50 papers most relevant to the query.

Screening

We screened in sources that met these criteria:

- **Cox Model Usage**: Does this study explicitly use Cox-Proportional Hazard models or Cox regression as a primary or secondary analytical method?
- **Healthcare Cost Evaluation**: Does this study evaluate, analyze, or model healthcare costs, medical expenses, or healthcare resource utilization as an outcome variable?
- **Healthcare Setting and Population**: Is this study conducted in a healthcare setting (hospitals, clinics, community care, long-term care, etc.) and involves human participants?
- **Appropriate Study Design**: Is this study an empirical study (randomized controlled trial, cohort study, case-control study, cross-sectional study, observational study) or a systematic review/meta-analysis?

- **Combined Methodology and Outcome**: Does this study use BOTH Cox-Proportional Hazard models AND evaluate healthcare costs (not just one of these components)?
- **Empirical Application**: Does this study include empirical application to healthcare cost data (i.e., is it NOT purely theoretical or methodological without practical application)?
- **Adequate Methodological Detail**: Does this study provide sufficient methodological detail (i.e., is it NOT a conference abstract, editorial, commentary, or letter without adequate methodological information)?

We considered all screening questions together and made a holistic judgement about whether to screen in each paper.

Data extraction

We asked a large language model to extract each data column below from each paper. We gave the model the extraction instructions shown below for each column.

- **Cox Model Application**:

  Extract details about how Cox-Proportional Hazard models were used for cost analysis including:

- Whether Cox-PH was actually applied to cost data (yes/no)
- Specific methodology for applying Cox-PH to costs (e.g., treating costs as time-to-event, using costs as covariates, novel adaptations)
- Any modifications or adaptations made to standard Cox-PH approach
- Technical details about implementation (e.g., handling of cost distributions, censoring approach)

- **Healthcare Cost Type**:

  Identify the specific type of healthcare costs being analyzed:

- Direct medical costs, indirect costs, or total costs
- Specific cost categories (hospitalization, surgery, treatment, etc.)
- Disease/condition context (e.g., CABG surgery, diabetes, myocardial infarction)
- Time frame for cost measurement (e.g., 1-year costs, lifetime costs, episode-based)

- **Comparison Methods**:

  Extract all statistical methods compared against Cox-PH for cost analysis including:

- Other regression models tested (linear, GLM, etc.)
- Specific distributions used (Weibull, gamma, Poisson, etc.)
- Non-parametric approaches (e.g., median regression, Aalen additive)
- Whether Cox-PH was the primary focus or just one of many methods compared

- **Performance Results**:

  Extract findings about Cox-PH performance for cost analysis including:

- Specific performance metrics used (MSE, MAE, AIC, etc.)
- How Cox-PH ranked compared to other methods
- Quantitative results (effect sizes, prediction accuracy, etc.)
- Statistical significance of cost associations found
- Model fit diagnostics or validation results

- **Cox Model Advantages**:

  Extract any advantages or strengths of using Cox-PH for cost analysis mentioned by authors including:

- Specific benefits identified (flexibility, robustness, interpretability, etc.)
- Situations where Cox-PH performed better than alternatives
- Theoretical or practical advantages discussed
- Unique capabilities for cost analysis

- **Cox Model Limitations**:

  Extract any limitations, challenges, or disadvantages of using Cox-PH for cost analysis including:

- Specific problems or weaknesses identified
- Assumptions that were violated or problematic
- Situations where other methods performed better
- Technical challenges in implementation
- Interpretability issues with coefficients

- **Study Population**:

  Extract key characteristics of the study population relevant to cost analysis:

- Sample size and data source
- Patient population/condition studied
- Healthcare setting (hospital, outpatient, etc.)
- Geographic location and healthcare system context
- Time period of data collection

- **Main Conclusions**:

  Extract the authors' main conclusions specifically about using Cox-PH models for healthcare cost analysis including:

- Overall assessment of Cox-PH for cost analysis
- Recommendations for when to use Cox-PH vs alternatives
- Implications for future cost analysis methodology
- Any calls for further research or methodological development

Results

Characteristics of Included Studies

Study	Healthcare Domain	Patient Population	Study Design	Primary Objective	Full text retrieved
Gregori et al., 2006	Diabetes	3892 diabetic patients (hospitalised)	Retrospective observational	Evaluate impact of model choice on hospitalisation cost estimates in diabetes care	No
Gregori et al., 2008	Myocardial infarction	487 post-acute myocardial infarction patients (uncomplicated)	Secondary analysis of trial data	Assess how model choice affects 1-year post-infarction cost estimates	No
Dudley et al., 1993	Cardiac surgery	155 coronary artery bypass graft surgery patients	Comparative analytic study	Compare analytic models for estimating effect of clinical factors on coronary artery bypass graft surgery cost	No
Austin et al., 2003	Cardiac surgery	1959 coronary artery bypass graft surgery patients (Calgary, Alberta)	Retrospective cohort	Compare regression models for analysing cost of coronary artery bypass graft surgery	No
Zigon et al., 2005	Diabetes	2550 type 2 diabetic patients (hospitalised)	Retrospective observational	Compare survival and regression models for hospitalisation cost estimation in diabetes	Yes
Marshall et al., 2004	Geriatric care	1392 geriatric inpatients (Northern Ireland)	Retrospective cohort	Estimate costs for geriatric patients using Coxian phase-type distribution	No

