


Survival Prediction in Home Hospice Care Patients with Lung Cancer Based on LASSO Algorithm

Cancer Control
Volume 29: 1–10
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10732748221124519
journals.sagepub.com/home/ccx


Yicheng Zeng^{1, #}, Weihua Cao^{2, #}, Chaofen Wu¹, Muqing Wang¹, Yanchun Xie¹,
Wenxia Chen¹, Xi Hu¹, Yanna Zhou¹, Xubin Jing¹, and Xianbin Cai^{1, 3} 

Abstract

Purpose: The aim of the present study was to develop a nomogram for prognostic prediction of patients with lung cancer in hospice.

Methods: The data was collected from 1106 lung cancer patients in hospice between January 2008 and December 2018. The data were split into a training set, which was used to identify the most important prognostic factors by the least absolute shrinkage and selection operator (LASSO) and to build the nomogram, while the testing set was used to validate the nomogram. The performance of the nomogram was assessed by c-index, calibration curve and the decision curve analysis (DCA).

Results: A total of 1106 patients, including 835 (75%) from the training set and 271 (25%) from testing set, were retrospectively analyzed in this study. Using the LASSO regression, 5 most important prognostic predictors that included sex, Karnofsky Performance Scale (KPS), quality-of-life (QOL), edema and anorexia, were selected out of 28 variables. Validated c-indexes of training set at 15, 30, and 90 days were .778 [.737-.818], .776 [.743-.809], and .751 [.713-.790], respectively. Similarly, the validated c-indexes of testing set at 15, 30, and 90 days were .789 [.714-.864], .748 [.685-.811], and .757 [.691-.823], respectively. The nomogram-predicted survival was well calibrated, as the predicted probabilities were close to the expected probabilities. Moreover, the DCA curve showed that nomogram received superior standardized net benefit at a broad threshold.

Conclusions: The study built a non-lab nomogram with important predictor to analyze the clinical parameters using LASSO. It may be a useful tool to allow clinicians to easily estimate the prognosis of the patients with lung cancer in hospice.

Keywords

lung cancer, hospice care, nomogram, machine learning, prognosis

Received April 11, 2022. Received revised July 26, 2022. Accepted for publication August 4, 2022.

Introduction

Lung cancer is one of the malignant cancers that seriously threaten human health. According to the latest data from the International Agency for Research on Cancer, it is estimated that there were 2.09 million new lung cancer cases were reported in 2018,¹ and 1.76 million patients die out of it. In China, the lung cancer incidence and mortality was relatively high compared to most countries.^{2,3} More than 70% of patients with lung cancer diagnosed as advanced tumors, resulting in only 16.1% of lung cancer patients could survive

¹Department of Gastroenterology, The First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong, P.R. China

²Department of Hospice, The First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong, P.R. China

³Clinical Research Center, The First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong, P.R. China

#Yicheng Zeng and Weihua Cao were co-first authors and contributed equally to this work.

Corresponding Author:

Xianbin Cai, Department of Gastroenterology, The First Affiliated Hospital of Shantou University Medical College, 57 Changping Road, Shantou, Guangdong 515041, P.R. China.
Email: cxbin1@qq.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE

and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

over 5 years after diagnosis, which was lower than that of developed countries in Europe and the United States (20.1%).⁴

More and more attention was paid to patients with terminal-stage lung cancer in hospice in China. In China, the hospice is an institution that provides free hospice care for advanced cancer patients having financial conditions, funded by the Li Ka-shing Foundation. To improve the management of patients in hospice, research on the survival time of patients in hospice is indispensable. There were several tools that had been used to evaluate prognosis in terminally ill patients, such as the Palliative Prognostic Index (PPI), the Palliative Performance Scale (PPS) and modified Glasgow Prognostic Score (mGPS).⁵⁻⁷ The Zhou LJ et al constructed a simple Chinese Prognostic Scale (ChPS), to predict the survival rate of patients with terminal-stage cancer, of which accuracy rate of prediction wasn't satisfying.⁸ Jing C et al developed a new prognostic scale for ChPS (new-ChPS Scale) by a prospective survey on the prognostic factors.⁹ However, none of these models were tailor-made for patients with advanced lung cancer and the applicability of them still need to be verified. Some of the Scale incorporated biologic and imaging parameters, however, the patients in Chinese hospice could not afford their follow-up blood testing or imaging examination. Therefore, an effective and economical tool was proposed to evaluate the outcome of advanced cancer in our study, which could also triage the patients and inform their family members.

Cox proportional hazard model is the most common method for assessing the effects of various factors in the survival analysis. However, under the condition that the number of independent variables was high while sample size was low. The Cox proportional hazard model was associated with limitation such as multicollinearity, reduction in estimation precision, and non-interpretability of the coefficients.¹⁰ Least Absolute Shrinkage and Selection Operator (LASSO) is an advanced machine learning method, which can overcome the above problems by adding a penalized function to the estimation of the partial maximum likelihood. In this way, the coefficients of redundant variables become exactly zero and the most probable prognostic factors are retained in the model.^{11,12} In some of the previous research, LASSO method was applied and its superiority over traditional methods was confirmed through different studies.^{13,14} In Viet-Huan Le's study, LASSO regression model was applied to find out the best CT-based radiomics features for predicting the overall survival (OS) of lung cancer.¹⁵

