STRATOS TG8 PAPER 2

REGRESSION TECHNIQUES FOR CENSORED FAILURE TIMES DATA

Michal Abrahamowicz, Richard Cook, Malka Gorfine,

Pierre Joly and Other Members of TG8

November 2, 2021

UNIVERSITY OF WATERLOO

REGRESSION FOR CENSORED FAILURE TIME DATA

PURPOSE

To review and compare approaches for regression analysis involving censored failure time data

RATIONALE

Raising awareness of the *full spectrum of modeling options* available for censored data will:

- *encourage critical thinking* about modeling assumptions and interpretation of covariate effects
- help practitioners select modeling approaches facilitating inferences of interest (e.g. causal)

We will also *review software for model fitting and diagnostics*

INTRODUCTION

In the introduction we will discuss the various goals of regression modeling with censored data which may be

a. to gain scientific insights into failure time processes

b. to study associations

c. causal inference

d. prediction

Much work has been carried out for uncensored data for meeting these objectives

- We focus primarily on settings where emphasis is on interpretation, so the *approximate "validity"* of models is key
- initial focus is on *fixed covariates*
- We comment on causal inference, prediction and time-dependent covariates in latter sections

HAZARD BASED MODELS

$$\lambda(t \mid X_i) = \lambda_0(t)\exp(X_i'\beta)$$ COX (1972)

$$\lambda(t \mid X_i) = \bar{X}_i'\alpha(t) \quad \text{where} \quad \bar{X}_i = (1, X_i')'$$ AALEN (1989)

$$\lambda(t \mid X_i) = \bar{X}_i'\alpha(t)\exp(X_i'\beta)$$ SCHEIKE AND ZHANG (2003)

TRANSFORMATION MODELS

FINE ET AL. (1998)

$$S(t \mid X) = P(T \geq t \mid X) = \exp(-\Lambda(t \mid X)) \quad \text{where} \quad \Lambda(t \mid X) = \int_0^t \lambda(s \mid X)\, ds$$

$$g(S(t \mid X)) = \alpha_0(t) + X'\beta$$

$g(u) = \log(-\log u)$ proportional hazards

$g(u) = \log(u/(1-u))$ proportional odds

Accelerated Failure Time Models

Lin et al. (1998)

$$S(t \mid X) = S_0(t \cdot g(X; \beta))$$

Regression based on Pseudo-values

Andersen and Perme (2010)

Restricted Mean Survival Time Regression

Zucker (1998)

Quantile Regression

Peng (2021)

Let $Q(\tau \mid Z)$ satisfy $P(\log T \leq Q(\tau \mid Z) \mid Z = z)$

Specify

$$Q(\tau \mid Z) = Z'\beta$$

MODEL DIAGNOSTICS

- Assessing whether the right class of model (e.g. PH/AFT/additive) is used
- Checking the functional form of covariate effects within a given class

RESIDUALS

- adaptations of linear regression diagnostics
- hazard-based residuals
- Schoenfeld residuals
- Martingale residuals

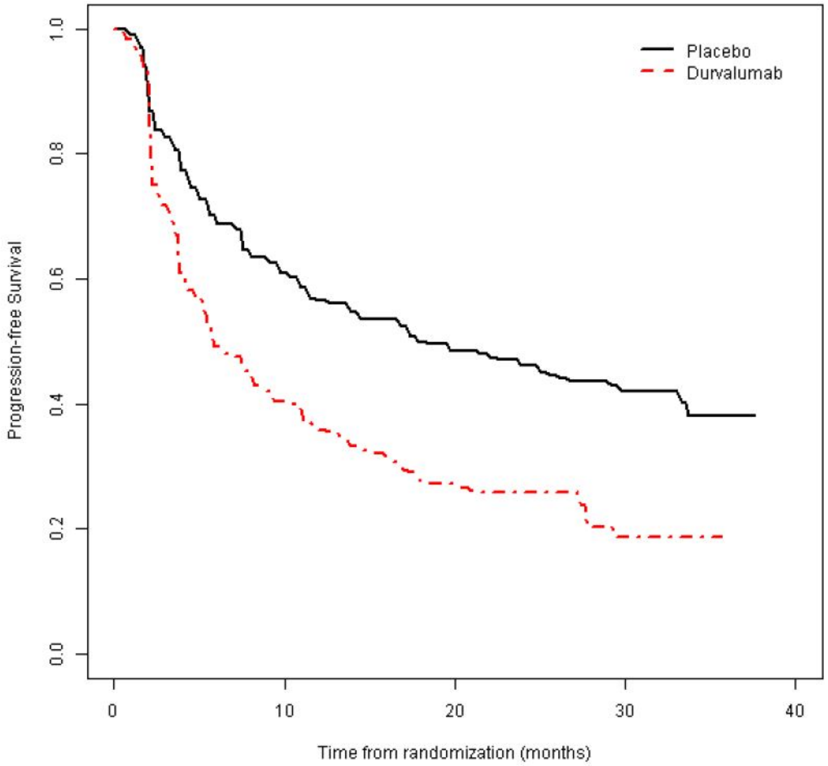
Effect of Durvulumab in Stage II Non-small Cell Lung Cancer

