Prognostic analysis and clinical characteristics of dual primary lung cancer: a population study based on surveillance, epidemiology, and end results (SEER) database

Guanghui Wang1,2 · Yukai Zeng1 · Haotian Zheng1 · Xiaogang Zhao3 · Yadong Wang1 · Hongchang Shen1 · Jiajun Du1,2

Received: 18 October 2021 / Accepted: 20 February 2022 / Published online: 22 March 2022 © The Author(s), under exclusive licence to The Japanese Association for Thoracic Surgery 2022

Abstract

Background The peculiarity and the lack of clinical studies of dual primary lung cancer (DPLC) led to limited knowledge about its clinical characteristics and prognosis. This study performed a retrospective analysis to assess the prognostic factors and clinical characteristics of DPLC.

Methods A total of 1419 DPLC patients from SEER were analyzed by univariate and multivariable Cox regression analyses. The independent prognostic factors were included to establish a nomogram. The accuracy and reliability of prognostic model were evaluated by C indexes, calibration plots, receiver operating characteristic (ROC) curves, decision curve analyses (DCA) and integrated discrimination improvement (IDI) scores. Chi-square test was used to assess the differences between DPLC and single primary lung cancer (SPLC) or synchronous DPLC (sDPLC) and metachronous DPLC (mDPLC).

Results Cox regression analysis showed that age, sex, histological type, stage, lymph node (LN) metastasis, surgery, chemotherapy were independent prognostic factors, we included these factors to establish a nomogram. In the training cohort, the C index was 0.690, and the area under curves (AUC) of 3 and 5-year survival time were 0.720 and 0.723. The calibration plots in training cohort and validation cohort were in excellent agreement. DCA and IDI showed that the predictive effect of the novel prognostic model was better than the model based on 8th AJCC TNM system. Chi-square test indicated that DPLC and SPLC had statistical differences on pathological and clinical features.

Conclusions The clinical and pathological characteristics of DPLC were different from the SPLC. The nomogram could provide accurate and individualized survival predictions for DPLC.

Keywords Dual primary lung cancer · Prognostic model · Nomogram · COX analysis · The Surveillance · Epidemiology · End Results database

Introduction

Lung cancer has the high incidence and mortality, and is a leading cause of cancer death worldwide [1]. The incidence of dual primary lung cancer (DPLC) has also increased due to the improvement of medical level and the extension of survival time of lung cancer patients [2]. DPLC occupies the majority of multiple primary lung cancer (MPLC). And there are two subtypes of DPLC including synchronous DPLC (occurring at the same time) and metachronous DPLC (occurring at different times) [5, 6]. The incidence of DPLC is approximately 0.8–14.5% per year for per non-small cell lung cancer patient who accesses to treatment [3, 4]. Nevertheless, little studies pay attention to DPLC, clinicians still don’t have a clear idea of the clinical features and
prognosis about DPLC. Although the TNM stage system is used to evaluate prognosis widely, this criterion is more applicable to single primary lung cancer (SPLC). There are some limitations and disadvantages using TNM stage system to assess DPLC, because DPLC has different clinical features and pathological characteristics from SPLC.

It is necessary to study the differences between SPLC and DPLC, and a dedicated clinical prognostic model for DPLC is urgently needed. The nomogram is a good choice for this purpose which evaluates the prognosis of DPLC patients owing to its convenience and accuracy [8, 9]. Given the resulting paucity of data on prognosis and optimal management of DPLC, the aim of this study was to establish an accurate and reliable prognostic model which could predict the overall survival (OS) of DPLC patients and explore their clinical features. We performed a retrospective analysis based on the complete clinical information of DPLC and SPLC from the Surveillance, Epidemiology, and End Results (SEER) database and Provincial Hospital Affiliated to Shandong University. Verified nomogram which has compared with TNM stage system was used to predict the OS, and chi-square test was performed to research the clinical features of DPLC.

**Patients and methods**

**Data source**

The relevant clinical information of DPLC and SPLC was extracted from SEER database between 2004 and 2015 using SEER*Stat program version 8.3.5 (http://seer.cancer.gov/seerstat/). In addition, MPLC patients’ information was collected and selected from Provincial Hospital Affiliated to Shandong University. Verified nomogram which has compared with TNM stage system was used to predict the OS, and chi-square test was performed to research the clinical features of DPLC.