In this study, through collecting and analyzing the data of the patients with lung cancer in hospice, we tried to identify the most important prognostic factors by LASSO method. Furthermore, we constructed a user-friendly nomogram with the selected variables, which help clinicians to give rise to rapid computation and evaluate the prognosis of patient. Predicting the prognosis of patient had the following meanings. First, medical staff in the hospice made follow-up strategies according to patients' conditions. Patients with

poor prognosis need more frequent visits, and the prediction model could provide a certain reference. Second, in China, many patients were very concerned about their survival time, which was related to whether they need to deal with several personal matters urgently, such as the disposal of property and the fulfillment of last wishes. Third, the prediction of survival time help patients prepare for the future challenges psychologically and practically.¹⁶ A good prognosis not only could increase patients' confidence in survival but reduce their mental burden.¹⁷

Material and Method

Study Population

We restricted our study cohort to 1106 patients who were diagnosed with primary lung cancer between January 2008 and December 2018. The information of patients was obtained from the Hospice Unit of the First Affiliated Hospital of Shantou University. The study was approved by the Ethics Committee of the First Affiliated Hospital of Shantou University (approval number: B-2022-164). Requirement for informed consent was waived because the study was retrospective and the identity of all patients remained undisclosed.

Variables Extraction

The baseline demographics included age, gender, ethnicity, literacy, history of alcohol use, smoking, history and effect of analgesic treatment, awareness of the disease and past medical history (hypertension or diabetes). The cancer-related information included metastasis, previous cancer treatment, duration of pain, concomitant symptoms, previous analgesic treatment, and its effect. The Karnofsky Performance Scale (KPS) was used to assess patient's performance status, which was translated into Chinese.¹⁸ The lowest score of KPS is 0 and the highest score is 100. The higher the score, the better the health status of the patient. The quality-of-life (QOL) scale in the study was developed by Dr Sun Yan in the 1990s by adapting widely used international scales to a version suitable for China.¹⁹ The QOL scale consists of 12 items (energy, sleep, appetite, activities of daily life, perception of cancer, attitude toward treatment, facial expression, fatigue, work relationships, pain, side effects of treatment and family relationships), with a total score of 60. X-tile 3.6.1 software (Yale University, New Haven, CT, USA) was employed to determine the best cutoff for KPS/QOL classified as different groups.²⁰ Karnofsky Performance Scale was categorized as 30 or lesser, 40, and 50 or more. Quality-of-life was divided into 3 levels: 30 or lesser, 31-35, 36 or more. The numeric rating scale (NRS) score was used to evaluate the level of pain. The score of 0-3 is mild pain, a score of 4-7 is moderate pain, and a score of 8-10 is severe pain.²¹ The survival time was defined as the number of days from registration to an event (dead or service paused). All the information was collected and

Table 1. Demographic, Clinical, and Tumor Characteristics of Patients with Lung Cancer in Hospice between the Training and Testing Set.

Characteristic	Overall (n = 1106)	Testing set (n = 271)	Training set (n = 835)	P
Time (median [IQR]) days	45 [21, 84]	45 [19, 84]	45 [21, 84]	.765
Sex (%)				
Female	301 (27.2)	73 (26.9)	228 (27.3)	.968
Male	805 (72.8)	198 (73.1)	607 (72.7)	
Age (median [IQR])	62 [56, 71]	62 [56, 70]	62 [56, 71]	.784
Area (%)				
Rural	330 (29.8)	89 (32.8)	241 (28.9)	.243
Urban	776 (70.2)	182 (67.2)	594 (71.1)	
Education (%)				
Illiteracy	159 (14.4)	34 (12.5)	125 (15.0)	.231
Primary school	542 (49.0)	145 (53.5)	397 (47.5)	
Middle school	276 (25.0)	68 (25.1)	208 (24.9)	
High school	106 (9.6)	18 (6.6)	88 (10.5)	
High school above	23 (2.1)	6 (2.2)	17 (2.0)	
Awareness of the disease (%)				
Full understanding	577 (52.2)	138 (50.9)	439 (52.6)	.841
Partial understanding	145 (13.1)	38 (14.0)	107 (12.8)	
Complete ignorance	384 (34.7)	95 (35.1)	289 (34.6)	
Metastasis (%)	986 (89.2)	246 (90.8)	740 (88.6)	.380
Operation (%)	104 (9.4)	25 (9.2)	79 (9.5)	.999
Chemotherapy (%)	385 (34.8)	84 (31.0)	301 (36.0)	.149
Radiotherapy (%)	172 (15.6)	45 (16.6)	127 (15.2)	.650
Duration (%)				
<1 month	216 (19.5)	47 (17.3)	169 (20.2)	.016
1-6 months	730 (66.0)	179 (66.1)	551 (66.0)	
6-12 months	111 (10.0)	24 (8.9)	87 (10.4)	
>12months	49 (4.4)	21 (7.7)	28 (3.4)	
Formal palliative care (%)				
None	156 (14.1)	43 (15.9)	113 (13.5)	.702
NSAIDs	170 (15.4)	39 (14.4)	131 (15.7)	
Weak opioids	370 (33.5)	86 (31.7)	284 (34.0)	
Strong opioids	410 (37.1)	103 (38.0)	307 (36.8)	
Analgesic effect (%)				
Bad	158 (14.3)	43 (15.9)	115 (13.8)	.670
Average	154 (13.9)	33 (12.2)	121 (14.5)	
Good	661 (59.8)	164 (60.5)	497 (59.5)	
Excellent	133 (12.0)	31 (11.4)	102 (12.2)	
Hypertension (%)	161 (14.6)	50 (18.5)	111 (13.3)	.046
Diabetes (%)	86 (7.8)	20 (7.4)	66 (7.9)	.881
Smoke (%)	327 (29.6)	80 (29.5)	247 (29.6)	.999
Drink (%)	73 (6.6)	12 (4.4)	61 (7.3)	.129
Formal constipation (%)	584 (52.8)	157 (57.9)	427 (51.1)	.060
Weight lose (%)	963 (87.1)	238 (87.8)	725 (86.8)	.748
Insomnia (%)	569 (51.4)	143 (52.8)	426 (51.0)	.667
Anorexia (%)	880 (79.6)	211 (77.9)	669 (80.1)	.475
Nausea (%)	206 (18.6)	46 (17.0)	160 (19.2)	.475
Vomiting (%)	207 (18.7)	48 (17.7)	159 (19.0)	.691
Abdominal distention (%)	39 (3.5)	8 (3.0)	31 (3.7)	.689
Tachypnea (%)	632 (57.1)	161 (59.4)	471 (56.4)	.425
Edema (%)	111 (10.0)	31 (11.4)	80 (9.6)	.442
QOL (%)				

