Model Evaluation

Jane Ma

2024-12-12

Background

Cirrhosis Overview

Related research

Study Objective

Methods

Dataset Description

- Data source, variables, and measures
- Missing data

Survival Analysis

Kaplan-Meier Estimates

Log-Rank Test

Cox PH

After the proportional hazards assumption is met, interaction terms between model covariates are considered using the likelihood ratio test. Suppose there are p covariates in the model. First, each one of the $\frac{p(p-1)}{2}$ interaction terms is added to the model to obtain $\frac{p(p-1)}{2}$ p value, and the interaction term with the lowest p value is added to the model. Then, the rest of the interaction terms are added to the new model one by one to obtain the p value. The process is repeated until $p > 0.05$ holds for all the likelihood ratio tests. The model with the added interaction terms is the final model.

With the final model, model evaluation is conducted. Deviance residuals and Cox-Snell residuals are used to evaluate model fit. For models with good fit, deviance residuals should be randomly distributed around 0, and for KM survival estimates using Cox-Snell residuals as the pseudo survival time, $log(-log(S(t)))$ should be approximately equal to $log(t)$. Influence diagnostics with LD option is used to identify influential individuals. It evaluates how much the log-likelihood would change if the i^{th} person was removed from the sample. After identifying outliers and influential individuals, these subjects are removed from the sample and the model is re-fit to see if the results would change.

Results

Descriptive Statistics (EDA)

KM and Log-Rank Test

Cox Model

Now, interaction terms are considered between the covariates in the model. During the first iteration, five interaction terms, including interaction for Copper with age, Albumin, SGOT, Prothrombin, and diease stage are selected using criteria $p < 0.05$, and the interaction between Albumin and Copper is added to the model as it has the lowest p value. During the second iteration, none of the interaction terms gets selected, and the iteration ends. The final model can then be specified as:

$$\begin{aligned} log(\text{HR}) =& \beta_1 I(\text{drug=D-penicillamine}) + \beta_2 age + \beta_3 \text{billirubin} + \beta_4 \text{albumin} + \beta_5 \text{copper} \\ &+ \beta_6 \text{sgot} + \beta_7 \text{prothrombin} + \beta_8 I(\text{stage=2}) + \beta_9 I(\text{stage=3}) + \beta_{10} I(\text{stage=4}) \\ &+ \beta_{11} \text{bilirubin : n\_days} + \beta_{12} \text{albumin : copper} \end{aligned}$$

Table 1 shows the model estimates. It can be concluded that, with $p < 0.05$:

- The primary variable of interest, drug, has a negative yet insignificant impact on survival. Other covariates that impose a negative significant impact include age, Bilirubin, SGOT, Prothrombin, Stage 4 (compared with Stage 1), and interaction between Albumin and Copper. Albumin, Copper, and interaction between Bilirubin and number of days instead have a protective significant impact.
- The hazard for PBC patients treated with D-penicillamine is 1.2720 times that of PBC patients treated with Placebo, holding other covariates constant.
- The hazard ratio for PBC patients with one year increase in age is 1.0343, holding other covariates constant.
- The hazard ratio for PBC patients with 1 mg/dl increase in Bilirubin is 1.2798, holding other covariates constant.
- The hazard ratio for PBC patients with 1 gm/dl increase in Albumin is 0.2316, holding other covariates constant.
- The hazard ratio for PBC patients with 1 ug/day increase in Copper is 0.9779, holding other covariates constant.
- The hazard ratio for PBC patients with 1U/ml in SGOT is 1.0065, holding other covariates constant.
- The hazard ratio for PBC patients with 1s increase in Prothrombin is 1.3257, holding other covariates constant.
- The hazard for PBC patients at Stage 2, 3, and, 4 is 3.7034, 5.4604, and 8.0139 times that of PBC patients at Stage 1, holding other covariates constant.
- For the same level of Bilirubin, a unit increase in survive time results in -0.0002Bilirubin change in the effect of Bilirubin on log hazard ratio for PBC patients, holding other covariates constant.
- For the same level of Albumin, a unit increase in Copper results in 0.0076Albumin change in the effect of Albumin on log hazard ratio for PBC patients, holding other covariates constant.

Model evaluation is then conducted on this final model. Figures 1 and 2 show the deviance residuals distribution and the KM estimates using Cox-Snell residuals as pseudo survival time respectively. As there is no obvious trend in the deviance plots and the line is close to the reference line in Cox-Snell plot, it can be concluded that the model is a relatively good fit.