**Patient and variable selection**

The following criteria were performed to select the DPLC patients in this study: (1) year of diagnosis was between 2004 to 2015; (2) site and morphology (site recoded ICD-O-3/WHO 2008) were lung and bronchus; (3) site and morphology (behavior recode for analysis) were malignant; (4) the lung cancers of patients were primary base on ‘Sequence number’, ‘First Malignant Primary Indicator’, ‘primary by International rules’ from SEER database; (5) the patients were defined as DPLC which their IDs were registered two times and the patients were defined as SPLC which their IDs were registered once. Patients with one primary lung cancer or two primary lung cancer were included for this study, and the information of status, survival time, AJCC stage, and AJCC N and grade was clear. A total of 1419 DPLC patients and 70,198 SPLC patients with pathological confirmation were selected from SEER database and 173 MPLC patients and 173 SPLC patients were selected from Provincial Hospital Affiliated to Shandong University. The variables included age at diagnosis, sex, race record, primary site label, laterality, grade, ICO-O-3 Hist/behav, malignancy, the time interval (months since the first tumor to the second tumor of the DPLC), AJCC TNM system, COD to site recode, Rx Summ-Surg Prim Site (1998 +), radiation record, chemotherapy record, smoke, chronic disease, familial-hereditary disease and pre-cancer symptoms. We selected the grade with poorer differentiation and the later stage of the first primary lung cancer and second primary lung cancer as the final grade and stage of the patient. The patients’ stage information recorded by 6th AJCC system was modified into 8th AJCC system.

**Statistical analysis**

All statistical analyses were performed by SPSS (version 24.0; SPSS Inc., Chicago, IL) and R software (4.0.5). Survival time was counted from date of diagnosis (the second tumor of DPLC) to date of death. Simple random sampling was performed by SPSS. The univariate analysis and the COX multivariate analysis were used to explored the independent prognostic factors of SPLC and DPLC. OS analysis was performed by Kaplan–Meier method. The propensity score matching (PSM) was used to eliminate the differences between SPLC and DPLC. We used the chi-square test to examine the variables. X-tile software was used to calculate cut-off value of age. Results were reported by the hazard ratio and 95% confidence interval. All p values were two-sided, and p values less than 0.05 were considered statistically significant.

**Results**

**Univariate and multivariate COX analysis for overall survival**

According to the univariate analysis of 1419 DPLC patients, the results showed age, race, sex, histological type, grade, stage, LN metastasis, surgery, chemotherapy were associated with the overall survival of DPLC patients. After univariate analysis, Cox multivariate analysis were performed to examine the effect of each clinical variables on OS, and the analysis clarified that age, sex, histological type, stage, LN metastasis, surgery, chemotherapy were significant prognostic factors (log-rank test, all p value < 0.05; Table 1). Aged ≥ 72 years old (HR 1.531; 95% CI 1.271–1.844; p < 0.001) was associated with worse OS. The female patients were associated with better prognosis (HR 0.654; 95% CI 0.566–0.755; p < 0.001). The prognosis of patients
Table 1 Univariate analysis and multivariate COX analysis of DPLC