(continued)

Table 1. (continued)

Characteristic	Overall (n = 1106)	Testing set (n = 271)	Training set (n = 835)	P
≤30	376 (34.0)	91 (33.6)	285 (34.1)	.978
31-35	437 (39.5)	107 (39.5)	330 (39.5)	
≥36	293 (26.5)	73 (26.9)	220 (26.3)	
NRS (%)				.136
≤3	88 (8.0)	14 (5.2)	74 (8.9)	
4-7	821 (74.2)	205 (75.6)	616 (73.8)	
≥8	197 (17.8)	52 (19.2)	145 (17.4)	
KPS (%)				.083
≤30	318 (28.8)	92 (33.9)	226 (27.1)	
40	484 (43.8)	107 (39.5)	377 (45.1)	
≥50	304 (27.5)	72 (26.6)	232 (27.8)	

Abbreviations: Values are presented as no. (%) or median (Q1, Q3)

recorded by 2 qualified doctors during the first follow-up visit. Multiple Imputation was used to handle the missing data.²²

Statistical Analysis

The patients were split into a training set and testing set in a random manner without replacement at a ratio of 3:1. To evaluate the differences between the training and testing sets, continuous variables with normal distribution were presented as the mean (\pm standard deviation) using student t-tests, while continuous variables with skewed distribution were presented as the median interquartile range (IQR) using the Mann-Whitney U test. Categorical variables were presented as frequency (proportion) and chi-square tests was applied for their comparisons. Kaplan-Meier curves with risk table were utilized to display the survival of the patients from training and testing sets respectively.

Least absolute shrinkage and selection operator regression was a machine learning algorithm first proposed by Robert Tibshirani in 1996. In this study, the LASSO regression was used to estimate the coefficients of COX regression model. With LASSO method, coefficients of unimportant variables were penalized to zero and important variables were retained, which enabled to adjust for model's over fitting and avoid extreme predictions. For our analysis, the lasso method was used to screen out the most representative variables for further multivariate COX regression analysis and construction of nomogram which enabled to predict the 15-days, 30-days and 90-days survival probability of the patients.

In evaluating the performance of the proposed nomogram, we employed both calibration, which was performed using 1000 bootstrap resamples, and the concordance index (C-index) which measured the classification accuracy. Furthermore, we also applied the decision curve analysis (DCA), a novel method to evaluate the nomogram from the perspective of clinical consequences by calculating the net benefit.

All analyses were carried out with R (Version 3.6.2, R Foundation, Vienna, Austria) and R packages ('survminer',

'glmnet', 'rms', 'timeROC', 'mice', 'ggDCA'). *P*-value < .05 was considered statistically significant.

Result

A total of 1106 patients, including 835 (75%) from the training set and 271 (25%) from testing set, were retrospectively analyzed in this study. Characteristics of patients were described in Table 1. The overall population consisted of 805 males (72.8%) and 301 females (27.2%), with a median survival time of 45 days. Most of them lived in urban areas and had accepted a basic level of education (primary school or above). However, only 52.2% of the patients had fully understood their physical conditions. In terms of treatment, 9.4% had undergone surgery, 34.8% had undergone chemotherapy, and 15.6% of the patients had undergone radiotherapy, while the majority (89.2%) had evidence of tumor metastasis. What's more, 92% of the patients suffered from moderate or severe pain, which led to the long-term use of analgesic medicine. As for evaluation scale, 376 (34.0%) patients' QOL score were less than 34 and 318 (28.8%) patients' KPS score were less than 30. The general characteristics of the patients showed that both datasets were similar with respect to all variables except hypertension and duration of pain. In the Kaplan-Meier plot (Figure 1), the survival situation of patients in the training set and testing set was visually displayed. The result also showed that there was no difference between these 2 group (*P* = .6). In addition, only 63 (7.5%) out of 835 patients survived beyond 6 months in the training set while 15 (5.5%) out of 271 patients survived beyond 6 months in the testing set.

