

## Research Article

# Long-Term Survival of Patients with Metastatic Non-Small-Cell Lung Cancer over Five Decades

Jair Bar <sup>1,2</sup> Damien Urban <sup>1,2</sup> Uri Amit,<sup>3</sup> Sarit Appel <sup>3</sup> Amir Onn,<sup>4</sup> Ofer Margalit <sup>1,2</sup>  
Tamar Beller,<sup>1</sup> Teodor Kuznetsov,<sup>1</sup> and Yaacov Lawrence<sup>2,3,5</sup>

<sup>1</sup>Institute of Oncology, Chaim Sheba Medical Center, Tel Hashomer, Ramat Gan 5262000, Israel

<sup>2</sup>Affiliated with Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel

<sup>3</sup>Department of Radiation Oncology, Chaim Sheba Medical Center, Tel Hashomer, Ramat Gan 5262000, Israel

<sup>4</sup>Pulmonology Institute, Chaim Sheba Medical Center, Tel Hashomer, Ramat Gan 5262000, Israel

<sup>5</sup>Department of Radiation Oncology, Sidney Kimmel Medical College at Thomas Jefferson University, Philadelphia, PA, USA

Correspondence should be addressed to Jair Bar; [yair.bar@sheba.health.gov.il](mailto:yair.bar@sheba.health.gov.il)

Received 14 July 2020; Revised 5 December 2020; Accepted 14 December 2020; Published 13 January 2021

Academic Editor: Xiaosheng Wang

Copyright © 2021 Jair Bar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Objective.** Novel therapeutics and supportive care improved outcomes for metastatic non-small-cell lung cancer (mNSCLC) patients. Major advances over the past five decades include the introduction of combination chemotherapy, small molecules targeting mutant proteins, especially EGFR, and more recently immunotherapy. We aim to document real-world long-term survival over the past five decades. **Methods.** Survival statistics were extracted from the Survival, Epidemiology, and End Results (SEER) database for mNSCLC patients during 1973–2015. Two- and five-year survival (2yS and 5yS) were analyzed using Kaplan–Meier and proportional hazard models. **Results.** The study population consisted of 280,655 mNSCLC patients diagnosed during 1973–2015. Longer survival was seen in younger, female, married, Asian/Pacific Islander race, adenocarcinoma, lower grade, more recent diagnosis, higher income, and chemotherapy-treated patients. 2yS increased during the study period from 2.6% to 12.9%, and 5yS increased from 0.7% to 3.2%. 2yS of patients <50 years of age rose from 2.1% to 22.8%, and their 5yS rose from 0.7% to 6.2%. 2yS of adenocarcinoma patients improved from 2.7% to 16.2%, and their improved 5yS from 1.1% to 3.9%. **Conclusions.** Between 1973 and 2015, there was a dramatic improvement in long-term survival, with an approximately five-fold increase in both 2yS and 5yS. Nonetheless, absolute numbers of long-term survivors remained low, with less than 4% living 5 years. This provides a baseline to compare long-term outcomes seen in the current generation of clinical trials.

## 1. Introduction

Lung cancer is the number one cause of cancer-related death worldwide, inflicting about 1.6 million deaths annually [1]. In 2016, 158,080 lung cancer patients died in the USA alone [2]. Lung cancer patients are diagnosed as metastatic disease in 40–50% of the cases [3, 4], with a median survival of less than nine months [3, 5]. The most common type of lung cancer is non-small-cell lung cancer (NSCLC), comprising 85–89% of lung cancers. Significant advances in the care of metastatic NSCLC (mNSCLC) patients over the recent decades include the advent of palliative chemotherapy [5], histology-directed chemotherapy agents [6], improvements

in supportive care [7], more recently targeted agents [8], and immunotherapy [9]. Modern treatments have achieved long-term survival, for instance, EGFR-mutant metastatic lung adenocarcinoma treated with erlotinib or gefitinib has a 14.6% five-year survival, and immunotherapy has achieved 13.4% five-year survival amongst patients with mixed levels of programmed death ligand-1 (PD-L1) expression [10].

Remarkably, metastatic anaplastic lymphoma kinase (ALK-) positive NSCLC patients were reported to have a median overall survival (mOS) of around 4–5 years [11]. The European Society of Medical Oncology (ESMO) Magnitude of Clinical Benefit Scale suggests a treatment that increases the two-year survival of patients who should receive

maximal support from the Oncologist community [12]. However, long-term survival of mNSCLC is rarely reported in studies and is not the focus of most retrospective database analyses.

The critical evaluation of reports of long-term survival rates of patients with novel therapies requires valid comparators; real-world data collected prior to the immunotherapy era can serve this purpose. The Surveillance, Epidemiology, and End Results (SEER) registry has been collecting data on cancer across the United States since 1973. Over time, the area covered has substantially increased, presently including 28% of the U.S. population. The population covered by the SEER registry is comparable to the general U.S. population with regard to measures of income and education, but has a higher proportion of foreign-born persons compared to the general U.S. population [13]. Some previous reports of lung cancer in the SEER database focus on incidence trends [14, 15]. A more recent publication reports also on survival, however in an aggregate manner, including all types of lung cancer (small cell and NSCLC), as well as all stages [16]. Other reports focused on survival and treatment costs of squamous cell NSCLC [17] or older NSCLC patients [18]. A mathematical model predicting survival was suggested, including all stages of lung cancer, and was limited to patients diagnosed in 1998–2001 [19]. The impact of tumor size on the outcome of early stage cancer was analyzed in another report [20]. The influence of treatments on short-term (one year) survival [21] or on median survival [18] of mNSCLC was reported. Other reports focus on the factors impacting chosen treatments [22]. Another study of the SEER database queried the prognostic role of race on lung cancer survival [23], and another examined age [24], both including all stages of disease.

Unlike previous studies, we aimed to focus on advanced NSCLC, the largest group of lung cancer patients, where dramatic changes have occurred recently regarding the standard of care systemic therapy. Accordingly, we wanted to describe the changes in the relevant prognostic parameters and outcome along the 40 years leading to the current era. For this goal, we utilized the SEER database to investigate long-term survival of mNSCLC prior to the immunotherapy era. Trends over time and the impact of potential prognostic factors, including pathologic grade and subtypes, socioeconomic and social parameters, and treatments administered, were studied in a comprehensive manner. We report of one of the largest cohorts of lung cancer patients, focusing on the less well-documented long-term survival. The data presented herein can serve as a basis of comparison when evaluating long-term survivor rates of mNSCLC patients treated with immunotherapy and other novel agents.

## 2. Materials and Methods

The study utilized data from the 18 registries that comprise the SEER database (Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, and Utah Alaska, San Jose-Monterey, Los Angeles, Rural Georgia, Greater California, Kentucky, Louisiana, New Jersey, and Greater Georgia).