| Variables               | n   | OS | Univariate analysis | Multivariate analysis |
|-------------------------|-----|----|---------------------|-----------------------|
|                         |     |    | HR(95% CI)          | HR(95% CI)            |
|                         |     |    | p Value             | p Value               |
| Age ≤64                 | 370 | 1  | 0.052               | 1.130 (0.930–1.374)   | 0.217 |
| Age 65–71               | 429 | 1  | 1.612 (1.354–1.919) | 1.531 (1.271–1.844)   | <0.001 |
| Age ≥72                 | 620 | 1  | 1.612 (1.354–1.919) | 1.531 (1.271–1.844)   | <0.001 |
| Race                    | 1419| 0.035 | 0.165             |                       |
| White                   | 1225| 1  | 1.208 (0.998–1.461) | 1.046 (0.808–1.354)   | 0.732 |
| Black                   | 113 | 1  | 0.657 (0.468–0.922) | 0.723 (0.513–1.021)   | 0.065 |
| Sex                     | 1419| 1  | 0.597 (0.521–0.685) | 0.654 (0.566–0.755)   | <0.001 |
| Race                    | 1419| 0.692 |                   |                       |
| Primary site            | 1419| 1  | 1.158 (1.027–1.305) | 1.130 (1.027–1.305)   | <0.001 |
| Histological type       | 1419| 0.007 |                   |                       |
| Ad + Ad                 | 763 | 1  | 1.531 (1.231–1.905) | 1.359 (1.082–1.708)   | 0.008 |
| Ad + Sq                 | 154 | 1  | 1.759 (1.490–2.077) | 1.284 (1.033–1.596)   | 0.024 |
| Sq + Sq                 | 323 | 1  | 1.381 (1.123–1.699) | 1.248 (1.033–1.596)   | 0.024 |
| Grade I                 | 110 | 1  | 1.846 (1.330–2.561) | 1.544 (1.102–2.165)   | 0.012 |
| Grade II                | 588 | 1  | 2.156 (1.559–2.982) | 1.566 (1.144–2.203)   | 0.010 |
| Grade III               | 699 | 1  | 2.210 (1.221–4.002) | 1.528 (0.821–2.845)   | 0.181 |
| Grade IV                | 22  | 1  | 1.864 (1.620–2.144) | 1.540 (1.254–1.891)   | <0.001 |
| Stage                   | 1419| 1  | 1.359 (1.082–1.708) |                       | <0.001 |
| Stage I                 | 686 | 1  | 1.270 (1.034–1.560) | 0.988 (0.766–1.274)   | 0.926 |
| Stage II                | 228 | 1  | 1.917 (1.614–2.277) | 1.422 (1.120–1.805)   | 0.004 |
| Stage III               | 310 | 1  | 2.388 (2.127–3.147) | 1.961 (1.563–2.460)   | <0.001 |
| Stage IV                | 195 | 1  | 2.156 (1.559–2.982) | 1.566 (1.144–2.203)   | 0.010 |
| LN metastasis           | 1419| 1  | 2.210 (1.221–4.002) | 1.528 (0.821–2.845)   | 0.181 |
| Radiotherapy No + No    | 893 | 1  | 1.177 (0.917–1.511) | 0.199                |
| Radiotherapy Yes + Yes  | 116 | 1  | 0.940 (0.755–1.171) | 0.580                |
| Radiotherapy Yes + No   | 169 | 1  | 1.023 (0.848–1.235) | 0.809                |
| Radiotherapy No + Yes   | 241 | 1  | 1.613 (1.312–1.985) | 1.214 (0.967–1.524)   | 0.095 |
| Chemotherapy            | 1419| 0.007 |                   |                       |
| Chemotherapy No + No    | 833 | 1  | 1.748 (1.475–2.071) |                       | <0.001 |
| Chemotherapy Yes + Yes  | 317 | 1  | 1.615 (0.948–2.751) | 1.447 (0.830–2.522)   | 0.192 |
| Chemotherapy No + Yes   | 25  | 1  | 2.189 (1.789–2.678) | 1.546 (1.236–1.934)   | <0.001 |
| Chemotherapy No + No    | 177 | 1  | 1.023 (0.848–1.235) | 0.809                |
| Chemotherapy No + Yes   | 241 | 1  | 1.613 (1.312–1.985) | 1.214 (0.967–1.524)   | <0.001 |
| Surgery                 | 1419| 1  | 1.375 (1.146–1.649) | 1.025 (0.824–1.275)   | 0.001 |
| Surgery No + No         | 900 | 1  | 1.748 (1.475–2.071) | 1.546 (1.236–1.934)   | <0.001 |
| Surgery Yes + Yes       | 317 | 1  | 1.615 (0.948–2.751) | 1.447 (0.830–2.522)   | 0.192 |
| Surgery No + Yes        | 25  | 1  | 2.189 (1.789–2.678) | 1.546 (1.236–1.934)   | <0.001 |
| Surgery No + No         | 177 | 1  | 1.023 (0.848–1.235) | 0.809                |
| Surgery No + Yes        | 241 | 1  | 1.613 (1.312–1.985) | 1.214 (0.967–1.524)   | <0.001 |
| Surgery No + No         | 833 | 1  | 1.748 (1.475–2.071) | 1.546 (1.236–1.934)   | <0.001 |
| Surgery Yes + Yes       | 201 | 1  | 1.635 (1.356–1.972) | 1.025 (0.824–1.275)   | 0.001 |
| Surgery Yes + No        | 166 | 1  | 0.962 (0.780–1.187) | 0.751 (0.598–0.943)   | 0.014 |
| Surgery No + Yes        | 219 | 1  | 1.613 (1.312–1.985) | 1.214 (0.967–1.524)   | <0.001 |
| Surgery No + No         | 177 | 1  | 1.613 (1.312–1.985) | 1.214 (0.967–1.524)   | <0.001 |
| Interval                | 1419| 0.163 |                   |                       |
| < 6 months              | 549 | 1  | 1.104 (0.960–1.270) | 0.163                |
| ≥6 months               | 870 | 1  | 1.104 (0.960–1.270) | 0.163                |