Using the LASSO regression, 5 prognostic predictors which included sex, KPS, QOL, edema and anorexia, were selected out of 28 variables which were probably associated with OS in the training set (Figure 2). The optimal λ value for LASSO regression with 10-fold cross-validation was .1026. According to the coefficients obtained from LASSO method, we inferred that KPS was the most important factor in

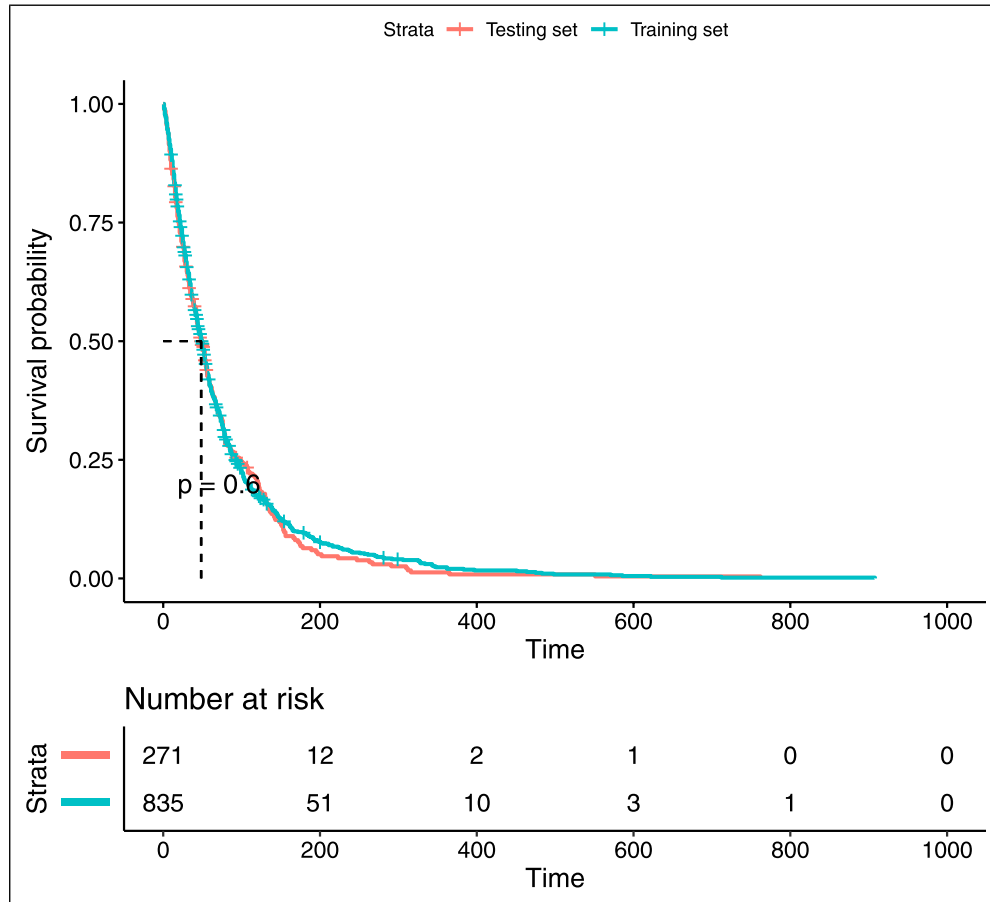


Figure 1. Kaplan-Meier curves with risk table for patients with lung cancer in training set and testing set.

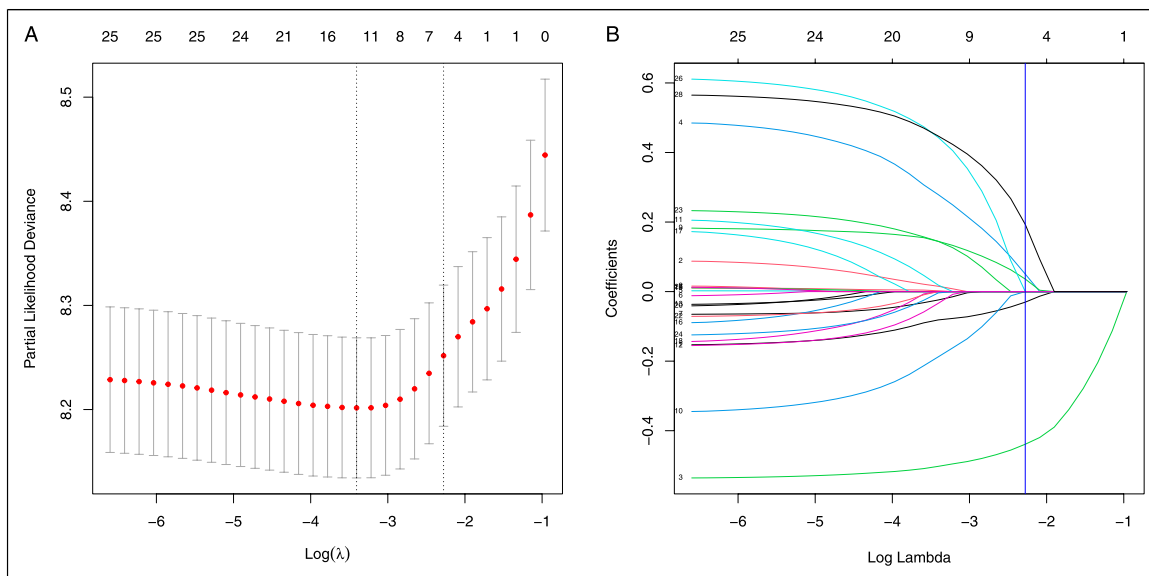


Figure 2. Selection of predictors using the LASSO regression analysis in patients with lung cancer (A) Using 10-fold cross-validation, the dotted vertical lines were drawn at the optimal values by minimum criteria and 1-s.e. Criteria (B) LASSO coefficient profiles of the 28 variables. The vertical line was drawn in terms of the formula ($x = \log(\lambda_{1-s.e.})$). At the optimal values $\lambda_{1-s.e.} = .1026$, 5 variables (sex, anorexia, edema, QOL and KPS) with a nonzero coefficient were finally identified.

