ARTICLE

Potential Influence on Clinical Trials of Long-Term Survivors of Stage IV Non-small cell Lung Cancer

Jennifer S. Davis*, Erin Prophet*, Ho-Lan Peng, Hwa Young Lee, Rebecca S. S. Tidwell, J. Jack Lee, Anish Thomas, Eva Szabo†, Shine Chang†

See the Notes section for the full list of authors' affiliations.
*Authors contributed equally to this work and are listed in alphabetical order.
†Co-senior authors.
Correspondence to: Jennifer S. Davis, PhD, Department of Epidemiology, The University of Texas MD Anderson Cancer Center, PO Box 301439, Houston, TX 77230-1439 (e-mail: JSDavis@MDAnderson.org).

Abstract

Background: New, effective treatments have resulted in long-term survival for small subgroups of metastatic non-small cell lung cancer (NSCLC) patients. However, knowledge of long-term survivor frequency and characteristics prior to modern therapies is lacking.
Methods: Surveillance Epidemiology and End Results (SEER) patients with stage IV NSCLC diagnosed from 1991 to 2007 and followed through 2012 were dichotomized by survival time into the 10% who lived 21 months or longer (long-term survivors) vs the remaining 90% and compared with participants in a representative clinical trial of molecular profiling and targeted therapies (CUSTOM).
Results: Among the 44 387 SEER patients, the 10% identified as long-term survivors were distinguishable from the remaining 90% by younger age, female sex, Asian race, adenocarcinoma histology, tumor grade, tumor site, and surgery. From 1991–1994 to 2003–2007, median survival increased by 6 months from 30 to 36 months among long-term survivors but by only 1 month from 3 to 4 months among the remaining 90%. Among the 165 participants in the CUSTOM trial, 54% met our SEER criterion of long-term survival by living for 21 months or longer.
Conclusions: Among SEER patients with stage IV NSCLC, long-term survivors had a median survival approximately 10 times that of the remaining 90%. Long-term survivors accounted for more than one-half of the participants in a representative clinical trial. Caution is required when extrapolating the outcomes of participants in clinical trials to patients in routine clinical practice.

The introduction of newer, effective metastatic non-small-cell lung cancer (NSCLC) treatment modalities, including targeted therapies and immunotherapy, has improved survival for patient subgroups (1–3). However, contextual benchmarks of long-term survival rates from late-stage disease prior to newer treatment modalities are lacking. The overall poor survival of late-stage NSCLC patients (5-year survival is 55% when diagnosed early [4] and <15% when diagnosed at stages III and IV [5]) and focus on the overwhelming number with short survival have long overshadowed the few patients who experience long survival, the latter being regarded as rare outliers. Nevertheless, some stage IV patients survive as long as 10 years (6).

Previously identified prognostic factors for NSCLC long-term survival include diagnosis with adenocarcinoma rather than squamous cell carcinoma (SCC), fewer involved lymph nodes, fewer metastatic sites or metastasis limited to the thorax or brain, normal lactate dehydrogenase levels at diagnosis, and

non- or former smoking status (2,7–9). Treatment factors associated with long-term survival include curative surgery, better or longer response to first-line chemotherapy, greater number of chemotherapy agents received, use of maintenance therapy, treatment with epidermal growth factor receptor tyrosine kinase inhibitors or immunotherapy, and better performance status (2,3,7–11). However, variation in long-term survival definitions, differences among patients (eg, stage at diagnosis, histology, treatment), and limited sample sizes restrict interpretation. Reports from single institutions and cohorts receiving specific treatments also constrain comparison (2,11,12). A better understanding of long-term survivors of advanced NSCLC diagnosed and treated prior to current therapies can help improve research methodologies and provide a critical benchmark for measuring patient benefit.

We first evaluated, using data from the Surveillance Epidemiology and End Results (SEER) database, whether NSCLC patients initially diagnosed with metastatic disease who attained long-term survival were distinguishable from those with shorter survival by demographic, tumor, or treatment characteristics, including patients with the two most common NSCLC histologies, SCC and adenocarcinoma (13). Second, we examined how patients surviving at least 5 years differed from other long-term survivors. Third, we evaluated improving survival patterns over time in patients with newly diagnosed stage IV NSCLC. Finally, we examined survival in a lung cancer molecular profiling clinical trial to evaluate its applicability to NSCLC patients in the general population (14).

Methods

SEER Database and Study Population

The SEER database includes information collected from 1973 to 2012 in nine cancer registries (15). The earliest year data with sufficient refinement are available is 1991. To allow at least 60 months of follow-up through 2012, we studied individuals newly diagnosed with stage IV NSCLC between 1991 and 2007 using data from the respiratory dataset (RESPIR).