Ad Lung adenocarcinoma, Sq Lung squamous cell carcinomas

For radiotherapy, Yes + No means that the 1st tumor received chemotherapy and 2nd tumor didn’t received chemotherapy and the rest can be explained in the same manner.
with lung squamous cell carcinomas (either the first tumor or the second tumor) was associated with worse OS compared with lung adenocarcinoma patients. The following factors led to short OS: Stage III patients (HR 1.422; 95% CI 1.120–1.805; \( p = 0.004 \)) and Stage IV patients (HR 1.961; 95% CI 1.563–2.460; \( p < 0.001 \)), LN metastasis (HR 1.540; 95% CI 1.254–2.189; \( p < 0.001 \)), surgery (Yes–No, HR 1.375; 95% CI 1.146–1.649; \( p = 0.001 \); No–No, HR 1.546; 95% CI 1.236–1.934; \( p < 0.001 \)). To more vividly reflect the relationship between independent prognostic factors and OS, Kaplan–Meier method curves were used to show the survival time of different factors (Fig. 1).

Propensity score matching of SPLC and DPLC

In our study, the OS rates of DPLC and SPLC patients were compared by Kaplan – Meier method. Because of the confounding factors, the survival rates between the two groups showed incomprehensible results which the SPLC patients had poor prognosis than DPLC patients (Fig. 2A, \( p < 0.001 \)).

We extracted the 900 DPLC patients with surgery and 900 SPLC patients with surgery which were extracted by the random sampling method from 70,198 SPLC patients to perform PSM analysis. The factors of grade, LN metastasis and stage were included into the PSM analysis. The differences of factors between the two groups were eliminated by PSM (Table supplement 1). Kaplan–Meier method was used to explore the OS of matching data, and the result showed that the DPLC patients were associated with the poor prognosis (Fig. 2B).

Clinical and pathological differences between DPLC and SPLC

The 1419 SPLC patients were extracted from 70,198 SPLC patients by the random sampling method. The pathological differences among the single primary lung tumor, 1st tumor of DPLC and 2nd tumor of DPLC were examined by chi-square test (Table 2). According to Table 2, we could draw a conclusion that the main histologic type of DPLC was lung adenocarcinoma, the main stage of DPLC was stage I. And the distant metastasis of PDLC was rare.

The clinical characteristics were analyzed by the patients' information from Provincial Hospital Affiliated to Shandong University. A total of 173 SPLC patients and 173 MPLC patients were included into chi-square test (Table 3). The proportion of male and female patients with DPLC was approximately the same, but the male SPLC patients made up the majority. Most MPLC patients didn’t smoke and have chronic diseases before the occurrence of cancer.

There results also indicated that the pathogenesis and etiology might be different from MPLC and SPLC.

Prognostic nomogram for DPLC

A total of 1419 patients from SEER database were randomly divided into the training cohort and validation cohort. The training cohort accounted for 70% of the total cases. To predict the survival time of DPLC patients, the patients’ information of training cohort was used to establish a nomogram including aforementioned significant factors (Fig. 3). The prognostic nomogram showed that surgery contributed the most to the OS of DPLC, followed by age, sex, histological type, stage, LN metastasis. The 3- and 5- year OS rates of the prediction results were shown in Fig. 3. The different classification of each factor was assigned a score on the point scale. The total score was calculated by adding all the scores of these factors. The higher score was associated with the worse survival prognosis. The nomogram could be used to predict each patient’s survival time according to their own conditions.