Study	Healthcare Domain	Patient Population	Study Design	Primary Objective	Full text retrieved
Zheng et al., 2018	Methodological/General	Not specified (hospital discharge datasets)	Methodological/statistical	Propose novel Cox model approach for cost data with temporal covariates	No
Guyot et al., 2011	Cost-effectiveness review	Not specified (National Institute for Health and Care Excellence cost-effectiveness analysis submissions)	Review of cost-effectiveness analyses	Examine model choice for survival and cost-effectiveness analysis in randomized controlled trials and cost-effectiveness analyses	No

Distribution of healthcare domains among included studies:

- Diabetes: 2 studies
- Cardiac surgery: 2 studies
- Myocardial infarction: 1 study
- Geriatric care: 1 study
- Methodological/general (not disease-specific): 1 study
- Cost-effectiveness review (not disease-specific): 1 study

Study design breakdown:

- Retrospective observational: 2 studies
- Retrospective cohort: 2 studies
- Comparative analytic study: 1 study
- Secondary analysis of trial data: 1 study
- Methodological/statistical study: 1 study
- Review of cost-effectiveness analyses: 1 study

Primary objectives:

- Comparison of different models for cost estimation in specific clinical contexts: 5 studies
- Application of a specific model (Coxian phase-type distribution) to cost estimation in geriatric care: 1 study
- Proposal of a novel modeling approach for cost data with temporal covariates: 1 study
- Review of model choice for survival and cost-effectiveness analysis in randomized controlled trials and cost-effectiveness analyses: 1 study

Among the included studies, we didn't find studies outside these categories.

Effects

Applications of Cox Proportional Hazards Models in Healthcare Cost Analysis

Study	Statistical Method(s) Compared	Cost Prediction Accuracy/Findings	Model Performance Metrics Used	Study Context
Gregori et al., 2006	Weibull parametric survival, Cox proportional hazards, Aalen additive regression	Agreement on covariate effects; model choice can cause up to 20% difference in cost estimates	No mention found	Hospitalisation costs in diabetes
Gregori et al., 2008	Weibull parametric survival, Cox proportional hazards, Aalen additive regression	Cox and Aalen models provide lower mean cost estimates; agreement on covariate effects	No mention found	1-year post-myocardial infarction costs
Dudley et al., 1993	Ordinary least squares (with/without transformation), logistic regression, Weibull, Cox proportional hazards	Cox proportional hazards most accurate for mean, median, and high-cost prediction; ejection fraction and age significant predictors	No mention found	Coronary artery bypass graft surgery costs
Austin et al., 2003	Linear regression, log-linear, generalized linear model (Poisson, negative binomial, gamma), median regression, Cox proportional hazards	Cox proportional hazards among best for cost prediction (mean squared error, mean absolute error); all models consistent for factor identification	Mean squared error, mean absolute error	Coronary artery bypass graft surgery costs
Zigon et al., 2005	Ordinary least squares, logistic regression, Weibull, Cox proportional hazards, Aalen additive regression	Cox proportional hazards overestimates high-risk, underestimates low-risk; Aalen model best for mean cost estimation	No mention found; 20% overestimation	Hospitalisation costs in diabetes (4.5 years)

Study	Statistical Method(s) Compared	Cost Prediction Accuracy/Findings	Model Performance Metrics Used	Study Context
Marshall et al., 2004	Coxian phase-type distribution (examined previous methods, no mention found)	No specific performance metrics reported; model presented as beneficial for scenario testing	No mention found	Geriatric inpatient costs
Zheng et al., 2018	Novel Cox model with temporal covariates (no direct comparison)	Estimators consistent/asymptotically normal; approach useful for limited data scenarios	No mention found	Hospital discharge cost data
Guyot et al., 2011	Cox regression (efficacy), parametric models (cost-effectiveness analysis)	Discrepancy in model use for efficacy vs. cost-effectiveness analysis; proportional hazards assumption seldom checked	No mention found	Review of cost-effectiveness analyses with survival outcomes

Summary of statistical methods compared for cost prediction:

- Cox proportional hazards was evaluated in all 8 studies, either as a primary method or in comparison with others.
- Weibull parametric survival models: 4 studies
- Aalen additive regression: 3 studies
- Ordinary least squares or linear regression: 2 studies
- Logistic regression: 2 studies
- Generalized linear model (Poisson, negative binomial, gamma), median regression, Coxian phase-type, and novel Cox models: each in 1 study
- Parametric models for cost-effectiveness analysis: 1 study

Model performance metrics:

- Only 1 study (Austin et al., 2003) reported explicit performance metrics (mean squared error, mean absolute error).
- One study (Zigon et al., 2005) reported a 20% overestimation as a finding.
- In 7 studies, we didn't find mention of explicit model performance metrics.