Inclusion criteria were mNSCLC diagnosed between 1973 and 2015. Primary lung cancer was identified according to the ICD-0-3 codes: C34.0 (main bronchus), C34.1 (upper lobe, lung), C34.2 (right middle lobe, lung), C34.3 (lower lobe, lung), C34.8 (overlapping lesion of the lung), and C34.9 (lung, NOS). Identified cases were divided into histology groups based on the World Health Organization classification [25]. Exclusion criteria were unknown stage, non-metastatic disease, sarcomas, carcinoid histology, unknown histology, nonspecified malignant histology, rare or unspecified histologies (WHO ICD-O-3 morphological codes [26] 8000, 8001, 8002, 8003, 8004, 8011, 8022, 8030, 8031, 8032, 8033, 8035, 8245, 8490, 8982, 8982, and 9133), or those diagnosed only at autopsy. Data extracted included patient demographics, tumor grade, primary site, treatment category, year of diagnosis, and survival until death or last follow-up as of December 31, 2015. Treatments are reported in the SEER database as ‘chemotherapy’, ‘radiotherapy’, both of these, or none; of note, the SEER database does not differentiate between ‘unknown’ and ‘did not receive’ for these modalities. Use of oral drugs such as tyrosine kinase inhibitors is not registered in the SEER database. Social status was estimated by the median income at 2010 of the county of each patient. Patients’ missing data for a covariate were included in the study, aside from analyses that included that particular covariate.

For the purposes of descriptive (but not survival) analysis, continuous variables (e.g., age, year of diagnosis, and regional income) were converted into categorical variables. Statistical analyses were performed using the Stata statistical package, version IC 11.1 (Stata, College Station, TX). Chi-square tests were used to assess associations between categorical variables. The primary endpoint was survival, defined from the time of initial diagnosis to the date of death, and was calculated using the Kaplan–Meier method. mOS was calculated for each year separately, or for periods, as relevant for the presented analyses. Landmark analyses were conducted regarding two-year survival (2yS) and five-year survival (5yS). Since the current SEER data are updated till 2015, the most recent 5yS data were from patients diagnosed during 2010. To ensure that those who did not survive a full month after diagnosis were included in the analysis, patients coded in the SEER data set as having a survival time of zero were assigned a survival time of half a month. The effects of demographic, pathologic, and treatment variables on survival were tested with a Cox univariate analysis. Multivariate analysis was performed with a Cox proportional hazard model. All  $p$  values were 2 sided, and a  $p < 0.05$  was considered statistically significant.

## 3. Results

**3.1. Patient Characteristics.** A consort diagram detailing the selected population is presented in Supplementary Figure 1. A total of 280,655 subjects with mNSCLC were included in the analysis; the median age was 67 years, and 58% were male. Demographic characteristics changed significantly from 1973 to 2015 (summarized by decades in Supplementary Table 1): the percentage of females increased from

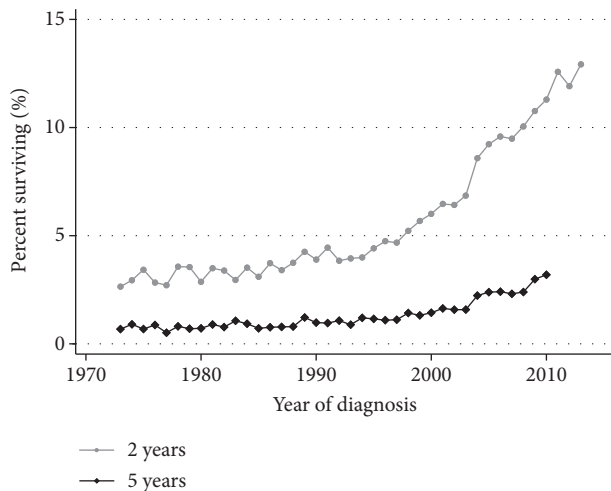


FIGURE 1: Changes in overall survival of metastatic lung cancer patients, 1973–2015. Trends in proportion of patients surviving 2 and 5 years, as a function of year of diagnosis. Both two- and five-year overall survival have improved over time (both  $p < 0.001$ , extended Wilcoxon rank-sum test).

25% in 1973 to 46% in 2015, median age increased from 62 to 68 years, proportion of white subjects decreased (from about 86% to 78%), and proportion of married decreased from 73% to 51% (all statistically significant; supplementary Table 1). The median income in 2010 at the county of residence of each patient decreased throughout the study period, probably reflecting the addition of counties with a lower median income to the SEER database.

**3.2. Changes in Overall Survival.** The mOS for the entire cohort ( $N=280,655$ ) was 4 months (survival data in the SEER database are rounded to full months). Over time, the mOS has improved from two months in 1973 to five months in 2015. Regarding long-term survival, a clear rise in 2yS is noted, increasing from 2.6% in 1973 to 12.9% in 2013 (latest year of which 2yS data can be calculated; Figure 1 and Tables 1 and 2), occurring mostly after the mid-1990s. A more modest increase is seen in the 5yS, from 0.7% in 1973 to 3.2% in 2010 (latest year of which 5yS data can be calculated), also seen mostly after the mid-1990s.

**3.3. Demographic Factors' Impact on Survival.** A clear survival advantage was seen in favor of the younger patients (Tables 1 and 2). For instance, amongst patients of 49 years or below, the 2yS has increased from 2.1% to 22.8% (Figure 2) and the 5yS has increased from 0.7% to 6.2%. Proportionally similar rises can be seen in the older age groups and are significant for the 2yS; however, the numerical improvements in the 5yS are minimal. Regarding sex, improvements in both 2yS and 5yS have been more profound amongst women (Supplementary Figure 2). Married people have a better survival than unmarried, and the difference and the improvement with time are seen mostly in the 2yS (Supplementary Figure 3).

Race was a significant prognostic factor, with black patients demonstrating a better 2yS than white patients and Asian or Pacific Islanders showing better 2yS and 5yS compared to white and black patients (Tables 1 and 2). Asian/Pacific Islanders had the clearest advantage compared to whites. The numbers of American Indian or Alaska Native patients were small in earlier periods, resulting in highly unstable survival rates when comparing different years. Therefore, only data from 2000 onwards are presented for this group of patients (Supplementary Figure 4, left lower panel). Over time, improvement in outcome is seen amongst all ethnic groups in 2yS. However, only for the Asian/Pacific Islanders, the 5yS gets to 5% (5.2% for the 2010 cohort, 95% CI 3.9–6.8). The group of 'unknown/other' race demonstrated better outcome but will not be discussed further due to the small number of patients.