Validity of predictive model’s accuracy

Accuracy of the nomogram was examined by the C index, the ROC curve, the calibration chart. The C index of training cohort was 0.690 (95% CI 0.678–0.702), and the C index of validation cohort was 0.681 (95% CI 0.663–0.699). Both of them had good predictive value. The 3- and 5- year AUC values of the ROC curves were 0.720 and 0.723 on training cohort (Fig. 4A). The 3- and 5- year AUC values of the ROC curves were 0.696 and 0.755 on validation cohort (Figure supplement 1A). Additionally, the calibration plots presented an acceptable and excellent consistency between predicted results and actual observations for 3- and 5- year OS rates in training cohort and validation cohort (Fig. 4B and Figure supplement 1B). We could observe that all calibration curves were closed to ideal the 45° dotted line.

To verify the predictive ability of the nomogram, the DCA was used to compare both the predictive model and the AJCC TNM system. The 3- and 5- year DCA curves of training cohort and validation cohort were shown on Fig. 4C and Figure supplement 1C. All the DCA curves presented that the new model had a more accurate prediction ability than the AJCC TNM system. The IDI values of 3- and 5- year were 7.8% (\( p < 0.001 \)) and 9.2% (\( p < 0.001 \)) on training cohort, and the IDI values of 3- and 5- year...
Fig. 1 Kaplan–Meier method of independent prognostic factors on DPLC
Fig. 2 Kaplan–Meier method of single primary lung cancer and double primary lung cancer. A Kaplan–Meier method of single primary lung cancer and double primary lung cancer without PSM (p < 0.001). B Kaplan–Meier method of single primary lung cancer and double primary lung cancer with PSM (p < 0.022)

Table 2 The analysis of pathological characteristics between SPLC patients and DPLC patients

| Variables | Primary | 1st | p   | 1st | 2nd | p   | Primary | 2nd | p   |
|-----------|---------|-----|-----|-----|-----|-----|---------|-----|-----|
| Laterality|         |     |     |     |     |     |         |     |     |
| Left      | 576     | 637 | 0.021 | 637 | 649 | 0.639 | 576     | 649 | 0.005 |
| Right     | 843     | 782 |     | 782 | 769 |     | 843     | 769 |     |
| Histologic type | |     |     |     |     |     |         |     |     |
| Ad        | 860     | 917 | 0.027 | 917 | 942 | 0.324 | 860     | 942 | 0.001 |
| Sq        | 559     | 502 |     | 502 | 477 |     | 559     | 477 |     |
| Grade     |         |     |     |     |     |     |         |     |     |
| I         | 142     | 222 | <0.001 | 222 | 313 | <0.001 | 142     | 313 | <0.001 |
| II        | 547     | 669 |     | 669 | 664 |     | 547     | 664 |     |
| III       | 718     | 512 |     | 512 | 434 |     | 718     | 434 |     |
| IV        | 12      | 16  |     | 16  | 8   |     | 12      | 8   |     |
| T         |         |     |     |     |     |     |         |     |     |
| T1        | 862     | 1101| <0.001 | 1101| 1184| <0.001 | 862     | 1184| <0.001 |
| T2        | 138     | 104 |     | 104 | 49  |     | 138     | 49  |     |
| T3        | 104     | 96  |     | 96  | 113 |     | 104     | 113 |     |
| T4        | 315     | 118 |     | 118 | 73  |     | 315     | 73  |     |
| N         |         |     |     |     |     |     |         |     |     |
| N0        | 732     | 950 | <0.001 | 950 | 950 | 1     | 732     | 950 | <0.001 |
| N1–3      | 687     | 469 |     | 469 | 469 |     | 687     | 469 |     |
| M         |         |     |     |     |     |     |         |     |     |
| M0        | 993     | 1339| <0.001 | 1339| 1271| <0.001 | 993     | 1271| <0.001 |
| M1        | 426     | 80  |     | 80  | 148 |     | 426     | 148 |     |
| Stage     |         |     |     |     |     |     |         |     |     |
| I         | 487     | 944 | <0.001 | 944 | 957 | <0.001 | 487     | 957 | <0.001 |
| II        | 145     | 169 |     | 169 | 159 |     | 145     | 159 |     |
| III       | 361     | 226 |     | 226 | 155 |     | 361     | 155 |     |
| IV        | 426     | 80  |     | 80  | 148 |     | 426     | 148 |     |
were 5.1% ($p < 0.001$) and 7.4% ($p < 0.001$) on validation cohort. The IDI results were in line with DCA curves.

