Comparing Deep Learning and Traditional Models for Predicting Survival in DLBCL: DeepSurv vs CoxPH

A. KASTURE (MBBS)[1], A. KAKADE (BTech, MS)[2] and A. BAHETI (MBBS, MD, DM, DNB)[3]

1. SBH Government Medical College & Hospital Dhule, India
2. University of Southern California, Los Angeles, USA
3. Deenanath Mangeshkar Hospital and Research Center, Pune, India

SBH Government Medical College & Hospital Dhule, India

Introduction

Artificial Intelligence, particularly machine learning and deep learning model like DeepSurv, is transforming healthcare by:

- Improving Survival Analysis - enhance the accuracy of survival predictions compared to traditional statistical methods.
- Supporting Clinical Decisions - aid in personalized treatment recommendations and risk stratification, helping clinicians optimize interventions.
- Handling Complex Data - improving outcomes in cancer research and other fields.

Aim

Accurate risk classification in DLBCL is critical for therapy, but CoxPH while being interpretable assumes linear covariate effects and proportional hazards. DeepSurv, an adaptation of CoxPH, can capture non-linear interactions without that assumption. Demonstrating a significant performance gain which could justify integrating deep survival models for the purpose of clinical prognostics

Objectives

To evaluate whether a deep learning based survival model (DeepSurv) can outperform the classical Cox proportional-hazards (CoxPH) model in predicting overall survival in diffuse large B-cell lymphoma (DLBCL) using routinely collected clinical and demographic variables.

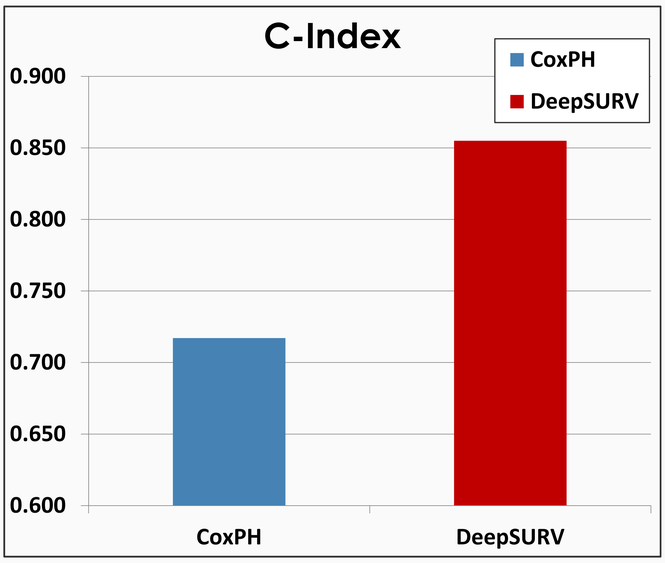
Methods

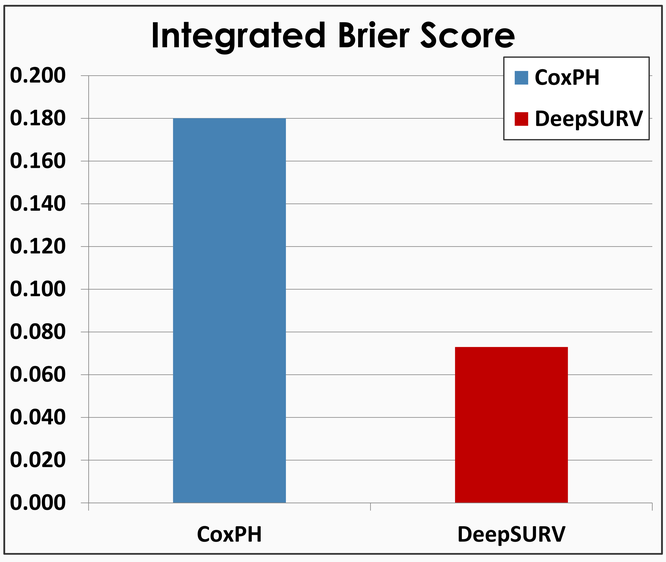
A retrospective analysis was done on 26,084 DLBCL cases from the SEER database (2000-2022). Raw data was standardised and cleaned, with missing values imputed and categorical variables one-hot encoded. Additional features were engineered to capture complex patterns in the data. For e.g., polynomial terms were included - such as 'age squared', interaction variables – such as 'age multiplied by survival time', and missingness indicators to flag absent values. Both Cox proportional hazards (CoxPH) and DeepSurv neural network models were trained using this enhanced feature set. All processing steps were applied equally to both models. Model performance was assessed by the concordance index (c-index) and integrated Brier score (IBS).

Results

- DeepSurv outperformed CoxPH in predicting survival outcomes.
- DeepSurv achieved a higher mean C-index (0.855) compared to CoxPH (0.717), indicating better survival discrimination and a lower mean IBS (0.073 vs 0.180), reflecting lower prediction error.

These results demonstrate the predictive accuracy of DeepSurv over CoxPH for complex non linear survival data.





Summary

DeepSurv delivered a clinically and statistically significant improvement in survival discrimination over CoxPH while sharply reducing prediction error. These findings support further evaluation of deep learning based survival models in DLBCL which can give each patient a more precise & personalized survival estimate by capturing complex patterns in the data that simple linear models might overlook.