For influence diagnostics, the individuals that provide the 5 largest absolute differences for the LD option are selected. After removing the outliers (identified by a deviance residual larger than 3, 2 are selected) and the 5 influential individuals, the model is re-fit. Table 2 compares the model estimates for the two models.

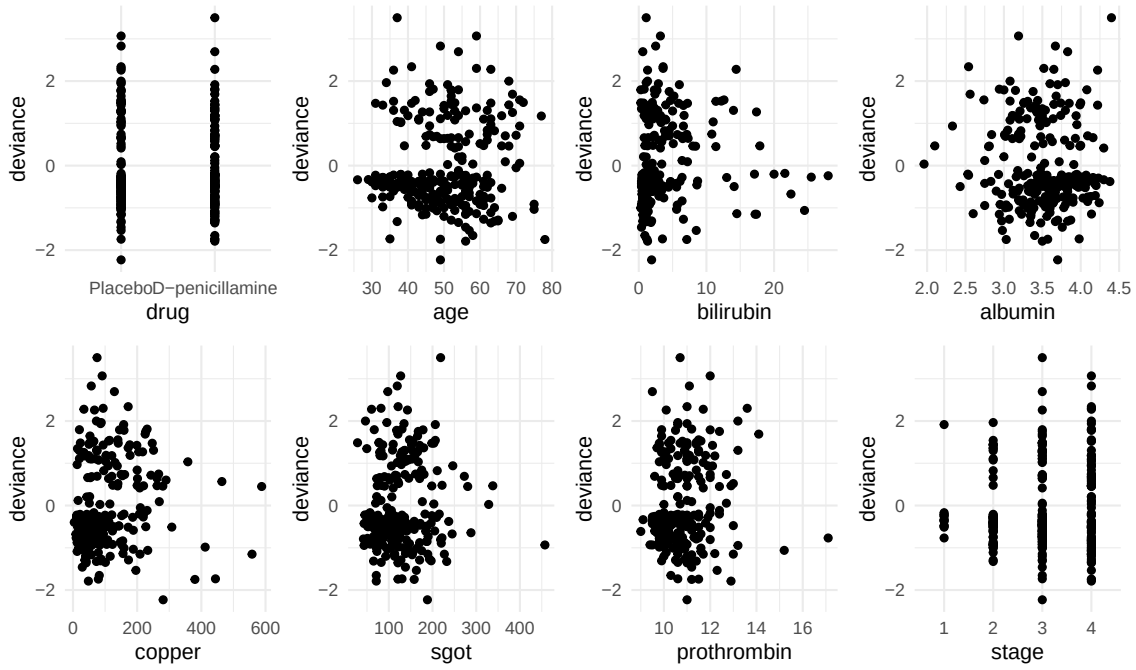

Figure 1: Deviance Residuals Scatterplot for Individual Variable

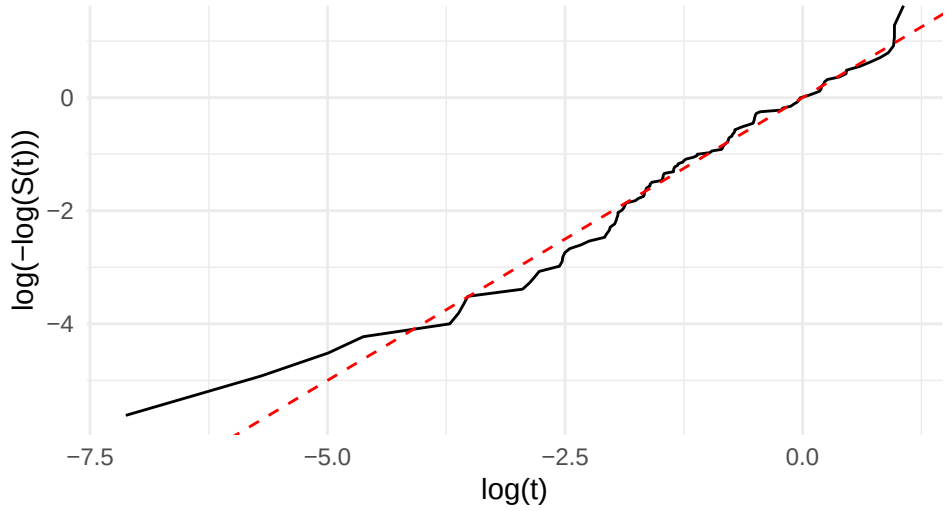

Figure 2: KM Estimates Using Cox-Snell Residuals

Table 1: Final Model Hazard Ratio Estimates

Characteristic	**HR**[1]	**95% CI**[1]	**p-value**
drug			
drug.L	1.2720	0.9451, 1.7118	0.1124
age	1.0343	1.0119, 1.0572	0.0025
bilirubin	1.2798	1.1927, 1.3731	0.0000
albumin	0.2316	0.1089, 0.4927	0.0001
copper	0.9779	0.9641, 0.9919	0.0020
sgot	1.0065	1.0026, 1.0104	0.0010
prothrombin	1.3257	1.0521, 1.6704	0.0168
stage			
1	—	—	
2	3.7034	0.4513, 30.392	0.2228
3	5.4604	0.6981, 42.713	0.1058
4	8.0139	1.0305, 62.324	0.0467
bilirubin * n_days	0.9998	0.9997, 0.9999	0.0000
albumin * copper	1.0076	1.0034, 1.0119	0.0004

[1]HR = Hazard Ratio, CI = Confidence Interval

There are subtle differences between model estimates, but the direction of impact stays the same for all the variables, thus resulting in similar conclusions.

Table 2: Model Parameter Estimates Comparison

	Original Model			New Model		
	Estimate	Hazard Ratio	p value	Estimate	Hazard Ratio	p value
drug.L	0.2406	1.2720	0.1124	0.2907	1.3374	0.0662
age	0.0337	1.0343	0.0025	0.0303	1.0308	0.0098
bilirubin	0.2467	1.2798	0.0000	0.3033	1.3543	0.0000
albumin	-1.4627	0.2316	0.0001	-1.5639	0.2093	0.0001
copper	-0.0224	0.9779	0.0020	-0.0226	0.9777	0.0021
sgot	0.0065	1.0065	0.0010	0.0063	1.0063	0.0024
prothrombin	0.2819	1.3257	0.0168	0.3428	1.4089	0.0063
stage2	1.3093	3.7034	0.2228	1.4232	4.1505	0.1921
stage3	1.6975	5.4604	0.1058	1.7935	6.0104	0.0930
stage4	2.0812	8.0139	0.0467	2.1404	8.5030	0.0440
bilirubin:n_days	-0.0002	0.9998	0.0000	-0.0002	0.9998	0.0000
albumin:copper	0.0076	1.0076	0.0004	0.0078	1.0078	0.0003

Discussion

The analysis demonstrates that D-penicillamine is inefficient in improving survival outcomes for patients with primary biliary cirrhosis (PBC). The hazard ratio for those treated with the drug is 1.2720 compared

to placebo, indicating no survival benefit and a potential negative effect. This finding suggests the need to reconsider its use and focus on alternative therapies that may offer greater efficacy.