Discussion

In recent years, a large number of studies focused on SPLC, but there were few studies on DPLC. To date, little is known about the prognostic factors and clinical characteristics of DPLC. An accurate predictive model and well-defined clinical features of DPLC are urgently needed. This study retrospectively analyzed the DPLC patients extracted from SEER database between 2004 and 2015. Univariate and multivariable Cox regression analysis were used to explore the significant clinical factors and these factors were used to establish a prognostic model. Apart from that, the comparison between DPLC and SPLC was analyzed according to the information from SEER database and Provincial Hospital Affiliated to Shandong University.

In this study, we found that age, sex, histological type, stage, LN metastasis, surgery, chemotherapy were significant factors which were associated with the prognosis of DPLC patients. These results were broadly consistent with other studies, but Cong-kuan Song et al. found that race, grade also had an impact on prognosis of DPLC [10].

Surgical resection is a main therapeutic method which benefits the most DPLC patients to prolong their OS [11]. Our study also showed that the survival time of patients with two operations was longer than other patients without surgery. And previous studies showed that radiation and chemotherapy emerged as the effective alternative methods for patients for whom surgery was contraindicated or was not in line with the patient’s wishes [12, 13]. However, we found that only chemotherapy could prolong survival time of DPLP, radiotherapy was not independent diagnostic factor for DPLC patients.

Astonishingly, the OS of SPLC was shorter than DPLC (Fig. 2A). We speculated that the clinical and pathological features between SPLC and DPLC were so different that the confounding factors confounded our judgment to survival time. After including SPLC and DPLC patients with surgery, the key prognostic factors such as grade, LN metastasis, stage were extracted to perform the PSM analysis, considering the interference of confounding factors. And the Kaplan–Meier curve showed reasonable outcome which the OS rate of DPLC was lower after PSM analysis (Fig. 2B). This may be because the second tumor of DPLC patients could not accept lobectomy and could only accept limited resection such as wide wedge resection.

Some previous study proposed a diagnostic criterion which DPLC could be divided into synchronous DPLC and metachronous DPLC according to interval time (6 months) of the 1st and 2nd tumor of DPLC [14]. Martini and Melamed thought the interval time of mMPLC

### Table 3 The analysis of clinical characteristics between SPLC patients and DPLC patients

| Variables             | SPLC     | MPLC     |
|-----------------------|----------|----------|
|                       | $n$      | $%$      | $n$      | $%$      |
| Sex                   |          |          |          |          |
| Male                  | 130      | 75.14    | 88       | 50.87    | $< 0.001$ |
| Female                | 43       | 24.86    | 85       | 49.13    |
| Age                   |          |          |          |          |
| $< 60$                | 82       | 47.40    | 95       | 54.91    | 0.162     |
| $\geq 60$             | 91       | 52.60    | 78       | 45.09    |
| Smoke                 |          |          |          |          |
| Yes                   | 119      | 68.79    | 65       | 37.57    | $< 0.001$ |
| No                    | 54       | 31.21    | 108      | 62.43    |
| Chronic disease       |          |          |          |          |
| Yes                   | 119      | 68.79    | 54       | 31.21    | $< 0.001$ |
| No                    | 54       | 31.21    | 119      | 68.79    |
| Familial-hereditary disease |    |          |          |          |
| Yes                   | 29       | 16.76    | 28       | 16.18    | 0.885     |
| No                    | 144      | 83.24    | 145      | 83.82    |
| Pre-cancer symptoms   |          |          |          |          |
| Yes                   | 118      | 68.21    | 93       | 53.76    | 0.006     |
| No                    | 55       | 31.79    | 80       | 46.24    |
| LN metastasis         |          |          |          |          |
| Yes                   | 68       | 39.31    | 44       | 25.43    | 0.006     |
| No                    | 105      | 60.69    | 129      | 74.57    |
was 2 years based on 50 cases in 1975 [16]. But the American College of Chest Physicians put forward the diagnostic criteria that the interval time was 4 years in 2007 [6]. Our study adopted the point of view which the time was 6 months. But it was confused that the interval time wasn’t significant prognostic factor, when it was...