**3.4. Pathologic Factors.** We noted an increase in adenocarcinomas (from 29% in 1973 to 66% in 2015,  $p < 0.001$ ) and a decrease in the proportion with squamous cell lung cancers (from 27.2% to 19.8%,  $p < 0.001$ ). There was no consistent trend of change regarding the proportion of patients with non-other specified (NOS) NSCLC or with SCLC. Regarding outcome, the improvement in 2yS is the most noticeable in the adenocarcinoma subgroup of patients, from 1995 and onward (Figure 3). However, the 5yS remains below 5% for all histologic groups.

Aiming to identify subgroups with better outcome in an exploratory manner, we combined some of the favorable histologic and demographic factors. Females younger than 50 years with adenocarcinoma diagnosed from 2000 onwards had a 5yS of 6.8% (95% CI 6.0–7.8; number at risk at five years: 172).

Pathological-grade data were missing for 58% of the study cohort (these patients were, nonetheless, included in the analysis of other covariates). Among the 121,583 patients with grade data, a significant impact of this factor can be demonstrated, with 5yS of 6.3% for the well-differentiated tumors subgroup considering the entire study cohort (i.e., years of diagnosis: 1973 till 2010; Table 2). In patients with well-differentiated tumors, the 2yS reached 30.5% and 5yS reached 10.6% at the end of the analyzed period (95% CI 7.4–14.4).

**3.5. The Role of Treatment Modalities.** Data about treatments entered in the SEER database include general chemotherapy and radiotherapy categories. The proportion of patients listed as receiving chemotherapy increased from 28% in 1973 to 52% in 2015 (Supplementary Table 1). As can be seen in Figure 4, the improvement in 2yS is significant in patients that received either chemotherapy or combined chemotherapy and radiotherapy. 5yS also improved, but at best reaches 4.6% (95% CI 4.1–5.1). The groups that received only radiotherapy had similar low outcomes as those receiving no therapy. We were unable to demonstrate that any particular histologic subtype especially benefited from chemotherapy.

TABLE 1: Univariate and multivariate analysis of variables associated with overall survival in metastatic lung cancer patients, 1973–2015, Cox model. The multivariate analysis incorporated following variables: age, sex, race, marital status, histology, grade, year of diagnosis, treatment modality (of note regarding treatment, for each modality the database does not differentiate between ‘treatment not given’ and ‘unknown whether given’), and income (relating to county median income, 2010, in 10,000\$). Patients with missing data for any one of these covariates were excluded from the multivariate analysis.

| Variable <sup>a</sup>    | Univariate analysis |                |                          | Multivariate analysis |                |                |
|--------------------------|---------------------|----------------|--------------------------|-----------------------|----------------|----------------|
|                          | HR                  | 95% C.I.       | <i>p</i> value           | HR                    | 95% C.I.       | <i>p</i> value |
| Age                      | 1.013               | (1.013, 1.013) | <0.001                   | 1.007                 | (1.006, 1.008) | <0.001         |
| Sex (female vs. male)    | 0.852               | (0.839, 0.852) | <0.001                   | 0.840                 | (0.825, 0.849) | <0.001         |
| <i>Race</i>              |                     |                |                          |                       |                |                |
| White                    |                     |                | 1.000 (comparator group) |                       |                |                |
| Black                    | 0.989               | (0.978, 1.001) | 0.063                    | 0.979                 | (0.960, 0.999) | 0.045          |
| Am. Indian/Alaska        | 1.011               | (0.952, 1.074) | NS (0.72)                | 1.143                 | (0.986, 1.325) | 0.076          |
| Asian/Pacific Islander   | 0.753               | (0.742, 0.765) | <0.001                   | 0.832                 | (0.808, 0.856) | <0.001         |
| Unknown/other            | 0.602               | (0.524, 0.693) | <0.001                   | 0.638                 | (0.475, 0.858) | 0.003          |
| Married                  | 0.880               | (0.873, 0.887) | <0.001                   | 0.906                 | (0.893, 0.919) | <0.001         |
| <i>Histology</i>         |                     |                |                          |                       |                |                |
| Lung carcinoma NOS       |                     |                | 1.000 (comparator group) |                       |                |                |
| Adenocarcinoma           | 0.736               | (0.729, 0.742) | <0.001                   | 0.860                 | (0.844, 0.875) | <0.001         |
| Squamous cell ca.        | 0.866               | (0.856, 0.875) | <0.001                   | 0.910                 | (0.892, 0.929) | <0.001         |
| Small cell carcinoma     | 0.858               | (0.849, 0.866) | <0.001                   | 1.029                 | (1.008, 1.05)  | 0.007          |
| <i>Grade</i>             |                     |                |                          |                       |                |                |
| Well diff.               |                     |                | 1.000 (comparator group) |                       |                |                |
| Moderately diff.         | 1.173               | (1.138, 1.209) | <0.001                   | 1.190                 | (1.149, 1.233) | <0.001         |
| Poorly diff.             | 1.512               | (1.470, 1.554) | <0.001                   | 1.484                 | (1.436, 1.535) | <0.001         |
| Undifferentiated         | 1.781               | (1.725, 1.840) | <0.001                   | 1.484                 | (1.425, 1.545) | <0.001         |
| <i>Year of diagnosis</i> |                     |                |                          |                       |                |                |
| 1973–9                   |                     |                | 1.000 (comparator group) |                       |                |                |
| 1980–9                   | 0.948               | (0.928, 0.968) | <0.001                   | 0.952                 | (0.918, 0.987) | 0.008          |
| 1990–9                   | 0.906               | (0.889, 0.924) | <0.001                   | 0.935                 | (0.903, 0.969) | <0.001         |
| 2000–9                   | 0.773               | (0.759, 0.787) | <0.001                   | 0.821                 | (0.793, 0.850) | <0.001         |
| 2010–5                   | 0.663               | (0.651, 0.676) | <0.001                   | 0.739                 | (0.712, 0.766) | <0.001         |
| <i>Treatment</i>         |                     |                |                          |                       |                |                |
| No treatment             |                     |                | 1.00 (comparator group)  |                       |                |                |
| Chemotherapy only        | 0.423               | (0.418, 0.428) | <0.001                   | 0.488                 | (0.478, 0.498) | <0.001         |
| Radiotherapy only        | 0.766               | (0.759, 0.774) | <0.001                   | 0.772                 | (0.758, 0.787) | <0.001         |
| Chemo + radio            | 0.445               | (0.440, 0.449) | <0.001                   | 0.517                 | (0.507, 0.527) | <0.001         |
| Income <sup>b</sup>      | 0.977               | (0.974, 0.980) | <0.001                   | 0.979                 | (0.974, 0.985) | <0.001         |

HR: hazard ratio. CI: confidence interval. Am: American. NOS: non-other specified. Ca: carcinoma. Diff: differentiated. Chemo: chemotherapy. Radio: radiotherapy. <sup>a</sup>Missing values: grade: 207,149 patients. Marital status: 12,727 patients. Treatment: 10,881 patients. Age: 3 patients. Income: 73,156 patients. <sup>b</sup>2010 yearly median county income (per 10,000 dollars).