Definition of Long-Term Survival

We measured survival months from date of cancer diagnosis until death or last follow-up, which was censored. To define long-term survivors, we used two approaches, statistical and clinically practical. The statistical approach employed multivariable Cox models with variable selection and Classification and Regression Tree (CART) analysis seeking factors associated with long-term survivors. The practical approach divided patients by survival time, defining patients in the top decile for longest survival as "long-term survivors" and comparing them with the remaining 90%. Among the long-term survivors, we compared patients by 5-year survival, a common benchmark of treatment success.

Covariates

Diagnosis year was grouped into four periods of approximately equivalent lengths (1991–1994, 1995–1998, 1999–2002, 2003–2007). Histology was summarized as "SCC" (including squamous and epidermoid histology) and "adenocarcinoma" (including adenocarcinoma and bronchioalveolar histology). SEER collected surgical details under different variables for cases diagnosed from 1983 to 1997 and from 1998 onward, summarized as follows:

"None," including no surgical procedure and non-cancer-directed surgery; "Diagnostic," including biopsies, exploratory procedures, and bypass surgeries; "Non-curative," including local tumor destruction and removal of less than one lobe; "Curative," including lobectomy and pneumonectomy; "Surgery of regional/distant site(s)/node(s) only;" "Resection of Lung, not otherwise specified (NOS);" "Surgery, NOS;" and "Unknown." Radiation was summarized as "any," "none," or "unknown."

Differences Between Long-Term and the Remaining 90%

Long-term survivors were compared with the remaining 90% by demographics (diagnosis age, sex, race, year of diagnosis), tumor characteristics (histology, grade, anatomical site, which was further characterized as main bronchus, upper lobe, middle lobe, lower lobe, overlapping lesion, NOS), and treatment (surgery, radiation) using χ^2 and Student t test for categorical and continuous variables, respectively. Median age at diagnosis was compared using a median two-sample test. Age at diagnosis was categorized by quartile for parametric analyses.

CUSTOM Clinical Trial Data

From February 2011 to December 2012, 647 patients were enrolled in the CUSTOM (Molecular Profiling and Targeted Therapies in Advanced Thoracic Malignancies) clinical trial (14) at the National Cancer Institute. Among them, 383 had complete information available. Of those, 165 had newly diagnosed stage IV NSCLC with SCC or adenocarcinoma, forming our analysis cohort.

Survival Analyses

Kaplan-Meier curves and log-rank tests were used for calculating median survival months and conducting unadjusted analyses. To identify long-term survivors statistically, we used two approaches to Cox regression modeling (backwards elimination and stepwise model building [16]) and CART analysis. All approaches included age at diagnosis, sex, race, time period, histology, and disease site as potential measures associated with long-term survivors. Covariate correlations were assessed and the absolute values were less than 0.3. The backwards elimination Cox model was explored with and without interaction terms. For the stepwise model selection, a best subsets regression was employed using Akaike information criterion to identify optimal model size. Finally, we used a multivariable CART analysis stratified by diagnosis time period, with a minimum terminal node size of 1000 patients, combining nodes with similar profiles. Within the most recent period (2003–2007), we implemented an additional multivariable CART analysis with a minimum node size of 200 to evaluate patients potentially receiving modern therapies.

Cox proportional hazards models were also used to evaluate survival, controlling for other factors among long-term survivors only (top 10%). Here, hazard for death among patients with SCC and adenocarcinoma formed hazard ratios comparing risk of death, defined as death from respiratory cancer (codes 22010–22060) (17). Because our test of the proportional hazards assumption was not met, we constructed an extended model with time-dependent covariates (histology*t and histology*$_{\sqrt{}}$t) to estimate hazard ratios for each interval of proportional hazards. We adjusted for secular trends by including grouped "years of diagnosis." Fully adjusted models included age at diagnosis,

sex, race, years of diagnosis, surgery, and radiation. We evaluated the effect of age at diagnosis on the association between histology and survival using the log-likelihood ratio test. *P* values less than .05 based on two-sided tests were considered statistically significant. Analyses were conducted using SAS 9.4 (SAS Institute).

Results

Study Population and Long-Term Survival Definitions

Of 624 997 individuals in the SEER RESPIR file, we analyzed 44 387 stage IV NSCLC patients after omitting those outside the period of interest (52.3%), patients with non-lung/bronchus tumors (3.6%), histology other than SCC or adenocarcinoma (22.3%), missing data for age or survival (0.1%), age younger than 20 years (two individuals, <0.01%), and those initially diagnosed with 0–III or unknown stage NSCLC (14.5%) (Supplementary Figure 1, available online).

Median survival for the study population was 4 months (quartiles: 1, 10), with 2.3% of patients (n = 1014) surviving 5 years or longer. Neither stepwise Cox regression models nor CART analysis revealed a distinct subgroup of long-term survivors across time periods, but both analyses found younger age, female sex, and Asian race associated with increased overall survival (Supplementary Tables 1 and 2, available online). Because both methods found consistent themes but no definable cutoff for long-term survival, we used an arbitrary definition of clinical practicality with patients in the top decile of longest survival (>21 months) considered "long-term survivors," who were compared with the remaining 90% (<21 months survival).