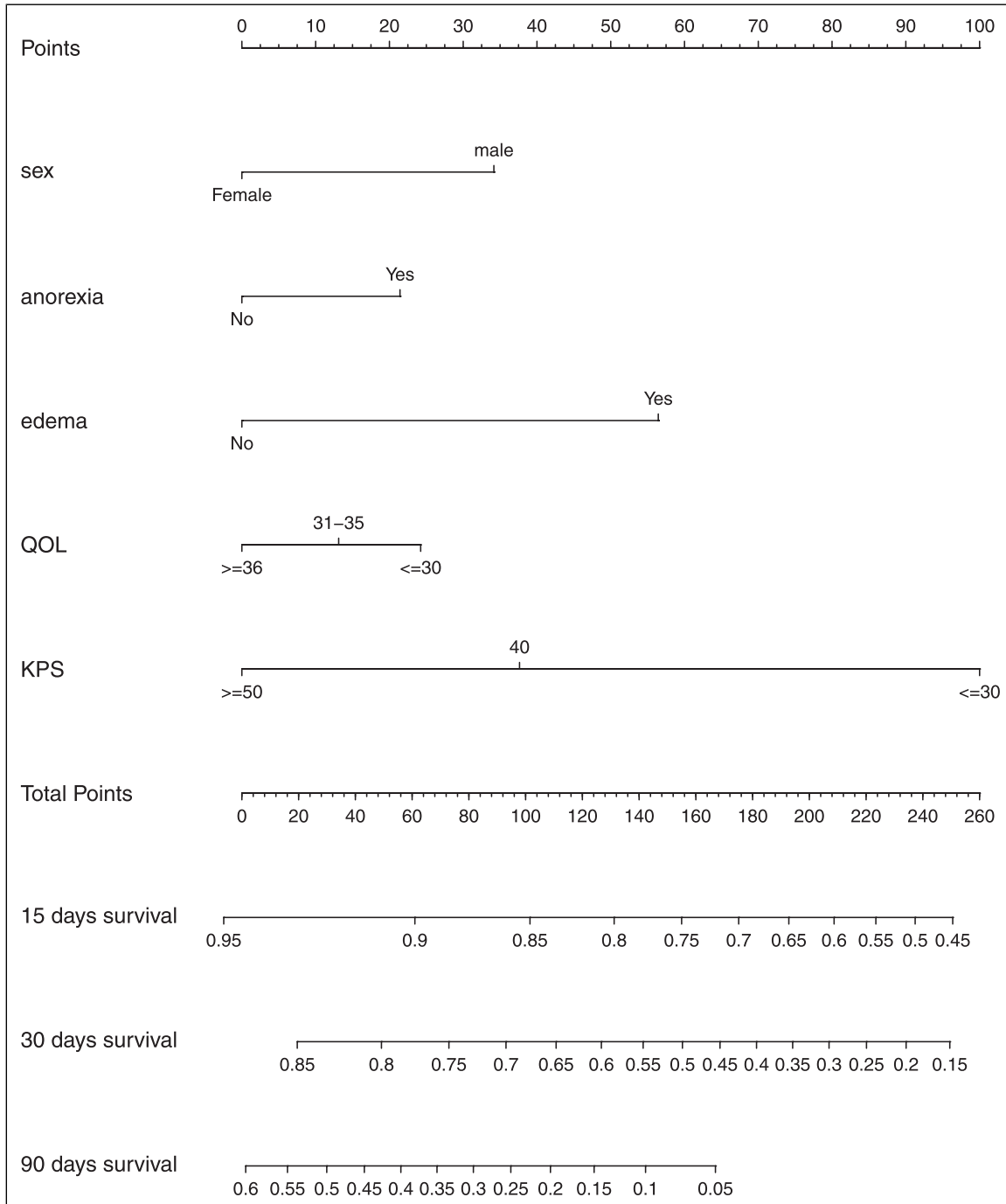


Figure 3. The Nomogram for predicting 15-days, 30-days and 90-days OS.

predicting survival probability. Furthermore, a nomogram with 5 prognostic predictors above selected by LASSO regression was constructed based on COX regression model for predicting the survival rate (Figure 3). To use the nomogram, a patient can obtain each variable score by matching its value to the top points axis. The total sum of each variable score was marked on total points axis and a line was drawn downward to determine the probability of median survival time. To examine

the performance of our predictive nomogram, we employed both discrimination and calibration assessments. As shown in Table 2, C-index analysis for the nomogram showed a good discrimination at 15, 30, and 90 days in both training set (C-index = .778 (95% CI .737-.818), .776 (95% CI .743-.809), and .751 (95% CI .713-.790), respectively) as well as in testing set (C-index = .789 (95% CI .714-.864), .748 (95% CI .685-.811), and .757 (95% CI .691-.823), respectively). The

nomogram-predicted survival was well calibrated at 15, 30, and 90 days by the training and testing sets, and the predicted probabilities were close to the expected probabilities (Figure 4). Moreover, the DCA curve was used to assess the clinical utility of the nomogram by calculating the net benefit. The

result showed that nomogram received superior standardized net benefit at a broad threshold (Figure 5).

Discussion

The incidence and mortality of lung cancer were both currently ranking first among all cancer reported worldwide. In China, the lung cancer mortality was relatively high compared to most countries.⁴ However, China's hospice system was established late, that results in limited research on lung cancer patients. Most of the study on hospice patients in China were the traditional survival analysis without offering a practical tool to evaluate the prognosis of these patients.