## 4. Discussion

We have described the 2yS and 5yS of lung cancer patients over a period of 43 years, with improvement seen mostly in the later fifteen years. Better outcomes are most pronounced among younger, married patients (Figure 2 and Supplementary Figure 3), in those bearing adenocarcinoma histology (Figure 3) and lower-grade tumors. Better outcome is seen in patients that have received chemotherapy (Figure 4). Race impacted survival, with a generally better outcome for nonwhite patients. The latest years analyzed here were 2013 regarding 2yS and 2010 regarding 5yS, before immunotherapy entered the arena of lung cancer treatments. Another large study from the preimmunotherapy period, of the International Association for the Study of Lung Cancer (IASLC), summarized world-wide data of more than 90,000 lung cancer patients diagnosed between 1999 and 2010. In that report, 2yS of 23% is reported for stage IVA and 6% for

stage IVB and 5yS of 10% of stage IVA and 0% for stage IVB (clinical stage according the 8<sup>th</sup> staging edition) [27]. The distinction between stage IVA and IVB is not available within the SEER database. However, our general conclusion that even among mNSCLC patients, specific subgroups have demonstrated nonnegligible long-term survival during the preimmunotherapy period is supported by the IASLC study. Recently, Howlader et al. linked death certificate data to the SEER database in order to study changes in NSCLC incidence and mortality over the period 2001–2016, and their study was not confined to metastatic disease. Similar to our findings, they demonstrated marked improvements in 2-year mortality amongst both men and women [28].

The subgroup that stands out with a long 5yS is the younger patients, with age being a significant prognostic factor, as reported previously [29, 30]. Among the younger patients, the improvement in outcome over the years is the most striking. Multiple factors probably contribute to this



TABLE 2: Percentage survival at two and five years by prognostic factors.

|                               | 2yOS |             | 5yOS |            |
|-------------------------------|------|-------------|------|------------|
|                               | %    | 95% CI      | %    | 95% CI     |
| Entire cohort                 | 8.0  | (7.9–8.2)   | 2.1  | (2.0–2.1)  |
| <i>Age</i>                    |      |             |      |            |
| <67                           | 9.5  | (9.3–9.6)   | 2.7  | (2.6–2.7)  |
| ≥67                           | 6.6  | (6.5–6.7)   | 1.5  | (1.5–1.6)  |
| <i>Sex</i>                    |      |             |      |            |
| Male                          | 6.3  | (6.1–6.4)   | 1.6  | (1.5–1.7)  |
| Female                        | 10.5 | (10.3–10.7) | 2.8  | (2.7–2.9)  |
| <i>Race</i>                   |      |             |      |            |
| White                         | 7.5  | (7.4–7.7)   | 2    | (1.9–2.0)  |
| Black                         | 7.2  | (6.9–7.5)   | 1.8  | (1.7–2.0)  |
| American Indian/Alaska native | 8.6  | (7.0–10.4)  | 1.5  | (0.8–2.6)  |
| Asian/Pacific Islander        | 15.1 | (14.5–15.6) | 4.1  | (3.7–4.4)  |
| Unk/oth                       | 23.1 | (17.6–29.1) | 15.1 | (9.8–21.5) |
| <i>Marital status</i>         |      |             |      |            |
| Single                        | 7.0  | (6.9–7.2)   | 1.8  | (1.7–1.9)  |
| Married                       | 8.7  | (8.6–8.9)   | 2.3  | (2.2–2.4)  |
| <i>Histology</i>              |      |             |      |            |
| Adenocarcinoma                | 10.8 | (10.6–11)   | 2.7  | (2.6–2.8)  |
| Lung carcinoma NOS            | 5.1  | (5.0–5.3)   | 1.3  | (1.3–1.4)  |
| Squamous cell carcinoma       | 6.2  | (6.0–6.4)   | 1.8  | (1.7–1.9)  |
| Small-cell carcinoma          | 4.4  | (4.3–4.6)   | 1.3  | (1.2–1.4)  |
| <i>Grade</i>                  |      |             |      |            |
| Well differentiated           | 18.6 | (17.6–19.7) | 6.3  | (5.6–7.0)  |
| Moderately differentiated     | 13.2 | (12.8–13.7) | 3.8  | (3.5–4.1)  |
| Poorly differentiated         | 7.1  | (6.9–7.3)   | 2.0  | (1.9–2.1)  |
| Undifferentiated              | 4.5  | (4.1–4.8)   | 1.3  | (1.1–1.5)  |
| <i>Year of diagnosis</i>      |      |             |      |            |
| 1973–79                       | 3.1  | (2.9–3.4)   | 0.7  | (0.6–0.9)  |
| 1980–89                       | 3.5  | (3.3–3.7)   | 0.9  | (0.8–1.0)  |
| 1990–99                       | 4.5  | (4.3–4.7)   | 1.1  | (1.0–1.2)  |
| 2000–09                       | 8.4  | (8.3–8.6)   | 2.1  | (2.0–2.2)  |
| 2010–15                       | 12.6 | (12.3–12.9) | 3.7  | (3.5–3.9)  |
| <i>Treatment</i>              |      |             |      |            |
| None                          | 4.5  | (4.3–4.6)   | 1.6  | (1.5–1.7)  |
| Chemotherapy only             | 14.9 | (14.6–15.2) | 3.3  | (3.1–3.5)  |
| Radiotherapy (RT) only        | 4.0  | (3.9–4.2)   | 1.0  | (1.0–1.1)  |
| Chemotherapy + RT             | 11.7 | (11.5–12)   | 2.9  | (2.8–3.1)  |
| <i>Income</i>                 |      |             |      |            |
| <median <sup>a</sup>          | 7.2  | (7.0–7.4)   | 1.9  | (1.8–2.0)  |
| ≥median <sup>a</sup>          | 8.5  | (8.4–8.7)   | 2.2  | (2.1–2.3)  |

2yOS: two-year overall survival. 5yOS: five-year overall survival. CI: confidence interval. NOS: non-other specified. Ca: carcinoma. <sup>a</sup>Annual county median in 2010 of \$52,595.

phenomenon, including better performance status, less comorbidities, and a higher rate of targetable mutations. The improvement in the younger patients is seen starting from the mid-90s, similar to the rest of our cohort, prior to the addition of the targeted agents to the treatment options.