Factors Associated with Long-Term and 5-Year Survival

Among patients diagnosed with stage IV NSCLC, Kaplan-Meier analysis revealed varying survival by histology among long-term survivors (Figure 1). Compared with the remaining 90%, long-term survivors were younger, more likely to be female or Asian, and diagnosed more recently. Their tumors were more likely to be adenocarcinoma, located in the upper lobes, with lower tumor grade, and to have been treated with surgery, but less likely to have been treated with radiation (Supplementary Table 3, available online). Among long-term survivors, those surviving at least 5 years were further distinguishable by younger age at diagnosis, lower tumor grade, and surgical treatment (Supplementary Table S3, available online). Of long-term survivors with SCC, 25.5% lived 5 or more years in contrast to 21.3% of long-term survivors with adenocarcinoma (Supplementary Table 3, bottom right, available online).

Median Survival and 5-Year Survival Over Time

Between long-term survivors and the remaining 90%, an approximately 10-fold difference was found in median survival. To reveal trends in survival over time, we compared median survival of patients diagnosed in the earliest period with those diagnosed in the most recent interval (1991–1994 vs 2003–2007). Median survival increased from 30 to 36 months for long-term survivors and from 3 to 4 months for the remaining 90%. Among long-term survivors, we observed an increase in median survival for those surviving less than 5 years, from 27 to 31 months, but no difference in median survival for those surviving at least 5 years (Table 1). However, the proportion of patients surviving
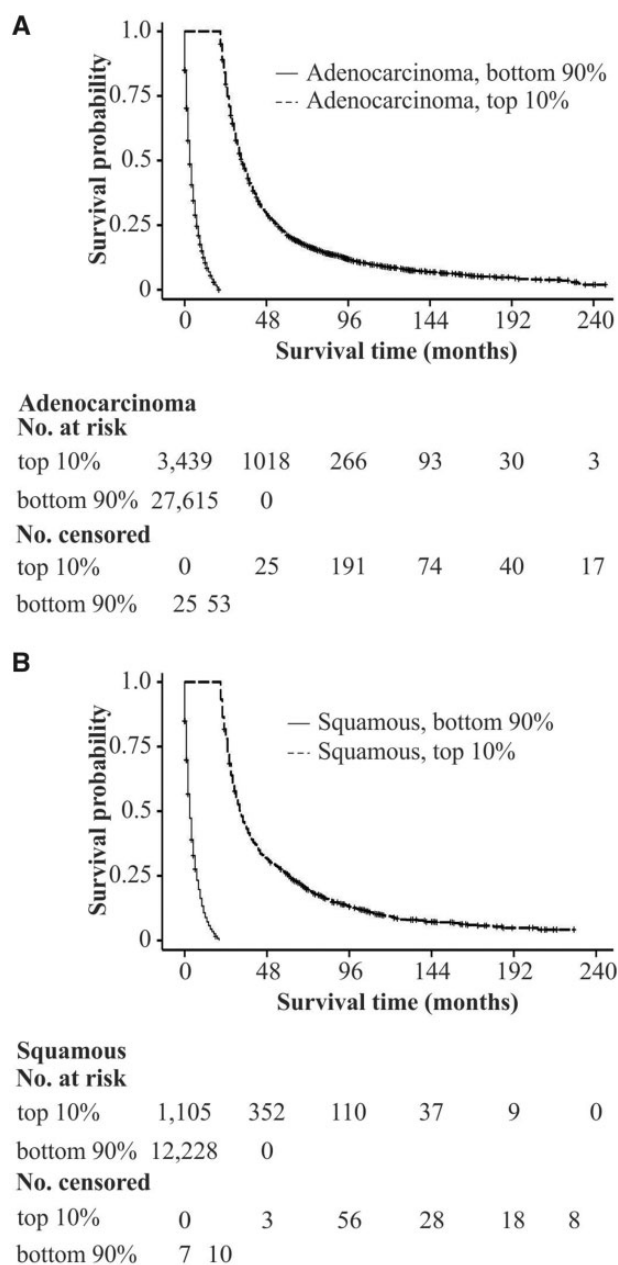

Figure 1. Survival by histology and long-term survival status. Kaplan-Meier curves are presented for survival time (months) stratified by long-term survival status and adenocarcinoma **A**) and squamous **B**) histology. The numbers of patients at risk and censored are aligned with survival time in months indicated in the figure.

at least 5 years doubled over time (1.5 to 3.1%) and for both adenocarcinoma and SCC (Table 1, right side). To provide context for survival changes, we mapped the percentage of patients surviving at least 5 years to milestones in NSCLC detection and treatment (Figure 2), finding that survival for both histologies appears to have increased prior to major therapeutic changes.