Fig. 4 Validation of predictive model’s accuracy. Accuracy of the nomogram was examined by the C index, the ROC curve, the calibration chart according to training cohort.
included into Cox regression analysis. Considering the influence of confounding factors, we performed the PSM to eliminate the difference between sDPLC and mDPLC, however, there still remained some bias after PSM. And the Kaplan–Meier curve showed that there were not any differences between their OS rates after PSM \( (p > 0.05) \). Cong-kuan Song et al. also found the results \[10\]. There were two possible reasons for this, firstly, we acknowledged the statistical limitations of the study due to the small sample size, secondly, the diagnostic criteria of interval were unreasonable. There results inspired that the further research could pay an attention to define new interval time which has more prognostic significance according to more data.

The chi-square analysis was used to examine the differences of SPLC and DPLC. We used the data from SEER database to analyze the pathological differences, and used the data from Provincial Hospital Affiliated to Shandong University to analyze the clinical differences (Tables 2 and 3).

The pathological differences between the SPLC and DPLC were mainly reflected in laterality, histologic type, grade and TNM stage. The histologic type of DPLC mainly was adenocarcinoma, and this conclusion was consistent with previous research \[15\]. The differentiated grade of DPLC was mainly II and III, and the number of early-stage patients made up the majority.

The clinical differences between the DPLC and SPLC mainly were sex, smoke, chronic disease, pre-cancer symptoms. In SPLC patient’s cohort, the male patients accounted for 75.14\%, the female patients accounted for 24.86\%. And the 119 patients with smoking habit accounted 68.79\%. In DPLC patient’s cohort, the proportion of sex was about the same, and the majority of patients hadn’t the smoking habit. Most MPLC patients didn’t have chronic diseases and pre-cancer symptoms, which brought difficulties to the timely diagnosis for patients' diseases.

It is noteworthy that there are still no specific treatment guidelines and plan for DPLC. The nomogram was widely used to predict the prognosis for cancer, primarily because of their ability to include statistical predictive factors into a multivariate visualization gram \[7\]. Considering the advantages of nomogram, we included aforementioned independent prognostic factors to establish a nomogram for predicting the OS of patients. The effectiveness was assessed by C index, ROC curve and calibration plot, and the assessment results indicated that the nomogram had the excellent accuracy and predictability. Clinicians used the traditional 8th AJCC TNM system to access the progression of cancer, but the system couldn’t predict the OS of patients and provide clinical guidance. And the assessment ability of the system lacked specificity for cancer subtypes. The nomogram made up for the shortcomings of the 8th AJCC TNM system, because we included clinical factors and therapy methods into the model. DCA curves were used to certify that the predictability of the new model was better than the 8th AJCC TNM system in training and validation cohort (Fig. 4 and Figure supplement 1). The IDI values also proved these results.

Our study has the following advantages. Firstly, this study established a nomogram to predict survival time of DPLC. This model could be used to improve prediction capabilities. Then, this study was the first attempt to performed the chi-square analysis and PSM analysis for comparing the differences of DPLC and SPLC. This work expanded our knowledge of DPLC. Certainly, there are some shortcomings in our research. We did not clearly explore the differences between sDPLC and mDPLC because of the lack of data. And the patients’ number from Provincial Hospital Affiliated to Shandong University and SEER database was so insufficient that the clinical features and pathological characteristics couldn’t be explored clearly. Considering the shortcomings of this study, further prospective analysis should be recommended.

**Conclusion**

The clinical and pathological characteristics of DPLC were different from the SPLC. Therefore, AJCC TNM system was not suitable for prognosis assessment of DPLC patients. And the nomogram based on significant factors could provide accurate and individualized survival predictions for DPLC.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11748-022-01795-6.

**Acknowledgements** We thank all patients, investigators, and institutions involved in this study, especially the SEER database and Provincial Hospital Affiliated to Shandong University for providing platforms and valuable data sets.

**Funding** This study was supported by the Jinan Science and Technology Plan (grant no. 202019058) and the Natural Science Foundation of Shandong Province (grant no. ZR2019MH026).