Race is well known to correlate with cancer incidence and outcome [19]. Regarding 5yS, black patients had a similar outcome to white patients in our data, as reported in other publications [29]. A markedly better outcome was found in our study for Asian or Pacific Islanders, as reported

earlier [31]. The group of ‘Asian or Pacific Islanders’ has been reported to have a relatively low incidence of lung cancer [32], comprising 6.7% of all lung cancer patients in our cohort. This is a complex group consisting, among others, of patients originated from eastern Asia, possibly including a higher proportion of EGFR mutation-positive patients. However, as discussed for the younger patients and as seen for the entire cohort, the improvement can be seen prior to the emergence of EGFR inhibitors or other targeted agents. Our findings resonate with other observations suggesting better outcome of Asian lung cancer patients compared to patients from the Western Hemisphere [33].

A recurrent observation is the prognostic value of socioeconomic status [34, 35] and, more interestingly, of marital status. Marital status has been found to predict lung cancer survival for NSCLC [35], as well as SCLC [36], although not in all reports [29]. Being married correlates with better survival for cancer patients in general [37], with diverse possible explanations, including better social support systems for married patients and better socioeconomic status, as well as selection bias; individuals who marry are more likely to be healthier than those that do not marry [38]. The impact of being married, as well as the significant impact of the patients’ social status, stresses the importance of providing comprehensive support to cancer patients, above and beyond the provision of the medical care.

A significant and strong prognostic factor in our cohort is the pathologic grade of the tumor, as reported previously in smaller data sets [30, 39, 40]. Tumor grade was available only for 43% of the study cohort, thus reducing the validity of this observation. Lack of clear definitions for pathologic grading of lung cancer impedes the inclusion of this important factor in a standardized manner in pathologic reports and in the staging system. It should be noted that only 5% of the patients in our cohort had well-differentiated tumors, and even moderately differentiated tumors constituted only 20.6% of the cohort. However, the better differentiated tumors have a markedly better prognosis. We, therefore, suggest that grade should be clearly identified in routine pathologic assessments of lung cancer specimens, as well as in clinical trials.

The reasons for the OS improvement seen along the years are most likely multifactorial. Conceivably, some of the improvement are the result of improved systemic treatments. This is supported by the correlation between chemotherapy treatments and improved outcome along the years (Figure 4). Clearly a selection bias exists in this analysis, as chemotherapy administration is usually limited to patients with relatively good performance status, a recognized prognostic factor by itself [2]. Regarding NSCLC, the use of chemotherapy vs. supportive care alone gradually gained support during the 1990s [15], which is also the period when the survival curves start to rise. Tailoring of chemotherapy drugs to tumor histology [6] and addition of anti-VEGF for nonsquamous NSCLC [41] are the major changes in chemotherapy choices that entered practice during 2000–2010; each of these had been demonstrated to improve median survival of clinical trial patients.

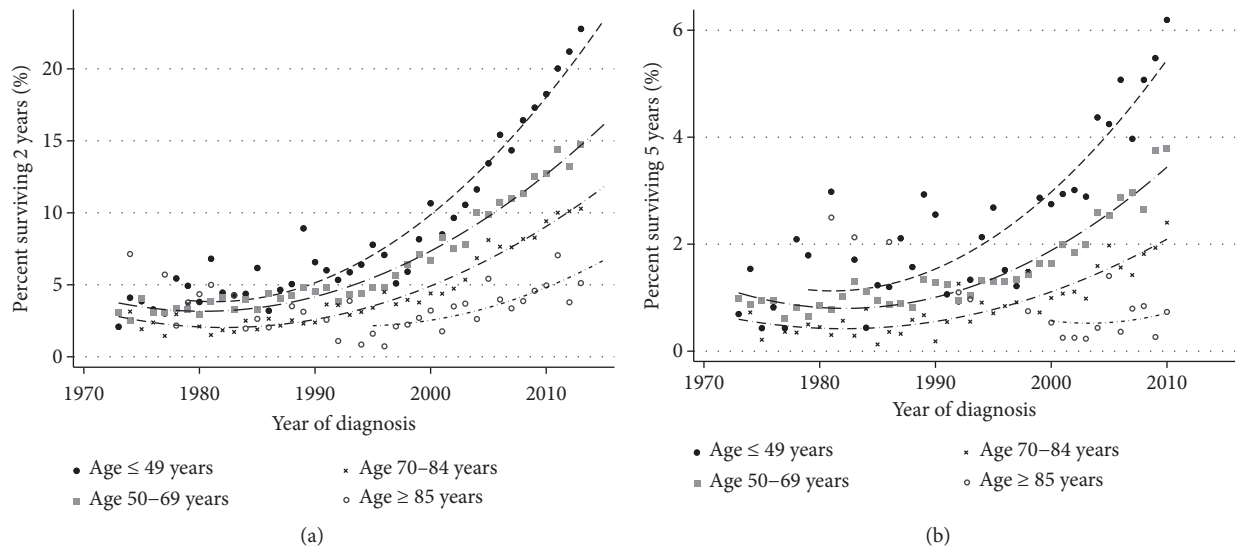


FIGURE 2: (a) Changes in the proportion of metastatic lung cancer patients surviving two years as a function of age and year of diagnosis. The symbols represent data for a specific year, and the lines indicate a trend line. (b) Changes in the proportion of metastatic lung cancer patients surviving five years as a function of age and year of diagnosis. The symbols represent data for a specific year, and the lines indicate a trend line.

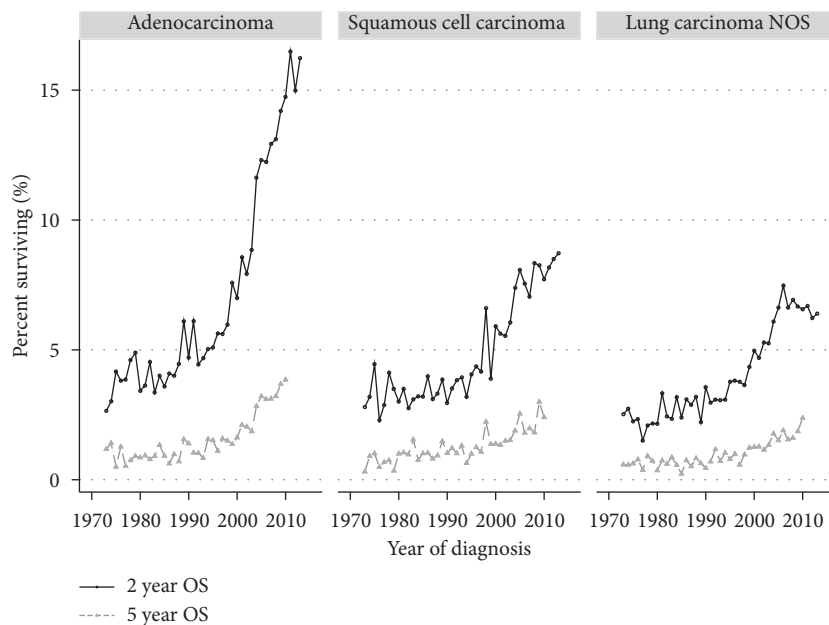


FIGURE 3: Trends in two-year and five-year survival 1973–2015 in different histologic subgroups. Yearly data are presented for each group.