Factors Associated with Longer Survival Among Long-Term Survivors

After adjusting for grouped year of diagnosis, sex, race, and surgery, the extended Cox model found that age at diagnosis and

Table 1. Median survival months and 5-year survival have improved over time for stage IV NSCLC

Time period of diagnosis	Total No. (%) n = 44 387 (100)	Median survival in months by patient group					No. (%) of all stage IV patients surviving at least 5 years by histology*		
		Overall n = 44 387	Remaining 90%† n = 39 843	Long-term survivors‡ n = 4544	Long-term survivors		Total No. (%)‡	Squamous No. (%)‡	Adenocarcinoma No. (%)‡
					<5-year survival n = 3530	≥5-year survival n = 1014			
1991–1994	9444 (21.3)	4‖	3‖	30§	27‖	103	139 (1.5)‖	40 (1.2)‖	99 (1.6)‖
1995–1998	9852 (22.2)	4	3	32	28	99	188 (1.9)	53 (1.8)	135 (2.0)
1999–2002	10 469 (23.6)	4	3	33	29	106	238 (2.3)	76 (2.5)	162 (2.2)
2003–2007	14 622 (32.9)	5‖	4‖	36§	31‖	99	449 (3.1)‖	113 (2.8)‖	336 (3.2)‖

*Percent shown (%) indicates percentage of all patients within that group who survived at least 5 years.
†Remaining 90% are patients surviving less than 21 months (90% of patients). NSCLC = non-small cell lung cancer.
‡Long-term survivors are patients surviving at least 21 months (10% of patients).
§P < .01 between earliest and most recent time periods.
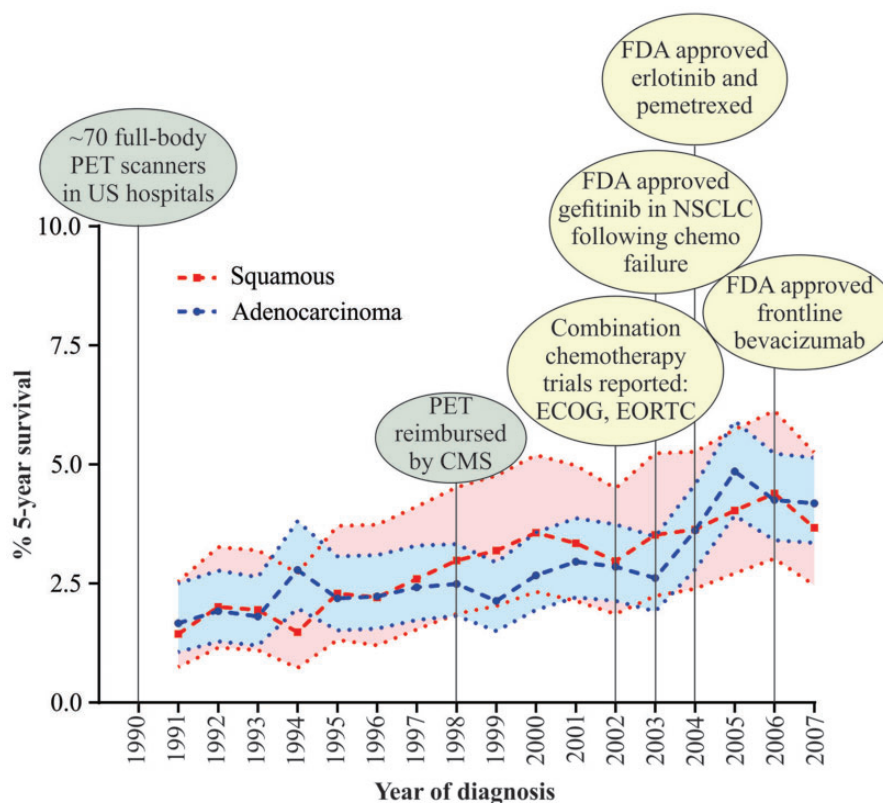‖P < .001 between earliest and most recent time periods.

Figure 2. Five-year survival has increased over time for both squamous and adenocarcinoma histology. Percent 5-year cancer survival and 95% confidence intervals are plotted from Surveillance Epidemiology and End Results data for cases diagnosed between 1991 and 2007 with follow-up through 2012. Milestones in non-small cell lung cancer (NSCLC) treatment have been overlaid on the graph to provide treatment context. CMS = Center for Medicare and Medicaid Services; ECOG = Eastern Cooperative Oncology Group; EORTC = European Organisation for Research and Treatment of Cancer; FDA = Food and Drug Administration; PET = positron emission tomography; US = United States.

histology interacted to influence survival—although with borderline significance (P = .06) (Supplementary Table 4, available online). Specifically, among long-term survivors, we observed periods when the likelihood of death for patients with SCC was lower than that for patients with adenocarcinoma, and this depended on age at diagnosis (Supplementary Table 4, available online). This model also revealed patients diagnosed in the most recent period, women, Asian/Pacific Islanders, and those

receiving diagnostic or noncurative surgery had a lower likelihood of death (Supplementary Table 4, available online).