Stage III non-small-cell lung cancer patients were randomized after standard chemotherapy

473 patients assigned to durvalumb

237 patients assigned to placebo

Kaplan-Meier estimates of progression-free survival rates by randomized arm in the overall study population



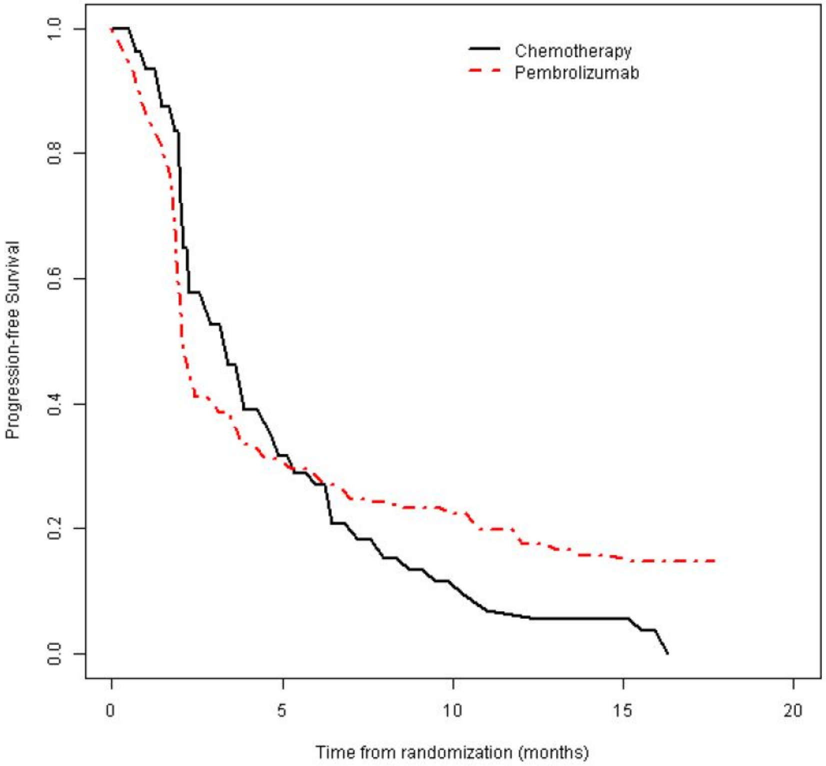
Antonia SJ, Villegas A, Daniel D, et al. (2017). Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *New England Journal of Medicine*, **377**(20): 1919–1929.

Effect of Pembrolizumab in Advanced Urothelial Carcinoma

A total of 542 patients with advanced urothelial cancer were randomized to receive

- pembrolizumab ($n = 270$)
- investigator's choice of chemotherapy ($n = 272$)

Kaplan-Meier estimates for the co-primary endpoint in the overall study population



Bellmunt J, de Wit R, Vaughn DJ, et al. (2017). Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *New England Journal of Medicine*, **376**(11): 1015–1026.

Additional Important Issues

Variable Selection

Delayed entry (discuss in the context of fixed covariates)

Dealing with time-dependent covariates

Prediction

Collapsibility and causal inference

Competing risks

References

Aalen OO (1989). A linear regression model for the analysis of life times. *Statistics in Medicine*, **8**(8), 907–925.

Andersen PK and Pohar Perme M (2010). Pseudo-observations in survival analysis. *Statistical Methods in Medical Research*, **19**(1), 71–99.

Antonia SJ, Villegas A, Daniel D, et al. (2017). Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *New England Journal of Medicine*, **377**(20): 1919–1929.

Bellmunt J, de Wit R, Vaughn DJ, et al. (2017). Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *New England Journal of Medicine*, **376**(11): 1015–1026.

Cheng SC, Wei LJ and Ying Z (1995). Analysis of transformation models with censored data. *Biometrika*, **82**(4), 835–845.

Cox DR (1972). Regression models and lifetables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **34**(2), 187–202.

Fine JP, Ying Z and Wei LG (1998). On the linear transformation model for censored data. *Biometrika*, **85**(4), 980–986.

Lin DY, Wei LJ and Ying Z (1998). Accelerated failure time models for counting processes. *Biometrika*, **85**(3), 605–618.

Peng L (2021). Quantile regression for survival data. *Annual Review of Statistics and Its Application*, **8**, 413–437.

Scheike TH and Zhang MJ (2003). Extensions and applications of the CoxAalen survival model. *Biometrics*, **59**(4), 1036–1045.

Zucker DM (1998). Restricted mean life with covariates: modification and extension of a useful survival analysis method. *Journal of the American Statistical Association*, **93**(442), 702–709.