The next breakthrough in the care of NSCLC was targeted agents, initially the drugs targeting the EGFR receptor. Since the entrance of targeted agents, survival has extended for patients harboring the relevant genetic aberrations. Young, female, and adenocarcinoma are the groups where most of the improvement was seen in our cohort, and these are also the characteristics of patients harboring most of the driver mutations [42, 43]. The predictive power of EGFR mutations (representing 17% of the adenocarcinoma patients [44]) was discovered in 2004

[45, 46]. However, the EGFR mutation testing and appropriate therapy entered practice only in 2009 [4], thus, may be, affecting our latest 2yS data and minimally the most recent 5yS data. ALK rearrangement occurs in only 3–5% of NSCLC [47] and, hence, unlikely to impact the survival results of our study cohort. Importantly, the identification of tumors bearing ALK translocations as candidates for ALK inhibitor therapy occurred in 2010 [48], and the FDA approval of crizotinib for this indication occurred in 2011.

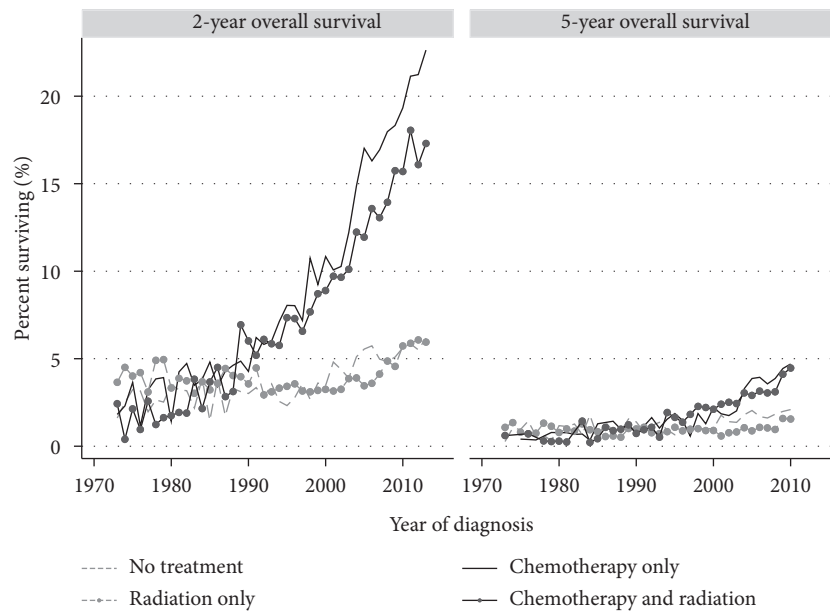


FIGURE 4: Proportion of metastatic lung cancer patients surviving two or five years according to treatment modality. Yearly data are presented for each subgroup, separated by the designated treatment modalities.

An additional potential explanation for improved mOS along the years is an overall more aggressive, less nihilistic approach to metastatic lung cancer, with a higher rate of patients receiving chemotherapy. A special example of aggressive therapy is the approach to oligometastatic disease [49, 50] with the use of locally ablative strategies, whether surgical, radiosurgical, or other. However, it should be noted that, when looking at the entire population, no correlation was found in our study between the use of radiotherapy and improved survival. Another relevant and related improvement that occurred along the years is the development of supportive care, recently demonstrated to prolong survival of lung cancer patients [7, 51]. As a result of improved supportive care, more aggressive treatments can be administered to less robust patients, potentially improving survival of some of these patients.

The increase of mOS might be related also to Will Roger's phenomenon, i.e., the increased likelihood with time to correctly classify minimally metastatic patients, whereas in the past such tumors may have been understaged. Such improved staging would include metastatic patients with a low burden of disease within the metastatic patient cohort, thus improving the median survival of the cohort. The increased use of CT-PET, mediastinoscopy, and brain MRI for staging of patients [52] is likely to have an impact on the overall results of this study.

The strengths of our study include the large size of the database and the inclusion of patients excluded from clinical trials due to comorbidities and/or poor performance status. Furthermore, this is a well-validated and reliable data set with long-term follow-up [13, 53]. The shortcomings of our study include its retrospective nature, lack of basic clinical data such as performance status and weight loss, lack of data regarding molecular subtypes (especially EGFR mutation status and other driver mutations), and lack of details

regarding systemic treatments administered. In addition, only a surrogate for socioeconomic status was captured, probably misclassifying many patients regarding this important prognostic parameter. Methods of staging have changed along the years as noted earlier, and these data are not captured in our data. As for any retrospective study, we cannot assign causative roles to the prognostic factors we have identified.

## 5. Conclusions

We have demonstrated substantial improvements over the last decades in the long-term survival of patients with metastatic lung cancer, mostly since the mid-90s. Most significant improvements were seen in younger, married patients, females, adenocarcinoma, low-grade tumors, and those receiving chemotherapy. While 2yS has bypassed 10% in recent years, 5yS is still in the range of 3–4%, stressing the need for further progress. Importantly, even prior to the era of immunotherapy and mostly before the emergence of targeted agents, subgroupsm NSCLC patients had long-term survival. The characteristics of these patients point to important prognostic factors, details of which should be collected and reported in current clinical trials.

## Data Availability

The raw data are freely available from the website of the NCI's Surveillance, Epidemiology, and End Results (SEER) Program <https://seer.cancer.gov/mortality/>.

## Ethical Approval

This study was exempt from ethics approval since it is based upon a publicly available anonymized database. A 'Data-Use Agreement' with SEER was signed.

## Conflicts of Interest

JB reports grants and personal fees from Merck Sharp and Dohme (MSD), Abbvie, AstraZeneca, Pfizer, Takeda, and Roche, grants from Bristol Myers Squibb (BMS), and personal fees from Boehringer Ingelheim (BI), VBL, Novartis, Bayer, and Lilly. AO is a consultant and member of the advisory board to Astra Zeneca, MSD, Roche, and Boehringer Ingelheim. DU reports personal fees from Merck Sharp and Dohme, Bristol Myers Squibb, Boehringer Ingelheim, Roche, Takeda, and AstraZeneca. YL reports research funding from Karyopharm Therapeutics, Checkmate Pharmaceuticals, Bristol-Myers Squibb, and pending from Merck Serono and honoraria/consultancy fees from Bristol-Myers Squibb, Clinigen Group and Roche Genetech.