Long-Term Survivors in a Lung Cancer Molecular Profiling Clinical Trial

To understand whether a disproportionate number of long-term survivors have been enrolled as clinical trial subjects, we

Table 2. Comparison of CUSTOM patient characteristics by 21-month survival*

Variable of Interest	Total	Short-term survival (<21 months)	Long-term survival (≥21 months)	P
No. (row %)	165	76 (46.1)	89 (53.9)	
Age at diagnosis, mean (SD), y	59.8 (12.1)	62.2 (12.9)	57.7 (11.0)	.02
Sex, n (column %)				.23
Male	72 (43.6)	39 (51.3)	35 (39.3)	
Female	93 (56.4)	37 (48.7)	54 (60.7)	
Race				.42
Caucasian	122 (73.9)	57 (75.0)	65 (73.0)	
African American	13 (7.9)	8 (10.5)	5 (5.6)	
Asian	24 (14.6)	8 (10.5)	16 (18.0)	
Other	6 (3.6)	3 (4.0)	3 (3.4)	
Smoking status				.71
Never	72 (43.6)	32 (42.1)	40 (44.9)	
Ever	93 (56.4)	44 (57.9)	49 (55.1)	
Median survival, mo	26.8	13.1	38.0	
No. censored (column %)	53 (32.1)	14 (18.4)	39 (43.8)	
Months from dx to enrollment				
Mean (SD)	12.8 (15.3)	4.3 (4.5)	20.1 (17.4)	<.0001
Median	8.8	2.0	17.5	
Range	0.03–109.0	0.03–17.2	0.3–109.0	
Histology				
Squamous	17 (10.3)	11 (14.5)	6 (6.7)	.10
Adenocarcinoma	148 (89.7)	65 (85.5)	83 (93.3)	
No. of prior lines of therapy				
Mean (SD)	1.4 (1.5)	0.9 (1.0)	1.9 (1.8)	<.0001
Median	1.0	1.0	1.0	
Range	0–7	0–4	0–7	
Categorical no. of prior therapies				.0003
0	49 (29.7)	31 (40.8)	18 (20.2)	
1	58 (35.15)	30 (39.5)	28 (31.5)	
2+	58 (35.15)	15 (19.7)	43 (48.3)	
Performance status at enrollment				.08
0	28 (17.0)	8 (10.5)	20 (22.5)	
1	123 (74.6)	61 (80.3)	62 (69.7)	
2	12 (7.3)	5 (6.6)	7 (7.9)	
3	2 (1.2)	2 (2.6)	0 (0)	
EGFR mutation	131			.01
WT	92 (70.2)	45 (81.8)	47 (61.8)	
Mutated	39 (29.8)	10 (18.2)	29 (38.2)	
KRAS	136			.86
WT	106 (77.9)	44 (77.2)	62 (78.5)	
Mutated	30 (22.1)	13 (22.8)	17 (21.5)	
ALK translocation	132			.96
Absent	116 (87.9)	50 (87.7)	66 (88.0)	
Present	16 (12.1)	7 (12.3)	9 (12.0)	

*CUSTOM = Molecular Profiling and Targeted Therapies in Advanced Thoracic Malignancies; dx = diagnosis; SD = standard deviation; WT = wild type.

evaluated data from the recent CUSTOM trial (14). Describing long-term survival using the 21-month cutoff revealed nearly 54% of stage IV NSCLC trial participants were long-term survivors (Table 2). Consistent with our SEER analysis, long-term survivors were statistically significantly younger at diagnosis and many had survived longer than expected at trial enrollment, with a median of 17.5 months from diagnosis to enrollment for long-term survivors compared with 2 months for the remainder. Mean prior lines of therapy was also different: 1.9 for long-term survivors and 0.9 for the remainder. In addition, statistically significantly more patients with *EGFR*-mutated tumors were long-term survivors (Table 2).

Of stage IV NSCLC patients enrolled in the trial, 35% had two or more prior lines of therapy, with a median of 19.3 months between diagnosis and study enrollment (Table 3). Comparing overall survival by prior lines of therapy (zero, one, or two or more) revealed better survival for patients with two or more lines of prior therapy (Figure 3). It is important to note that follow-up was approximately 30 months for patients in this trial, resulting in large numbers of censored patients.