**Declarations**

**Conflict of interests** The authors declare that they have no competing interests.

**Ethics approval and consent to participate** Not applicable.

**Consent for publication** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. And investigators had obtained informed consent before enrolling partici-
pants in clinical trials. The consent was for publication of their details under the Creative Commons Attribution License 4.0 (such that they will be freely available on the internet).

References

1. Qu J, Wang L, Jiang M, Zhao D, Wang Y, Zhang F, Li J, Zhang X. A review about pembrolizumab in first-line treatment of advanced NSCLC: focus on KEYNOTE studies. Cancer Manag Res. 2020. https://doi.org/10.2147/CMAR.S257188.

2. Wang H, Hou J, Zhang G, Zhang M, Li P, Yan X, Ma Z. Clinical characteristics and prognostic analysis of multiple primary malignant neoplasms in patients with lung cancer. Cancer Gene Ther. 2019;11:26.

3. Surapaneni R, Singh P, Rajagopalan K, Hageboutros A. Stage I lung cancer survivorship: risk of second malignancies and need for individualized care plan. J Thorac Oncol. 2012;7(8):1252–6.

4. Hamaji M, Allen MS, Cassivi SD, Deschamps C, Nichols FC, Wigle DA, Shen KR. Surgical treatment of metachronous second primary lung cancer after complete resection of non-small cell lung cancer. J Thorac Cardiovasc Surg. 2013. https://doi.org/10.1016/j.jtcvs.2012.12.051.

5. Shintani Y, Okami J, Ito H, et al. Clinical features and outcomes of patients with stage I multiple primary lung cancers. Cancer Sci. 2020. https://doi.org/10.1111/cas.14748.

6. Chen TF, Xie CY, Rao BY, Shan SC, Zhang X, Zeng B, Lei YY, Luo HH. Surgical treatment to multiple primary lung cancer patients: a systematic review and meta-analysis. BMC Surg. 2019. https://doi.org/10.1186/s12893-019-0643-0.

7. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol. 2008. https://doi.org/10.1200/JCO.2007.12.9791.

8. Callegaro D, Miceli R, Mariani L, Raut CP, Gronchi A. Soft tissue sarcoma nomograms and their incorporation into practice. Cancer. 2017. https://doi.org/10.1002/cncr.30721.

9. Caulfield S, Menezes G, Marignol L, Poole C. Nomograms are key decision-making tools in prostate cancer radiation therapy. Urol Oncol. 2018. https://doi.org/10.1016/j.urolonc.2018.03.017.

10. Song CK, Guo ZX, Shen XY, Wang YJ, Wang QW, Yu DH, Chen C, Liu XP, Huang JY, Li S, Hu W. Prognostic factors analysis and nomogram construction of dual primary lung cancer: a population study. Biomed Res Int. 2020. https://doi.org/10.1155/2020/7206591.

11. Muranishi Y, Sonobe M, Hamaji M, Kawaguchi A, Hijiya K, Motoyama H, Menju T, Aoyama A, Chen-Yoshikawa TF, Sato T, Date H. Surgery for metachronous second primary lung cancer versus surgery for primary lung cancer: a propensity score-matched comparison of postoperative complications and survival outcomes. Interact Cardiovasc Thorac Surg. 2018. https://doi.org/10.1093/icvts/ivx389.

12. Creach KM, Bradley JD, Mahasittiwat P, Robinson CG. Stereotactic body radiation therapy in the treatment of multiple primary lung cancers. Radiother Oncol. 2012. https://doi.org/10.1016/j.radonc.2011.12.005.

13. Nagasaka M, Gadgeel SM. Role of chemotherapy and targeted therapy in early-stage non-small cell lung cancer. Expert Rev Anticancer Ther. 2018. https://doi.org/10.1080/14737314.2018.1409624.

14. Detterbeck FC, Jones DR, Kernstine KH, Nauheim KS. American college of physicians lung cancer. Special treatment issues. Chest. 2003. https://doi.org/10.1378/chest.123.1_suppl.244S.

15. Shan S, She J, Xue ZQ, Su CX, Ren SX, Wu FY. Clinical characteristics and survival of lung cancer patients associated with multiple primary malignancies. PLoS ONE. 2017. https://doi.org/10.1371/journal.pone.0185485.

16. Martini N, Melamed MR. Multiple primary lung cancers. J Thorac Cardiovasc Surg. 1975;70(4):606.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.