The current study collected the follow-up data of 1106 patients with lung cancer from the Hospice Unit of the First Affiliated

Table 2. The Concordance Index (C-index) with 95% CI at 15, 30, and 90 Days in Both Training Set and Testing Set.

	C-Index (95% CI)		
	15 days	30 days	90 days
Training set	.778 (.737-.818)	.776 (.743-.809)	.751 (.713-.790)
Testing set	.789 (.714-.864)	.748 (.685-.811)	.757 (.691-.823)

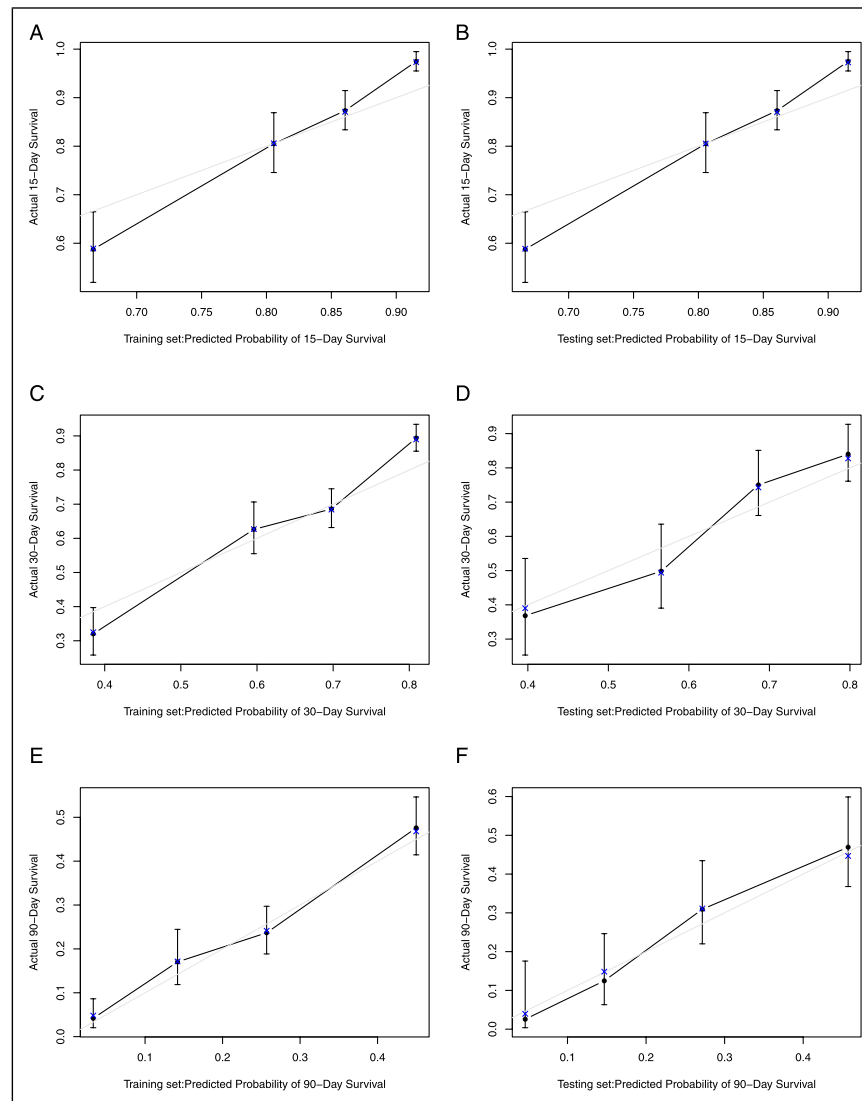


Figure 4. Calibration curves for predicting overall survival rate by the nomogram in the training and testing set. Calibration curves of the prognostic nomogram for 15-days overall survival (A), 30-days overall survival (C) and 90-days overall survival (E) in the training set; calibration curves for 15-days overall survival (B), 30-days overall survival (D), and 90-days overall survival (F) in the testing set.

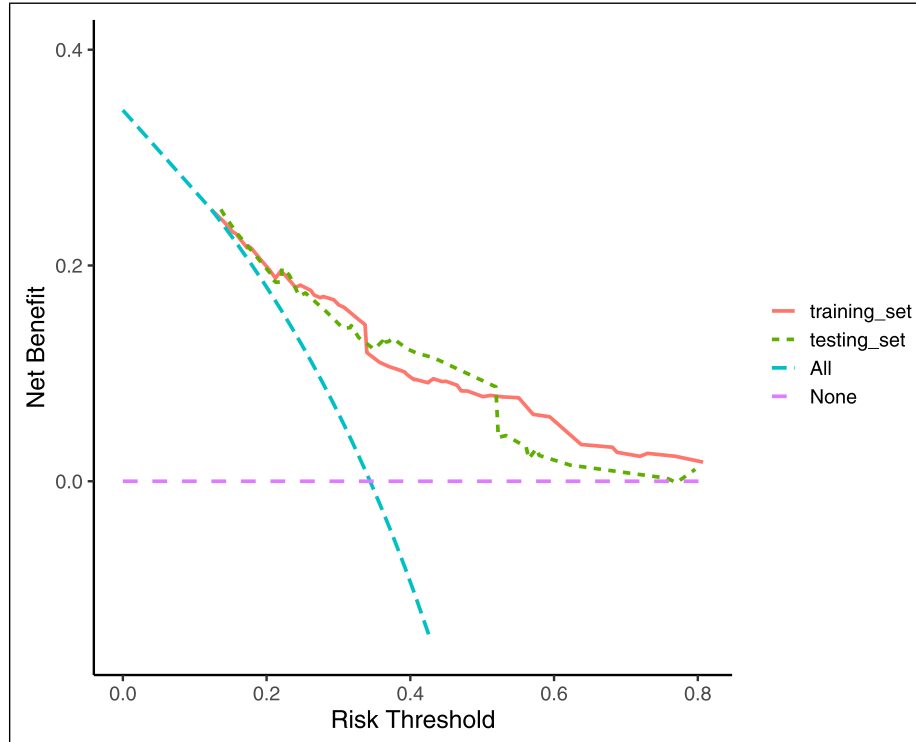


Figure 5. The decision curves analysis curve of the prognostic nomogram in the training and testing set.