Discussion

The rapidly changing lung cancer treatment landscape, with widespread use of targeted and immune therapies, has resulted in statistically significant median survival improvements, raising the prospect of long-term survival for some individuals diagnosed with metastatic disease (11,12). To evaluate this potential impact on the larger NSCLC population, we must

examine survival course and long-term survivor characteristics from eras prior to such treatments. In this SEER-based study, we identified and characterized a small but statistically significant number of long-term survivors. The percentage of long-term survivors has steadily increased over time both for adenocarcinoma and SCC; however, we found long-term survivors in the early 1990s, before development and widespread use of modern chemotherapy regimens, targeted therapy, and immunotherapy.
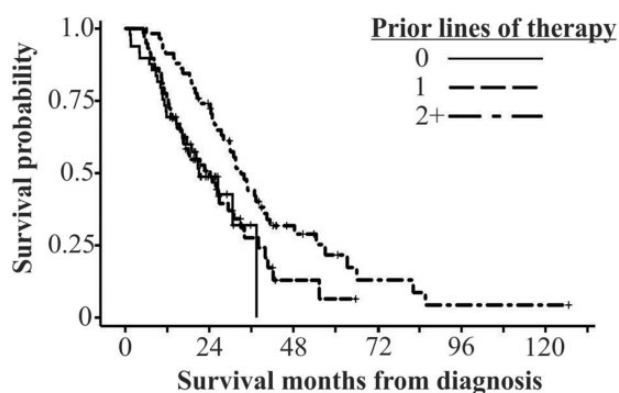
The top 10% had longer survival than the remainder (by definition); however, the range of their substantially longer survival indicates considerable heterogeneity, raising the possibility of identifiable subgroups with widely divergent prognoses. Previously reported demographic profiles (18) of long-term

Figure 3. Survival of CUSTOM (Molecular Profiling and Targeted Therapies in Advanced Thoracic Malignancies) trial participants by prior lines of therapy. Kaplan-Meier curves are presented for survival time (months) from diagnosis stratified by lines of therapy received between diagnosis and study enrollment. Patients with two or more lines of prior therapy survive longer. The numbers of patients at risk and censored are aligned with survival time in months indicated in the figure.

survivors mirror that of patients with EGFR-activating mutations (19). Our analysis of CUSTOM clinical trial data suggests a higher proportion of EGFR-mutated tumors diagnosed among long-term survivors in the trial. Improvements in median and 5-year survival in adenocarcinoma patients are also consistent with the hypothesis that EGFR-mutation-driven NSCLC is associated with long-term survival, possibly reflecting Food and Drug Administration approval of the first EGFR-targeted drug in 2003. However, a single institutional experience with epidermal growth factor receptor tyrosine kinase inhibitors for EGFR-mutated metastatic NSCLC reported substantially higher 5-year survival (14.6%) than we observed either overall or for adenocarcinoma (2). Because approximately 15% of adenocarcinomas are expected to have EGFR mutations (20), the combination of EGFR mutation and targeted treatment cannot be exclusively responsible for the observed survival gain. Therapies targeting other molecular aberrations were introduced outside our analytic timeframe, precluding their contribution to these survival trends.

As with patients diagnosed with stage IV adenocarcinomas, there is a subgroup of SCC patients who survive at least 5 years, and this subgroup size has also increased. This gain cannot be attributed to therapeutic advances beyond modern chemotherapy regimens, which do not provide long-term disease control (21). A recent single-institution report suggests that molecular heterogeneity may account for survival differences among stage IV SCC patients, with median survival of 8.6–21.3 months, depending on mutation profile (22). Such findings support the hypothesis that specific molecular subtypes have a prolonged natural history, irrespective of targeted treatment, and deserve future study.

The relationship of demographic and treatment characteristics to long-term survival is complex. For instance, although adenocarcinoma histology portended higher likelihood of surviving at least 21 months, among long-term survivors those with SCC were more likely to survive 5 years. The molecular basis for this is unclear (22) and bears further investigation. Similarly, we found both curative and noncurative surgery associated with longer survival, though surgery is not routinely used in this setting. This is consistent with a prior report of longer survival for stage IV NSCLC patients undergoing surgery as part of multimodality treatment (23), an association that may reflect treatment bias. Other treatments previously associated with survival, such as chemotherapy, were not captured in SEER, precluding their assessment (11).

We observed increasing 5-year survivor numbers over time, a trend that continues, with 4.9% of stage IV NSCLC patients

Table 3. CUSTOM stage IV NSCLC patients compared by lines of prior therapy

Variable	Total	Lines of prior therapy		
		0	1	2+
No. (row %)	165	49 (29.7)	58 (35.15)	58 (35.15)
Months from dx to enrollment				
Mean	12.8	2.0	10.3	24.5
Median	8.8	1.0	8.7	19.3
Range	0.03–109.0	0.03–18.6	0.8–52.6	4.2–109.0
Median survival (95% CI*), mo	26.8 (24.3 to 31.5)	21.1 (15.7 to 37.5)	24.3 (15.9 to 31.0)	33.0 (27.4 39.6)
No. surviving 5 years (column %)	7 (4.2)	0 (0)	1 (1.7)	6 (4.2)
No. censored (column %)	53	23 (46.9)	16 (27.6)	14 (24.1)
Maximum follow-up from diagnosis, mo	126.6	35.8	65.8	126.6
Maximum follow-up from enrollment, mo	30.9	27.9	29.2	30.9