Findings on cost prediction accuracy:

- Two studies reported agreement on covariate effects across models.
- Two studies found that model choice can cause substantial differences in cost estimates (up to 20%).
- Two studies found Cox proportional hazards to be the most accurate or among the best for cost prediction.
- One study found the Aalen model best for mean cost estimation.

- One study found all models consistent for factor identification.
- One study found Cox proportional hazards overestimates high-risk and underestimates low-risk costs.
- Two studies presented novel or alternative approaches without direct comparison.
- One study found a discrepancy in model use for efficacy versus cost-effectiveness analysis and noted that the proportional hazards assumption was seldom checked.

We didn't find mention of explicit model performance metrics in 7 of the 8 studies; only 1 study reported mean squared error and mean absolute error.

Comparative Performance Against Alternative Methods

Study	Cox Proportional Hazards Performance	Best Performing Method(s)	Key Findings
Gregori et al., 2006	Agreement on covariate effects; possible 20% over/underestimation	Aalen additive regression	Model choice crucial; Aalen model more flexible and less biased for cost estimation
Gregori et al., 2008	Lower mean cost estimates; agreement on covariate effects	Aalen additive regression	Model choice affects interpretation; Aalen model more flexible
Dudley et al., 1993	Most accurate for mean, median, and high-cost prediction	Cox proportional hazards	Cox proportional hazards shows promise for analyzing clinical predictors of cost
Austin et al., 2003	Among best for cost prediction (mean squared error, mean absolute error); consistent factor identification	Cox proportional hazards, generalized linear models, log-linear, median regression	All models consistent for factor identification; model choice should be data-driven
Zigon et al., 2005	Overestimates high-risk, underestimates low-risk; 20% overestimation overall	Aalen additive regression	Cox proportional hazards limited by non-proportional hazards; Aalen model preferred
Marshall et al., 2004	No mention found	Coxian phase-type distribution	Model beneficial for scenario testing; no direct comparison reported
Zheng et al., 2018	Estimators consistent/asymptotically normal; suited for limited data	Novel Cox model (proposed)	Approach useful for datasets with only event-time and covariate at event-time

Study	Cox Proportional Hazards Performance	Best Performing Method(s)	Key Findings
Guyot et al., 2011	Not directly used for cost-effectiveness analysis; proportional hazards assumption seldom checked	No mention found	Discrepancy in model use for efficacy vs. cost-effectiveness analysis; need for consistent model selection

Key findings across studies:

- Three studies (Gregori et al., 2006; Gregori et al., 2008; Zigon et al., 2005) reported Aalen additive regression as the best or preferred method, citing its flexibility and reduced bias, especially when proportional hazards assumptions were violated.
- Two studies (Dudley et al., 1993; Austin et al., 2003) found Cox proportional hazards to be the best or among the best for cost prediction.
- One study (Austin et al., 2003) also identified generalized linear models, log-linear, and median regression as among the best.
- One study (Marshall et al., 2004) found the Coxian phase-type model beneficial for scenario testing, but we didn't find mention of a direct comparison to other models.
- One study (Zheng et al., 2018) found their novel Cox approach useful for datasets with only event-time and covariate at event-time.
- One study (Guyot et al., 2011) noted a need for consistent model selection between efficacy and cost-effectiveness analysis contexts.

Additional insights:

- Three studies emphasized that model choice is crucial and should be data-driven.
- One study highlighted limitations of Cox proportional hazards under non-proportional hazards.
- One study recommended Cox proportional hazards for analyzing clinical predictors of cost.
- One study found all models consistent for factor identification and suggested tailoring model choice to the data.
- One study noted the need for consistent model selection between efficacy and cost-effectiveness analyses.

We didn't find direct head-to-head comparisons of all models in every study, and in some cases, model performance or selection criteria were not fully described in the available abstracts or full texts.