Hospital of Shantou University Medical College. Different from the traditional COX regression method for survival analysis, our study adopted an advanced algorithm of machine learning-LASSO, which can efficiently screen out key variables from many clinical indicators. Available pieces of evidence suggest that LASSO has better predictive performance than traditional models.^{13,14} Furthermore, based on the selected predictors, we interpreted a nomogram for clinicians to quickly assess the prognosis of individual with lung cancer. Considering that most patients in hospice could not afford their follow-up blood testing and imaging examination, we chose not to incorporate any laboratory indexes into our predictive model. Furthermore, the economical and practical model still performed well in prediction.

In this study, we identified sex as a significant prognostic predictor for patients with lung cancer in hospice. According to the statistics of World Health Organization (WHO), the cumulative mortality risk of males with lung cancer was far higher than females in 2018 (3.19% vs 1.32%), which proved that sex was a significant factor related to prognosis.²³ The KPS and QOL were widely recognized as an effective indicator for assessing the survival status of the patients, and the results of this study was consistent with previous study.²⁴ In the existing literature, KPS and QOL were also used to construct the prognostic models and had good predictive performance.²⁵⁻²⁷ Moreover, some evidence suggested that certain symptoms have an important impact on the prognosis of cancer patients.^{9,28,29} In our study, symptoms including edema and anorexia were selected by LASSO as key

predictors in the model. According to the existing studies anorexia is one of the typical manifestations of cachexia and cancer and anorexia-cachexia syndrome (CACS) is present in 57-61% of patients with lung cancer, which was directly attributable for 20% of cancer deaths.³⁰ Therefore, it is convincing that the anorexic was selected as one of the predictors in our study. We recommended a unique and robust model which consists of prognosis scales and symptoms for survival analysis in hospice patients with lung cancer.

There were still several limitations in our study. Firstly, this was a retrospective study, which may cause recall bias and prevent our model from getting better performance. Secondly, since most of our patients survive less than half a year, hence, this may lead to inaccurate predictions of the prognosis of the patients with longer survival times. To solve this problem, we have adopted 10-fold cross-validation to reduce this error. Thirdly, since our data was only from one single research center and in the absence of further external validation, the use of this prediction model in other hospice care center should be cautious. Last but not least, our nomogram didn't require any laboratory indicators, which probably prevented the nomogram from reaching excellent performance. However, the majority of home hospice care patients in China was low-income and it was unrealistic for them to afford the lab examination. Combined with the situation of Chinese hospice, our non-lab nomogram could be a compromise and economical tool to predict patients' prognosis. In future, studies involving large-sample size and multi-center still need to be carried out and incorporated to improve our nomogram.

Conclusion

This study identified the most important and non-lab based prognostic factors by LASSO method and built a nomogram for clinical use. Our finding might be an important contribution to the prediction of patients with lung cancer in hospice, allowing clinicians to easily estimate the status of their patients and to help adjust their follow-up management.

Abbreviations

LASSO: least absolute shrinkage and selection operator

KPS: Karnofsky Performance Scale

QOL: quality-of-life

NRS: numeric rating scale

IQR: interquartile ranges

OS: overall survival

DCA: decision curve analysis

CI: confidence interval

Authors' Contributions

Yicheng Zeng, Xianbin Cai, Weihua Cao, and Xubin Jing contributed to the idea and design. Yicheng Zeng, Chaofen Wu, Muqing Wang, Yanchun Xie, and Wenxia Chen contributed to the data collection. Yicheng Zeng, Xi Hu, Yanna Zhou contributed to the data analysis. Yicheng Zeng, Xubin Jing and Xianbin Cai contributed to the manuscript writing and revision. All authors approve the final version of the manuscript.

Availability of Data and Material

The data used and analysed in the study are available from the corresponding author on reasonable request.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the grants from the Medical Scientific Research Foundation of Guangdong Province, China (grant no.201811981727429), the Guangdong Science and Technology Department, China (grant no. 20200304–65), the Li Ka Shing Foundation “Heart of Gold” National Hospice Services Program.

ORCID iD

Xianbin Cai  <https://orcid.org/0000-0002-3967-8666>

Ethical Approval and Ethical Standards

This study was approved by the ethical review board of the First Hospital Affiliated of Shantou University Medical College (approval number:B-2022-164) and was conducted in accordance with the standards of the Declaration of Helsinki. Informed consent was waived because of the retrospective nature of the study.

