長庚大學醫學院臨床醫學研究所
畢業生研究成果

畢業年度：111學年度第二學期　　畢業研究生：楊霈莛　　學號：M0900502

現職：畢業生　　指導教授：林志榮

畢業論文題目（中文）：應用不同模型分析經動脈化療栓塞術之肝癌患者存活率與肌少症之相關性研究

畢業論文題目（英文）：Using Different Models to Approach the Association between Survival Rate and Sarcopenia in Patients with Hepatocellular Carcinoma Undergoing Transarterial Chemoembolization

Abstract

Liver cancer is one of the most prevalent cancers in the world. Developing prediction models has always been a topic of interest to experts. Accurately predicting the risk of death is of great benefit to both clinicians and patients, and therefore many studies have been conducted to develop effective risk prediction models. Many published papers showed sarcopenia negatively affects the prognosis of hepatocellular carcinoma, but there were no studies building risk prediction model associated with sarcopenia. Besides, there was no consensus on the criteria of sarcopenia in previous studies. In this study, we aimed to include sarcopenia in the prediction model and apply a mixture model approach to decrease the bias that was cause by ambiguous criteria for determining sarcopenia. Finally, we compared the risk prediction ability of mixture Cox model with basic Cox proportional hazards model.

Material and Methods

- Study population: Between January 2010 and August 2015 patients diagnosed with HCC who had previously undergone TACE treatment at Linkou Chang Gung Memorial Hospital in Taiwan were included in this study.
- Basic Cox proportion hazards model: Candidates for inclusion in the Cox proportional hazard model were variables with a p-value less than 0.05 in the univariate analysis. By stepwise selection (p-value of ≤0.05), a group of factors that put together in the model producing the optimal results was identified.
- Mixture Cox proportional hazards model: Under this framework, it is not necessary to know which individual belongs to which of the subgroups beforehand. Assume an independent right-censored data consisting of $(Y_i, \delta_i, x_i), i = 1, \dots, n$, where Y_i denotes the survival time, x_i denotes regression covariates with coefficients, and δ_i denotes censoring indicators. Each patient derived from one of K latent classes with probability π_k,$k = 1, \dots, K$, $\sum_{i=1}^{k} \pi_k = 1$, and the observed data had the density $f_k(y, \delta|x) = [h_{0k}(y)\exp(x'\beta_k)]^{\delta}\exp[-H_{0k}(y)\exp(x'\beta_k)]$, where $h_{0k}(t)$ is the baseline hazard and $H_{0k}(t)$ is baseline cumulative hazard for the kth class. The density of the mixture model can be given by $f(Y, \Delta|x) = \prod_{i=1}^{n}\sum_{k=1}^{K}\pi_k f_k(Y_i, \delta_i|x_i)$. If we observed the latent class $U = (U_1, \dots, U_n)$, where U is distributed to a multinomial distribution, the density of the complete data can be presented as $f(Y, \Delta|x, U) = \prod_{i=1}^{n}\prod_{k=1}^{K}[\pi_k f_k(Y_i, \delta_i|x_i)]^{u_{ik}}$. Then we set the initial values for EM-algorithm to estimate parameters we wanted to know.
- Model-Performance Measure: Two important terms used to describe the elements of predictive accuracy: calibration and discrimination. Calibration refers to the degree of bias in the predicted outcome, and we used calibration plot to assess model's calibration; discrimination refers to a predictor's ability to distinguish between patients' different responses, and concordance index (C-index) was calculated to assess model's discriminative ability.

Results

- Seven factors significantly associated with poor outcome were identified: tumor size, tumor number, alpha-fetoprotein, aspartate aminotransferase, end stage renal disease, albumin, and sarcopenia.

Table 4.3 Multivariate analysis of prognostic factors in patients with HCC

Characteristic	**Multivariate analysis**		
Parameter	HR	95% CI	P value
SizeMax ≧5	1.543	1.115-2.137	0.009
Tumor number >1	1.491	1.078-2.061	0.0158
Albumin <3.5(g/dL)	1.416	1.032-.934	0.031
AFP >200 (ng/ml)	2.265	1.634-3.136	<.0001
AST >47 (U/L)	1.492	1.089-2.044	0.0128
ESRD present	4.1	1.914-8.658	0.0002
Sarcopenia present	1.468	1.042-2.068	0.0282

- A prognostic nomogram was constructed based on the results of selected prognostic factors from multivariate analysis.