*CI = confidence interval; CUSTOM = Molecular Profiling and Targeted Therapies in Advanced Thoracic Malignancies; dx = diagnosis; NSCLC = non-small cell lung cancer.

surviving 5 years in a SEER analysis from 2006 to 2012 (24). Although some improvement may be related to stage migration (the "Will Rogers phenomenon") (25), improvements in systemic treatments and supportive therapies may also contribute. Despite this progress, we found the majority of patients with stage IV NSCLC had a median survival of 5 months, even as recently as 2003–2007, a sobering statistic that is substantially lower than the 8-to-12-month median survival reported in phase III clinical trials conducted in the same time period and in a systematic review of these trials (21,26,27).

This difference suggests that clinical trials disproportionately enroll participants with better prognosis than the majority (28). Although outside the timeframe of our SEER analysis, the CUSTOM trial, with a median of 8.8 months between participant diagnosis and study enrollment, supports this proposition. In the recently reported follow-up of the CA209-003 phase I nivolumab study, 16% of trial participants with lung adenocarcinoma or SCC survived at least 5-years from trial enrollment. Similar to the CUSTOM trial, the majority of participants had two or more prior lines of therapy (87.6% and 81.2% of all treated patients and those with ≥5 year survival, respectively) (3). It is highly likely that patients with poor prognosis, treatment-refractory disease are underrepresented in clinical trials, due in large part to strict eligibility criteria and, often, inability to travel.

A variety of factors contributed to our finding less robust long-term survival than previously reported (2,7,10–12). First, we limited the focus to patients with adenocarcinoma or SCC, thereby excluding less common histologies or inadequately characterized cases. Second, we analyzed only patients with newly diagnosed stage IV disease, excluding those who relapsed from early-stage disease whose course may differ. Also, we excluded cases with pleural effusion, which was categorized as stage IIIB disease during the timeframe of our study (29). Of critical significance, our analysis included all patients in the SEER database rather than patients presenting to a single hospital or outpatient specialty department, thus avoiding possible selection bias. That we found a 4-month median survival for 90% of the patient population supports the hypothesis that sicker patients and those who never achieve treatment referral are filtered out of clinical studies.

This SEER-based study's strengths include its population-level scope, long follow-up, large numbers, and potential generalizability. Nevertheless, the 9 SEER registries available in 1991 did not draw data from all 50 states. Important SEER limitations include lack of information on systemic treatment, comorbidities, performance status, smoking status, and occupational and socioeconomic data. Because tumor molecular data are not recorded in SEER, the effect of tumor genetics and targeted treatments could not be evaluated. Although the number of CUSTOM trial patients was relatively small, their better demographic and molecular characterization provides hints regarding the contribution of these factors to long-term survival.

Our results have several important implications. The existence of long-term survivors diagnosed with metastatic NSCLC demonstrates heterogeneity in disease pathogenesis and outcomes. The identification of such survivors prior to modern treatments suggests that some patients experience prolonged survival independent of treatment. Examination of patient characteristics and survival in a recent clinical trial shows a potential bias in clinical trial enrollment, with overrepresentation of long-term survivors.

Understanding tumor biology and other characteristics of long-term survivors may reveal new therapeutic avenues for the general population of NSCLC patients. Alternatively, prospectively identifying patients with longer-term survival potential may allow more tailored therapeutic approaches with improved outcomes. Studies of this unique patient subgroup focusing on host factors and molecular characteristics influencing survival are urgently needed.

Funding

Notes

Affiliations of authors: Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX (JSD, EP, HYL, SC); Department of Management, Policy, and Community Health, The University of Texas School of Public Health, Houston, TX (HLP); Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX (RSST, JJL); Center for Cancer Research (AT), and Division of Cancer Prevention (ES), National Cancer Institute, Bethesda, MD.

No authors have any conflicts of interest to report.

JSD, EP, SC, and ES contributed to conceptualization of the research question and approach. HYL, HLP, EP, JSD, ES, AT, RSST, and JJL contributed to data curation, methodology, and analysis. JSD, EP, HYL, HLP, RSST, AT, JJL, ES, and SC interpreted results, contributed to the development of the manuscript, and reviewed, edited, and approved the final manuscript.