## Supplementary Materials

Figure S1: consort diagram of patients participating in this study. SCLC: small-cell lung cancer; Figure S2: trends in two-year and five-year survival 1973–2015 in males and females. Yearly data are presented for each group; Figure S3: proportion of surviving patients at different time points, stratified by marital status. Yearly data are presented for each group; Figure S4: proportion of surviving patients at different time points, stratified by the ethnic group. Yearly data are presented for each group. The numbers of American Indian or Alaska Natives patients were small in earlier periods; thus, only data from 2000 onwards are presented for this group of patients; Table S1: patients characteristics of the entire study population and by decades (based on the year of diagnosis). (*Supplementary Materials*)

## References

- [1] J. Ferlay, I. I. Soerjomataram, R. Dikshit et al., “Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012,” *International Journal of Cancer*, vol. 5, 2012.
- [2] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2017,” *CA: A Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7–30, 2017.
- [3] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2016,” *CA: A Cancer Journal for Clinicians*, vol. 66, no. 1, pp. 7–30, 2016.
- [4] R. Meza, C. Meernik, J. Jeon, and M. L. Cote, “Lung cancer incidence trends by gender, race and histology in the United States, 1973–2010,” *PLoS ONE*, vol. 10, no. 3, 2015.
- [5] NSCLC Meta-Analyses Collaborative Group, “Chemotherapy in addition to supportive care improves survival in advanced non-small-cell lung cancer: a systematic review and meta-analysis of individual patient data from 16 randomized controlled trials,” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, vol. 26, no. 28, pp. 4617–4625, 2008.
- [6] C. Allemani, T. Matsuda, V. Di Carlo et al., “Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries,” *The Lancet*, vol. 391, 2018.
- [7] J. S. Temel, J. A. Greer, A. Muzikansky et al., “Early palliative care for patients with metastatic non-small-cell lung cancer,” *New England Journal of Medicine*, vol. 363, no. 8, pp. 733–742, 2010.
- [8] W. Pao and N. Girard, “New driver mutations in non-small-cell lung cancer,” *The Lancet Oncology*, vol. 12, no. 2, pp. 175–180, 2011.
- [9] P. Sharma, K. Wagner, J. D. Wolchok, and J. P. Allison, “Novel cancer immunotherapy agents with survival benefit: recent successes and next steps,” *Nature Reviews Cancer*, vol. 11, 2011.
- [10] S. Gettinger, L. Horn, D. Jackman et al., “Five-year follow-up of nivolumab in previously treated advanced non-small-cell lung cancer: results from the ca209-003 study,” *Journal of Clinical Oncology*, vol. 36, no. 17, pp. 1675–1684, 2017.
- [11] S. Gettinger, H. Borghaei, J. Brahmer et al., “OA14 04 five-year outcomes from the randomized, phase 3 trials CheckMate 017/057: nivolumab vs docetaxel in previously treated NSCLC,” *Journal of Thoracic Oncology*, vol. 14, no. 10, pp. S244–S245, 2019.
- [12] R. S. Herbst, I. Monnet, S. Novello et al., “LBA63 Long-term survival in patients (pts) with advanced NSCLC in the KEYNOTE-010 study overall and in pts who completed two years of pembrolizumab (pembro),” *Annals of Oncology*, vol. 29, no. suppl\_8, 2018.
- [13] J. Lin, D. Rodriguez–Abreu, A. G. Robinson et al., “Updated analysis of Keynote-024: pembrolizumab versus platinum-based chemotherapy for advanced non-small-cell lung cancer with PD-L1 tumor proportion score of 50% or greater,” *Journal of Clinical Oncology*, vol. 37, no. 7, pp. 537–546, 2019.
- [14] B. J. Solomon, D. W. Kim, Y. L. Wu et al., “Final overall survival analysis from a study comparing first-line crizotinib versus chemotherapy in alk-mutation-positive non-small-cell lung cancer,” *Journal of Clinical Oncology*, vol. 36, 2018.
- [15] N. I. Cherny, U. Dafni, J. Bogaerts et al., “ESMO-magnitude of clinical benefit Scale version 1.1,” *Annals of Oncology*, vol. 28, 2017.
- [16] NIH, *Overview of the SEER Program*, NIH, Washington, DC, USA, 2018, <https://seer.cancer.gov/about/overview.html>.
- [17] D. R. Lewis, D. P. Check, N. E. Caporaso, W. D. Travis, and S. S. Devesa, “US lung cancer trends by histologic type,” *Cancer*, vol. 120, 2014.
- [18] T. Lu, X. Yang, Y. Huang et al., “Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades,” *Cancer Management and Research*, vol. 11, 2019.
- [19] K. L. Davis, R. K. Goyal, S. L. Able, J. Brown, L. Li, and J. A. Kaye, “Real-world treatment patterns and costs in a US Medicare population with metastatic squamous non-small cell lung cancer,” *Lung Cancer*, vol. 87, no. 2, pp. 176–185, 2015.
- [20] M. A. Bittoni, A. Arunachalam, H. Li et al., “Real-world treatment patterns, overall survival, and occurrence and costs of adverse events associated with first-line therapies for medicare patients 65 Years and older with advanced non-small-cell lung cancer: a retrospective study,” *Clinical Lung Cancer*, vol. 19, 2018.
- [21] A. Agrawal, S. Misra, R. Narayanan, L. Polepeddi, and A. Choudhary, “A lung cancer outcome calculator using ensemble data mining on SEER data,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Beijing, China, August 2011.
- [22] J. Zhang, K. A. Gold, H. Y. Lin et al., “Relationship between tumor size and survival in non-small-cell lung cancer (NSCLC): an analysis of the surveillance, epidemiology, and