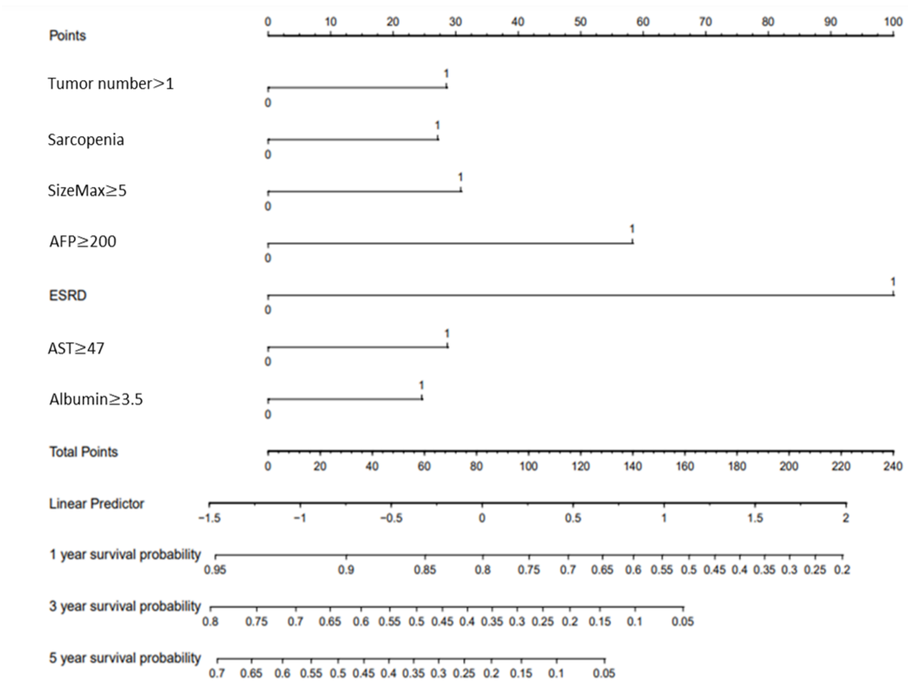

Figure 4.4 Nomogram predicting 1-, 3-, and 5-year survival probability

- We set the initial proportion of sarcopenia group as 50% and non-sarcopenia group as 50%, and the posterior probability of two components were 0.498% and 0.502%, respectively.
- The statistical significance of the majority of the factors was high in both components and these factors were associated with poor prognosis.
- Estimates of all the factors were in the same spectrum between two components and similar to basic Cox model.

Table 4.5 Mixture Cox proportional hazard model estimates by mixture group

	Component 1 P=0.498			**Component 2 P=0.502**		
Parameter	**HR**	**95% CI**	**P value**	**HR**	**95% CI**	**P value**
SizeMax ≧5	2.114	1.363-3.279	<.0001	1.002	0.592-1.695	0.995
Tumor number >1	1.144	0.727-1.799	0.489	1.876	1.164-3.002	0.011
Albumin >3.5(g/dL)	1.549	1.001-2.396	0.027	1.738	1.082-2.791	0.014
AFP >200 (ng/ml)	3.041	1.846-5.009	<.0001	2.305	1.433-3.708	0.001
AST >(47 U/L)	1,588	1.027-2.456	0.017	1.564	0.970-2.532	0.054
ESRD present	4.759	1.384-16.36	0.001	3.535	1.350-9.208	0.023

- The C-index of the nomogram to predict OS was 0.678 which was higher than those commonly used scoring or staging systems.
- The C-index of the mixture Cox proportional hazard model was 0.684 which was higher than that of basic Cox model.

Conclusions

- This is the first study to include sarcopenia in prognostic model for HCC patients who had received TACE therapy.
- With an easy-to-use presentation, model showed better performance compared with TNM staging system, BCLC system, ART scoring system, 6 and 12 score, up-to-7 criteria, etc.
- Mixture model can improve model's performance in the scenario of viewing sarcopenia as a cause of heterogeneity.