Age and disease stage emerged as critical determinants of survival, underscoring the importance of early detection and stage-specific care. The hazard of mortality increases significantly with each advancing stage, with Stage 4 patients facing an eightfold higher risk compared to Stage 1. Early interventions to halt disease progression are vital, as is tailoring treatment strategies to the patient's disease stage.

Liver function indicators such as Bilirubin, Copper, SGOT, Prothrombin, and Albumin are vital in survival outcomes. Elevated Bilirubin, SGOT, and Prothrombin levels are associated with higher hazards, reflecting liver damage and dysfunction. Conversely, higher albumin levels provide a strong protective effect, emphasizing the importance of maintaining good nutritional and synthetic liver function. The modest protective impact of Copper is proved by previous studies where Copper deficiency is identified as a risk factor for mortality in advanced liver disease (Yu et al. 2019). These findings highlight the necessity of monitoring liver function and metabolic health closely to identify high-risk patients and address reversible factors.

The analysis also revealed key interactions affecting survival. A negative Albumin-Copper interaction was observed, indicating a synergistic detrimental effect. This underscores the complexity of metabolic interactions in liver disease and the need to address deficiencies or toxicities in a balanced manner. Furthermore, continued investigation of other potential interactions is needed to uncover personalized treatment strategies.

To improve patient outcomes, it is crucial to monitor high-risk patients routinely, focusing on those with poor liver function indicators or advanced disease stages. Additionally, conducting biological investigations into the protective role of copper and its interactions with other variables could provide valuable insights. These efforts can pave the way for personalized therapies that address the unique risk profiles of individual patients and enhance survival outcomes.

Conclusion

References

Yu, L., I. W. Liou, S. W. Biggins, M. Yeh, F. Jalikis, L. N. Chan, and J. Burkhead. 2019. "Copper Deficiency in Liver Diseases: A Case Series and Pathophysiological Considerations." *Hepatology Communications* 3 (8): 1159–65. https://doi.org/10.1002/hep4.1393.