- end results (SEER) registry," *Journal of Thoracic Oncology*, vol. 10, 2015.
- [23] T. K. Owonikoko, C. Ragin, Z. Chen et al., "Real-world effectiveness of systemic agents approved for advanced non-small cell lung cancer: a SEER-medicare analysis," *The Oncologist*, vol. 18, 2013.
- [24] L. R. Shugarman, K. MacK, M. E. S. Sorbero et al., "Race and sex differences in the receipt of timely and appropriate lung cancer treatment," *Medical Care*, vol. 47, no. 7, pp. 774–781, 2009.
- [25] S. M. Gadgeel, R. K. Severson, Y. Kau et al., "Impact of race in lung cancer," *Chest*, vol. 120, 2001.
- [26] T. K. Owonikoko, C. C. Ragin, C. P. Belani et al., "Lung cancer in elderly patients: an analysis of the surveillance, epidemiology, and end results database," *Journal of Clinical Oncology*, vol. 24, 2007.
- [27] W. D. Travis, "WHO classification of tumours of the lung, pleura, thymus and heart," *International Agency for Research on Cancer*, vol. 10, 2015, <https://books.google.co.il/books?id=nKO1rQEACAAJ>.
- [28] P. Goldstraw, K. Chansky, J. Crowley et al., "The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM Classification for lung cancer," *Journal of Thoracic Oncology*, vol. 11, 2016.
- [29] A. Yoshizawa, N. Motoi, G. J. Riely et al., "Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases," *Modern Pathology*, vol. 24, 2011.
- [30] J. Feuer, C. A. Carter, K. A. McGlynn et al., "A prognostic model to predict mortality among non-small-cell lung cancer patients in the U.S. Military health system," *Journal of Thoracic Oncology*, vol. 10, 2015.
- [31] B. K. Edwards, A. M. Noone, A. B. Mariotto et al., "Annual report to the nation on the status of cancer, 1975–2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer," *Cancer*, vol. 120, 2014.
- [32] J. Putila, S. C. Remick, and N. L. Guo, "Combining clinical, pathological, and demographic factors refines prognosis of lung cancer: a population-based study," *PLoS ONE*, vol. 6, 2011.
- [33] L. A. Torre, R. L. Siegel, and A. Jemal, "Lung cancer statistics," *Advances in Experimental Medicine and Biology*, vol. 893, 2016.
- [34] W. Zhou and D. C. Christiani, "East meets west: ethnic differences in epidemiology and clinical behaviors of lung cancer between east Asians and Caucasians," *Chinese Journal of Cancer*, vol. 30, 2011.
- [35] T. Y. D. Cheng, S. M. Cramb, P. D. Baade, D. R. Youlten, C. Nwogu, and M. E. Reid, "The international epidemiology of lung cancer: latest trends, disparities, and tumor characteristics," *Journal of Thoracic Oncology*, vol. 11, 2016.
- [36] S. H. Ignatius Ou, J. A. Zell, A. Ziogas, and H. Anton-Culver, "Low socioeconomic status is a poor prognostic factor for survival in stage I nonsmall cell lung cancer and is independent of surgical treatment, race, and marital status," *Cancer*, vol. 112, 2008.
- [37] S. H. I. Ou, A. Ziogas, and J. A. Zell, "Prognostic factors for survival in extensive stage small cell lung cancer (ED-SCLC): the importance of smoking history, socioeconomic and marital statuses, and ethnicity," *Journal of Thoracic Oncology*, vol. 4, 2009.
- [38] S. L. Gomez, S. Hurley, A. J. Canchola et al., "Effects of marital status and economic resources on survival after cancer: a population-based Study," *Cancer*, vol. 122, 2016.
- [39] M. S. Rendall, M. M. Weden, M. M. Favreault, and H. Waldron, "The protective effect of marriage for survival: a review and update," *Demography*, vol. 48, 2011.
- [40] K. H. Yu, C. Zhang, G. J. Berry et al., "Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features," *Nature Communications*, vol. 7, 2016.
- [41] D. H. Harpole, J. E. Herndon, W. G. Wolfe, J. D. Iglehart, and J. R. Marks, "A prognostic model of recurrence and death in stage I non-small cell lung cancer utilizing presentation, histopathology, and oncoprotein expression," *Cancer Research*, vol. 55, 1995.
- [42] K. S. Albain, J. J. Crowley, M. LeBlanc, and R. B. Livingston, "Survival determinants in extensive-stage non-small-cell lung cancer: the southwest oncology group experience," *Journal of Clinical Oncology*, vol. 9, 1991.
- [43] Non-small Cell Lung Cancer Collaborative Group, "Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials," *BMJ*, vol. 311, no. 7010, pp. 899–909, 1995.
- [44] G. V. Scagliotti, P. Parikh, J. von Pawel et al., "Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer," *Journal of Clinical Oncology*, vol. 26, no. 21, pp. 3543–3551, 2008.
- [45] A. Sandler, R. Gray, M. C. Perry et al., "Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer," *New England Journal of Medicine*, vol. 355, no. 24, pp. 2542–2550, 2006, [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=17167137](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17167137).
- [46] T. Kosaka, Y. Yatabe, H. Endoh, H. Kuwano, T. Takahashi, and T. Mitsudomi, "Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications," *Cancer Research*, vol. 64, 2004.
- [47] S. J. Rodig, M. Mino-Kenudson, S. Dacic et al., "Unique clinicopathologic features characterize ALK-rearranged lung adenocarcinoma in the western population," *Clinical Cancer Research*, vol. 15, 2009.
- [48] M. G. Kris, B. E. Johnson, D. J. Kwiatkowski et al., "Identification of driver mutations in tumor specimens from 1,000 patients with lung adenocarcinoma: the NCI's Lung Cancer Mutation Consortium (LCMC)," *Journal of Clinical Oncology*, vol. 29, no. 18, Article ID CRA7506, 2011.
- [49] J. G. Paez, P. A. Janne, J. C. Lee et al., "EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy," *Science*, vol. 304, no. 5676, pp. 1497–1500, 2004, [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=15118125](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15118125).
- [50] T. J. Lynch, D. W. Bell, R. Sordella et al., "Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib," *The New England Journal of Medicine*, vol. 350, no. 21, pp. 2129–2139, 2004.
- [51] C. G. Azzoli, S. Baker, S. Temin et al., "American society of clinical oncology clinical practice guideline update on chemotherapy for stage IV non-small-cell lung cancer," *Journal of Clinical Oncology*, vol. 13, 2009.
- [52] D. R. Camidge, S. A. Kono, A. Flacco et al., "Optimizing the detection of lung cancer patients harboring anaplastic

- lymphoma kinase (ALK) gene rearrangements potentially suitable for ALK inhibitor treatment,” *Clinical Cancer Research*, vol. 16, no. 22, pp. 5581–5590, 2010.
- [53] D. E. Gerber and J. D. Minna, “ALK inhibition for non-small cell lung cancer: from discovery to therapy in record time,” *Cancer Cell*, vol. 18, 2010.
- [54] D. R. Gomez, G. R. Blumenschein, J. J. Lee et al., “Local consolidative therapy versus maintenance therapy or observation for patients with oligometastatic non-small-cell lung cancer without progression after first-line systemic therapy: a multicentre, randomised, controlled, phase 2 study,” *The Lancet Oncology*, vol. 37, 2016.
- [55] C. Collen, N. Christian, D. Schallier et al., “Phase II study of stereotactic body radiotherapy to primary tumor and metastatic locations in oligometastatic nonsmall-cell lung cancer patients,” *Annals of Oncology*, vol. 25, 2014.
- [56] G. A. Silvestri, A. V. Gonzalez, M. A. Jantz et al., “Methods for staging non-small cell lung cancer: diagnosis and management of lung cancer, 3rd ed: American college of chest physicians evidence-based clinical practice guidelines,” *Chest*, vol. 143, 2013.
- [57] J. L. Warren, C. N. Klabunde, D. Schrag, P. B. Bach, F. Gerald, and G. Riley, “Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population,” *Medical Care*, vol. 40, 2002.