Model Selection and Validation Outcomes

Study	Cox Proportional Hazards Performance Summary	Best Performing Method(s)	Key Findings
Gregori et al., 2006	Potential for over/underestimation; less flexible than Aalen model	Aalen additive regression	Model choice can affect cost estimates by up to 20%

Study	Cox Proportional Hazards Performance Summary	Best Performing Method(s)	Key Findings
Gregori et al., 2008	Lower mean cost estimates; not as flexible as Aalen model	Aalen additive regression	Importance of model choice for accurate cost determinant interpretation
Dudley et al., 1993	Most accurate for cost prediction	Cox proportional hazards	Cox proportional hazards recommended for analyzing clinical predictors of cost
Austin et al., 2003	Among best for cost prediction; consistent with other models	Cox proportional hazards, generalized linear models, log-linear, median regression	All models consistent for factor identification; model choice should be tailored to data
Zigon et al., 2005	Overestimates mean cost by ~20%; assumption violations	Aalen additive regression	Cox proportional hazards limited by non-proportional hazards; Aalen model preferred
Marshall et al., 2004	No mention found	Coxian phase-type distribution	Model useful for scenario testing; no direct comparison reported
Zheng et al., 2018	Consistent/asymptotically normal estimators; suited for limited data	Novel Cox model (proposed)	Approach useful for datasets with only event-time and covariate at event-time
Guyot et al., 2011	Not directly used for cost-effectiveness analysis; assumption checking lacking	No mention found	Discrepancy in model use for efficacy vs. cost-effectiveness analysis; need for consistent model selection

Summary across the 8 studies:

- Three studies found Cox proportional hazards less flexible or prone to over/underestimation compared to the Aalen additive model.
- One study found Cox proportional hazards overestimated mean cost by approximately 20% and was limited by assumption violations.
- Two studies found Cox proportional hazards to be among the most accurate or consistent for cost prediction.
- One study reported Cox proportional hazards estimators as consistent and asymptotically normal, and suitable for limited data.
- One study did not provide mention of Cox proportional hazards performance.

Best performing methods identified:

- Aalen additive regression: 3 studies
- Cox proportional hazards: 2 studies

- Generalized linear models, log-linear, and median regression: each identified once (all in the same study)
- Coxian phase-type distribution and a novel Cox model: each identified once
- In one study, we didn't find mention of a best performing method

Key findings:

- Two studies reported that model choice can affect cost estimates by up to 20%.
- One study emphasized the importance of model choice for interpreting cost determinants.
- One study recommended Cox proportional hazards for analyzing clinical predictors of cost.
- One study found all models consistent for factor identification and suggested tailoring model choice to the data.
- One study noted Cox proportional hazards was limited by non-proportional hazards and preferred the Aalen model.
- One study highlighted the utility of a model for scenario testing but did not provide direct comparisons.
- One study noted the need for consistent model selection between efficacy and cost-effectiveness analyses.

Based on available abstracts or full texts, we didn't find direct quantitative comparisons of model performance in 2 studies, and in 1 study, the best performing method was not specified.

References

A. Marshall, B. Shaw, and S. McClean. "Estimating the Costs for a Group of Geriatric Patients Using the Coxian Phase-type Distribution." *Statistics in Medicine*, 2004.

D. Gregori, A. Desideri, R. Bigi, M. Petrinco, L. Cortigiani, G. Zigon, and E. Pagano. "Proper Modeling Strategies Selection for the Assessment of Post-Infarction Costs." *International Journal of Cardiology*, 2008.

Dario Gregori, Eva Pagano, Rosalba Rosato, S. Bo, Giulia Zigon, and Franco Merletti. "Evaluating Hospital Costs in Type 2 Diabetes Care: Does the Choice of the Model Matter?*." *Current Medical Research and Opinion*, 2006.

G. Zigon, R. Rosato, S. Bo, and D. Gregori. "Comparison of Analytic Models for the Costs of the Hospitalized Diabetic Patients." *Advances in Methodology and Statistics*, 2005.

P. Austin, W. Ghali, and J. Tu. "A Comparison of Several Regression Models for Analysing Cost of CABG Surgery." *Statistics in Medicine*, 2003.

P. Guyot, N. Welton, M. Ouwens, and A. Ades. "Survival Time Outcomes in Randomized, Controlled Trials and Meta-Analyses: The Parallel Universes of Efficacy and Cost-Effectiveness." *Value in Health*, 2011.

R. Dudley, F. Harrell, L. Smith, D. Mark, R. Califf, D. Pryor, D. Glower, Joseph Lipscomb, and M. Hlatky. "Comparison of Analytic Models for Estimating the Effect of Clinical Factors on the Cost of Coronary Artery Bypass Graft Surgery." *Journal of Clinical Epidemiology*, 1993.

Yanqiao Zheng, X. Zhao, and Xiaoqi Zhang. "A Novel Approach to Estimate the Cox Model with Temporal Covariates and Application to Medical Cost Data." *Communications in Statistics - Theory and Methods*, 2018.