References

1. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med*. 2015;373(17): 1627–1639.
2. Lin JJ, Cardarella S, Lydon CA, et al. Five-year survival in *EGFR*-mutant metastatic lung adenocarcinoma treated with *EGFR*-TKIs. *J Thorac Oncol*. 2016;11(4): 556–565.
3. Gettinger S, Horn L, Jackman D, et al. Five-year follow-up of nivolumab in previously treated advanced non-small-cell lung cancer: results from the CA209-003 study. *J Clin Oncol*. 2018;36(17):1675–1684.
4. American Cancer Society. *Cancer Facts & Figures 2016*. Atlanta: American Cancer Society; 2016.
5. American Cancer Society. *Non-Small Cell Lung Cancer Survival Rates, by Stage*. http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/ non-small-cell-lung-cancer-survival-rates. Accessed August 2, 2016.
6. Frenkel M, Gross S, Popper Giveon A, et al. Living outliers: experiences, insights and narratives of exceptional survivors of incurable cancer. *Future Oncol*. 2015;11(12):1741–1749.
7. Kaira K, Takahashi T, Murakami H, et al. Long-term survivors of more than 5 years in advanced non-small cell lung cancer. *Lung Cancer*. 2010;67(1): 120–123.
8. Chen Y-Z, Feng X-B, Li Z-D, et al. Clinical study on long-term overall survival of advanced non-small-cell lung cancer patients treated with Chinese medicine and western medicine. *Chin J Integr Med*. 2014;20(3):179–183.
9. Okamoto T, Maruyama R, Shoji F, et al. Long-term survivors in stage IV non-small cell lung cancer. *Lung Cancer*. 2005;47(1):85–91.
10. Van Damme V, Govaerts E, Nackaerts K, et al. Clinical factors predictive of long-term survival in advanced non-small cell lung cancer. *Lung Cancer*. 2013; 79(1):73–76.
11. Giroux Leprieur E, Lavole A, Ruppert A-M, et al. Factors associated with long-term survival of patients with advanced non-small cell lung cancer. *Respirology*. 2012;17(1):134–142.

12. Satoh H, Ishikawa H, Ohara G, et al. Long-term survivors after chemotherapy in advanced non-small cell lung cancer. *Anticancer Res.* 2007;27(6C): 4457–4460.
13. National Cancer Institute. *Surveillance, Epidemiology and End Results Program.* http://seer.cancer.gov. Accessed March 03, 2017.
14. Lopez-Chavez A, Thomas A, Rajan A, et al. Molecular profiling and targeted therapy for advanced thoracic malignancies: a biomarker-derived, multiarm, multihistology phase II basket trial. *J Clin Oncol.* 2015;33(9):1000–1007.
15. Hankey BF, Ries LA, Edwards BK. The Surveillance, Epidemiology, and End Results program: a national resource. *Cancer Epidemiol Biomarkers Prev.* 1999; 8(12):1117–1121.
16. Shtatland ES, Kleinman K, Cain EM. Model building in PROC PHREG with automatic variable selection and information criteria. In: *SUGI 30. Philadelphia, Pennsylvania*, 2005. Abstract 30.
17. National Cancer Institute. *SEER Cause of Death Recode 1969+ (9/17/2004).* https://seer.cancer.gov/codrecode/1969+_d09172004/index.html. Accessed February 13, 2018.
18. Ries LAG, Young JL, Keel GE, Eisner MP, Lin YD, Horner M-J (eds). SEER Survival Monograph: Cancer Survival Among Adults: U.S. SEER Program, 1988–2001, Patient and Tumor Characteristics. Bethesda, MD: National Cancer Institute, SEER Program; 2007.
19. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst.* 2005;97(5):339–346.
20. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014;511(7511):543–550.
21. Schiller JH, Harrington D, Belani CP, et al. Comparison of four chemotherapy regimens for advanced non–small-cell lung cancer. *N Engl J Med.* 2002;346(2): 92–98.
22. Paik PK, Shen R, Won H, et al. Next-generation sequencing of stage IV squamous cell lung cancers reveals an association of PI3K aberrations and evidence of clonal heterogeneity in patients with brain metastases. *Cancer Discov.* 2015;5(6):610–621.
23. David EA, Canter RJ, Chen Y, et al. Surgical management of advanced non-small cell lung cancer is decreasing but is associated with improved survival. *Ann Thorac Surg.* 2016;102(4):1101–1109.
24. Howlader N, Noone AM, Krapcho M, et al. *SEER Cancer Statistics Review, 1975-2013.* Bethesda, MD: National Cancer Institute; 2016.
25. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med.* 1985;312(25): 1604–1608.
26. Sandler A, Gray R, Perry MC, et al. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. *N Engl J Med.* 2006;355(24): 2542–2550.
27. Kiely BE, Alam M, Blinman P, et al. Estimating typical, best-case and worst-case life expectancy scenarios for patients starting chemotherapy for advanced non-small-cell lung cancer: a systematic review of contemporary randomized trials. *Lung Cancer.* 2012;77(3):537–544.
28. Van Spall HG, Toren A, Kiss A, et al. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA.* 2007;297(11):1233–1240.
29. Postmus PE, Brambilla E, Chansky K, et al. The IASLC Lung Cancer Staging Project: proposals for revision of the M descriptors in the forthcoming (seventh) edition of the TNM classification of lung cancer. *J Thorac Oncol.* 2007; 2(8):686–693.