References

1. Ferlay J, Ervik M, Lam F, et al. *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer; 2018. [Cited 5 Nov 18.] Available from URL: <https://gco.iarc.fr/today>
2. Zeng H, Zheng R, Guo Y, et al. Cancer survival in China, 2003-2005: A population-based study. *Int J Cancer*. 2015;136(8):1921-1930.
3. Lachmann P. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK. *Lancet*. 2011;377(9772):127.
4. Cao M, Chen W. Epidemiology of lung cancer in China. *Thorac Cancer*. 2019;10(1):3-7.
5. Morita T, Tsunoda J, Inoue S, Chihara S. The palliative prognostic index: A scoring system for survival prediction of terminally ill cancer patients. *Support Care Cancer*. 1999;7(3):128-133.
6. Baik D, Russell D, Jordan L, Dooley F, Bowles KH, Masterson Creber RM. Using the palliative performance scale to estimate survival for patients at the end of life: A systematic review of the literature. *J Palliat Med*. 2018;21(11):1651-1661.
7. McMillan DC, Crozier JEM, Canna K, Angerson WJ, McArdle CS. Evaluation of an inflammation-based prognostic score (GPS) in patients undergoing resection for colon and rectal cancer. *Int J Colorectal Dis*. 2007;22(8):881-886.
8. Lingjun Z, Jing C, Jian L, Wee B, Jijun Z. Prediction of survival time in advanced cancer: A prognostic scale for Chinese patients. *J Pain Symptom Manag*. 2009;38(4):578-586.
9. Cui J, Zhou L, Wee B, Shen F, Ma X, Zhao J. Predicting survival time in noncurative patients with advanced cancer: A prospective study in China. *J Palliat Med*. 2014;17(5):545-552.
10. Cox DR. Regression models and life-tables. *J Roy Stat Soc B*. 1972;34(2):187-220.
11. Tibshirani R. Regression shrinkage and selection via the lasso. *J Roy Stat Soc B*. 1996;58(1):267-288.
12. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16(4):385-395.
13. Garcia-Carretero R, Barquero-Perez O, Mora-Jimenez I, Soguero-Ruiz C, Goya-Esteban R, Ramos-Lopez J. Identification of clinically relevant features in hypertensive patients using penalized regression: A case study of cardiovascular events. *Med Biol Eng Comput*. 2019;57(9):2011-2026.
14. Zhang Z, Hong Y. Development of a novel score for the prediction of hospital mortality in patients with severe sepsis: The use of electronic healthcare records with LASSO regression. *Oncotarget*. 2017;8(30):49637-49645.
15. Le VH, Kha QH, Hung TNK, Le NQK. Risk score generated from CT-based radiomics signatures for overall survival prediction in non-small cell lung cancer. *Cancers*. 2021;13(14).
16. den Daas N. Estimating length of survival in end-stage cancer: A review of the literature. *J Pain Symptom Manag*. 1995;10(7):548-555.
17. Stone PC, Lund S. Predicting prognosis in patients with advanced cancer. *Ann Oncol*. 2007;18(6):971-976.
18. Friendlander AH, Ettinger RL. Karnofsky performance status scale. *Spec Care Dentist*. 2009;29(4):147-148.

19. Sun Y. China: Status of cancer pain and palliative care. *J Pain Symptom Manag.* 1993;8:399-403.
20. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: A new bioinformatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res.* 2004;10(21):7252-7259.
21. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual analog scale for pain (VAS Pain), numeric rating scale for pain (NRS Pain), mcGill pain questionnaire (MPQ), short-form mcGill pain questionnaire (SF-MPQ), chronic pain grade scale (CPGS), short form-36 bodily pain scale (SF-36 BPS), and measure of intermittent and constant osteoarthritis pain (ICOAP). *Arthritis Care Res.* 2011;63(suppl 11):S240-S252.
22. Pedersen AB, Mikkelsen E, Cronin-Fenton D, et al. Missing data and multiple imputation in clinical epidemiological research. *Clin Epidemiol.* 2017;9:157-166.
23. World Health Organization. Lung cancer. Available at: <https://gco.iarc.fr/today/data/factsheets/cancers/15-Lung-fact-sheet.pdf>. Accessed March 12, 2021.
24. Yildiz Celtek N, Suren M, Demir O, Okan I. Karnofsky performance scale validity and reliability of Turkish palliative cancer patients. *Turk J Med Sci.* 2019;49(3):894-898.
25. Nie M, Liu C, Pan Y-C, et al. Development and evaluation of oral cancer quality-of-life questionnaire (QOL-OC). *BMC Cancer.* 2018;18(1):523.
26. Chambless LB, Kistka HM, Parker SL, Hassam-Malani L, McGirt MJ, Thompson RC. The relative value of postoperative versus preoperative karnofsky performance scale scores as a predictor of survival after surgical resection of glioblastoma multiforme. *J Neuro Oncol.* 2015;121(2):359-364.
27. Abernethy AP, Shelby-James T, Fazekas BS, Woods D, Currow DC. The Australia-modified karnofsky performance status (AKPS) scale: A revised scale for contemporary palliative care clinical practice [ISRCTN81117481]. *BMC Palliat Care.* 2005;4:7.
28. Ohde S, Hayashi A, Takahasi O, et al. A 2-week prognostic prediction model for terminal cancer patients in a palliative care unit at a Japanese general hospital. *Palliat Med.* 2011;25(2):170-176.
29. Huang Y, Xi Q, Xia S, et al. Development and validation of a prognostic scale for hospitalized patients with terminally ill cancer in China. *Support Care Cancer.* 2014;22(1):145-152.
30. Arrieta O, Luvián-Morales J, Turcott JG, Oñate-Ocaña LF. Quality of life and anorexia/cachexia in lung cancer: Validation of the Spanish version of the FAACT instrument. *Qual Life Res.* 2018;27(10):2709-2718.